

Systems Biology for DOE Energy Missions: Bioenergy

Bioenergy Research Centers

Great Lakes Bioenergy Research Center (GLBRC)

1

Discovery of Novel Genes Regulating Polysaccharide Biosynthesis and Secretion in Plants

Michael Held^{1,2*} (heldm@msu.edu), Starla Zemelis,^{1,2} Saunia Withers,^{1,3} Curtis Wilkerson,^{1,3} and **Federica Brandizzi**^{1,2,3}

¹Great Lakes Bioenergy Research Center, ²MSU-DOE Plant Research Laboratory, and ³Dept. of Plant Biology, Michigan State University, East Lansing

<http://www.msu.edu/~brandizz/index.html>

Project Goals: The long term goals of our project are to identify novel factors that facilitate the biosynthesis and/or secretion of hemicellulosic and pectic cell wall polymers. These advances will lay groundwork for designing new strategies for improving plant biomass yield and digestibility.

During cotton ovule development there is a rapid burst in the secretion and deposition of non-cellulosic cell wall polymers between 4 and 6 days post-anthesis (dpa). Within this developmental time period, the cotton Golgi swell and produce a large number of secretory vesicles to accommodate the increase in secretion. Differential proteomics of cotton Golgi at 4 and 6 dpa has generated an extensive list of ~800 proteins that increase in abundance and are hypothetically involved in non-cellulosic cell wall biosynthesis and/or secretion. While many of these 6dpa abundant proteins are already known to be involved in polysaccharide biosynthesis and secretion (i.e. nucleotide sugar interconverting enzymes, glycosyltransferases, and transport related proteins), most have been annotated as genes of unknown function. To identify novel proteins which facilitate and/or modulate polysaccharide transport, we have selected a subset of these 6dpa abundant unknown proteins. Currently, we have cloned seventeen *Arabidopsis* orthologs of these cotton candidates as cerulean cyan fluorescent protein fusions; fifteen of which have been localized to endomembranes by transient expression analyses followed by live cell imaging. t-DNA insertion lines have been identified for many of these candidates, and preliminary analyses show several lines having altered cell wall and/or growth phenotypes. This research puts sound foundations toward establishing the nature of factors that not only control cell wall biosynthesis, but also the

traffic of cell wall components through and from the Golgi apparatus.

2

Utilizing Biochemical Adaptations of Plants and Next-Generation Sequencing Instruments to Discover Enzymes and Transcription Factors Involved in Plant Cell Wall Biosynthesis

Jacob Kruger Jensen,^{1,2} Jean-Christophe Cocuron,^{1,2} and Curtis G. Wilkerson^{1,2,3*} (Wilk13@msu.edu)

¹DOE Great Lakes Bioenergy Research Center, ²Dept. of Plant Biology, ³Dept. of Biochemistry and Molecular Biology, Michigan State University, East Lansing

Project Goals: We are attempting to define the proteins required by plants to synthesize arabinoxylan. We are also attempting to discover the transcription factors that regulate the synthesis of these proteins. The ultimate aim is to alter the ratio of hexose to pentose sugars in the plant cell wall to increase the production of biofuels from these feedstocks by fermentation.

One way to discover enzymes and regulatory proteins involved in a biological pathway is to examine the mRNAs of cells actively engaged in the pathway of interest, ideally when the pathway is a major activity for the tissue. There are many such plant tissues that produce very large amounts of product from a single biochemical pathway. As an example, the mucilaginous layer of psyllium (*Plantago ovata* Forsk) seed contains about 60% arabinoxylan by weight (Fischer et al. 2004) and so is a good tissue to use for transcriptional profiling to discover genes involved in arabinoxylan biosynthesis. Similarly, the endosperm tissue of Fenugreek seeds contains 80% galactomannan by weight and hence provides a system to study mannan and galactomannan biosynthesis. The recent development of DNA sequencing instruments that can produce millions of sequences quickly and at moderate cost now allows us to exploit such systems to discover both biosynthetic enzymes from specific pathways and the transcription factors that control their expression. We are interested in using such a strategy to gain a greater understanding of plant cell wall biosynthesis.

Our initial work has been to gain a better understanding of the enzymes required for arabinoxylan biosynthesis using the psyllium system. Arabinoxylan is a major component of grass cell walls and as grasses are likely to be important as bioenergy crops such work could have a major impact on bioenergy research. We have successfully generated 4 cDNA libraries from psyllium mucilaginous tissue at 6, 8, 10 and 12

days post anthesis (DPA) and have obtained over 850,000 DNA sequences using the Roche GS-FLX sequencer. We have developed an analysis pipeline and query software to allow us to cluster, annotate and search these large datasets. An examination of these sequences using our software revealed that enzymes involved in the biosynthesis of UDP-xylose were highly represented in those cDNA libraries suggesting that these libraries likely are enriched in genes involved in arabinoxylan biosynthesis and its regulation. We find that homologues of the putative xylan synthases IRX10 and IRX10-like are very abundant these tissues. We find at least six genes encoding proteins in the glycosyl transferase family 61 at high abundance. We also find other genes that have been implicated in secondary cell wall biosynthesis that could be involved in xylan biosynthesis as well. One of these genes is related to the *Arabidopsis* gene At3g50220. The expression of this gene is highly correlated with the expression of IRX10 in *Arabidopsis* and so may represent a component of the xylan synthase. Since the psyllium mucilaginous layer is synthesizing almost exclusively arabinoxylan it is likely that the subset of genes present in this tissue compared to the larger set of genes found in cambial tissues defines a minimal set of genes required to synthesize arabinoxylan. We have also identified a subset of transcription factors known to be up regulated during secondary cell wall biosynthesis. Since psyllium is making only arabinoxylan and not a secondary cell wall likely the transcription factors we have identified are involved in the regulation of arabinoxylan biosynthetic enzymes. We are currently expressing the psyllium versions of these genes in various heterologous systems to establish the function of these proteins.

This research was funded by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494).

3 Generation and Analysis of Transgenic Poplars with Altered Wall Compositions

Jae-Heung Ko, Won-Chan Kim, Hyun-Tae Kim, and
Kyung-Hwan Han* (hanky@msu.edu)

Dept. of Forestry and Great Lakes Bioenergy Research
Center, Michigan State University, East Lansing

Project Goals: (1) Identify the key regulatory genes that control xylogenesis and secondary wall biosynthesis; (2) Characterize the biological functions of cell wall biosynthesis candidate genes identified in model plants by Thrust I collaborators; and (3) Develop biotechnology tools that allow us to create woody biomass feedstocks with altered cell walls that are more easily digestible thereby releasing higher quantities of fermentable sugars.

Wood is gaining popularity as a source of fermentable sugars for liquid fuel production. However, our current knowledge on the genetic control of woody biomass formation is limited. Secondary wall of wood consists of a complex mixture

of cellulose, hemicellulose, and lignin. Proportional variability within the mixture of the three major components varies depending on the species of feedstock used, growing site, climate, age and the part of the plant harvested. The essentially uncontrolled variability of biomass properties presents process design and operating challenges for the production of bioenergy from woody feedstocks. Better understanding of the molecular mechanisms underlying its biosynthesis will help us develop biotechnological means to genetically control key pathways that determine the quantity and quality of the biomass.

In an effort to identify the transcriptional regulatory network controlling the biosynthetic process, we developed an experimental system that induces ectopic development of secondary wall in *Arabidopsis thaliana*. Using this system, we carried out Affymetrix GeneChip and Illumina Digital Gene Expression analyses to identify a battery of genes differentially expressed during secondary wall biosynthesis. These analyses allowed us to identify a group of transcription factors whose expression is coincided or preceded with the induction of secondary wall biosynthetic genes. The list includes a novel transcription factor AtC3H14 that could activate the transcription of all of the secondary wall biosynthesis genes tested, suggesting its potential role as another master regulator of secondary wall biosynthesis. Based on these results, we derived a tentative hierarchical transcriptional regulatory network leading to biosynthesis of secondary wall components. In order to confirm the relationship between transcription factors and their target genes, we are using both *in vivo* transcriptional activation assay and electrophoretic mobility shift assay. The current study tested our hypothesis that the selected transcription factors are responsible for the activation of the individual genes involved in the biosynthesis of secondary wall, and generated additional testable hypotheses. This presentation will describe (1) our strategy to identify transcription factors regulating secondary wall biosynthesis by using inducible secondary wall thickening system and time-course whole genome transcriptome profiling and (2) functional characterization of selected candidate genes in the network.

Utility promoters with 'freedom-to-operate' constitute a key enabling tool for biotechnological improvement of bioenergy crops. In this project, we are developing strong utility promoters that can drive target gene in a developing xylem-specific manner. We have obtained tissue-specific transcriptome profiles in poplar stems and identified candidate promoters. These promoters were fused to GUS reporter gene and expressed in transgenic poplars, which confirm their tissue-specific expression. We are testing the utility of these promoters with an anthocyanin biosynthesis gene in transgenic poplars.

4

Transgenic Poplars with Altered Lignins for Improved Biomass Pretreatment and Saccharification

Hoon Kim^{1*} (hoonkim@wisc.edu), Fachuang Lu,¹ Ali Azarpira,¹ Jorge Rencoret,¹ **John Ralph**,¹ Bruce E. Dale,² Curtis G. Wilkerson,² Nick Santoro,² Shane Cantu,² Cliff Foster,² Rebecca Van Acker,³ Bart Ivens,³ Wout Boerjan,³ and Shawn D. Mansfield⁴

¹DOE Great Lakes Bioenergy Research Center and Dept. of Biochemistry, University of Wisconsin, Madison; ²DOE Great Lakes Bioenergy Research Center and Michigan State University, East Lansing; ³Dept. of Plant Systems Biology, Flanders Institute for Biotechnology (VIB), and Dept. of Plant Biotechnology and Genetics, Ghent University, Gent, Belgium; and ⁴Dept. of Wood Science, University of British Columbia, Vancouver, Canada

Project Goals: The project is aimed at reducing plant cell wall recalcitrance toward enzymatic saccharification by altering lignin composition and structure to allow pretreatment methods to be more efficient.

Introduction

Several approaches are being targeted to explore biomass crop improvement for more efficient conversion (GLBRC Roadmap Goal to “Apply cutting-edge research to help create a new generation of sustainable bioenergy feedstocks, processing technologies and fuels”). One approach, recognizing that lignin is a major factor in plant cell wall (CW) recalcitrance to breakdown (of the polysaccharides to simple sugars), capitalizes on mechanistic insight gained from examining lignin-pathway transgenics in extensive collaborative studies. It is now well recognized that massive compositional changes can be achieved, particularly by misregulation of the various hydroxylase genes.^{1,2} Wild-type poplar has a guaiacyl-syringyl lignin, i.e., one comprised of guaiacyl (G) and syringyl (S) units in comparable amounts, but with only traces of *p*-hydroxyphenyl (H) units; these units are derived from coniferyl, sinapyl, and *p*-coumaryl alcohols respectively. Downregulation of C3H produces lignins rich in the normally minor H units; downregulation of F5H produces G-rich lignins, and upregulation of F5H produces S-rich lignins. It has also become clear that monomer-substitution can occur. For example, COMT-deficient plants incorporate 5-hydroxyconiferyl alcohol into their lignins (replacing some of the sinapyl alcohol), CAD-deficient plants incorporate more hydroxycinnamaldehydes, and recent evidence is that CCR-deficient plants incorporate the hydroxycinnamic acids themselves into the polymer.^{3,4} Thus lignification is considered to be particularly metabolically plastic. Altering the lignin composition/structure, allowing it to be more readily freed (by pretreatments) from the polysaccharide components, can provide enormous energy savings for biomass conversion. Transgenics are therefore being examined for their improved pulping potential, and for improved digestibility, directly or after pretreatment.

F5H-upregulated Poplar

A lignin compositional change particularly targeted for improved alkaline pulping also results in a significant improvement in CW digestibility. F5H upregulation, driven by an appropriate promoter (C4H), produces lignins derived primarily from sinapyl alcohol; the Mansfield group has generated plants strikingly rich in syringyl units.⁵ As a result, the lignin chains are rather homogeneous, being composed essentially only from one resinol unit (from initial dimerization) and β -ether-linked units. The lignins also have only a low degree of polymerization (and are therefore low-molecular-weight). These and other factors result in cell walls that saccharify more efficiently following acidic or basic pretreatment methods⁶ – see Figure 1. We are currently evaluating the effects of the Ammonia Fiber Expansion (AFEX) pretreatment.

CCR-deficient Poplar

CCR-downregulated poplars have slightly lower lignin levels and incorporate ferulic acid into the polymer.^{3,4} CCR-deficient poplar materials saccharify particularly efficiently after mild basic pretreatments – Figure 2. AFEX pretreatment is again being examined.

Conclusions

Already, results from these transgenics (and others), targeted because their lignins were expected to have ‘favorable pretreatment properties’, suggest that altering lignification in biomass crops can significantly lower recalcitrance to pretreatment and saccharification, providing improved bioconversion efficiencies.

References

1. J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P.F. Schatz, J.M. Marita, R.D. Hatfield, S.A. Ralph, J.H. Christensen and W. Boerjan. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochemistry Reviews*, 2004, 3, 29-60.
2. W. Boerjan, J. Ralph and M. Baucher. Lignin biosynthesis. *Annual Reviews in Plant Biology*, 2003, 54, 519-549.
3. J. Ralph, H. Kim, F. Lu, J.H. Grabber, J.-C. Leplé, J. Berrio-Sierra, M. Mir Derikvand, L. Jouanin, W. Boerjan and C. Lapierre. Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl-CoA reductase deficiency). *The Plant Journal*, 2008, 53, 368-379.
4. J.-C. Leplé, R. Dauwe, K. Morreel, V. Storme, C. Lapierre, B. Pollet, A. Naumann, Gilles, K.-Y. Kang, H. Kim, K. Ruel, A. Lefebvre, J.-P. Josseleau, J. Grima-Pettenati, R. De Rycke, S. Andersson-Gunnerås, A. Erban, I. Fehrle, M. Petit-Conil, J. Kopka, A. Polle, E. Messens, B. Sundberg, S.D. Mansfield, J. Ralph, G. Pilate and W. Boerjan. Downregulation of cinnamoyl coenzyme A reductase in poplar; multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell*, 2007, 19, 3669-3691.
5. J.J. Stewart, T. Akiyama, C.C.S. Chapple, J. Ralph and S.D. Mansfield. The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiology*, 2009, 150, 621-635.
6. S.K. Huntley, D. Ellis, M. Gilbert, C. Chapple and S.D. Mansfield. Significant increases in pulping efficiency in

C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *Journal of Agricultural and Food Chemistry*, 2003, 51, 6178-6183.

This research was funded in part by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494).

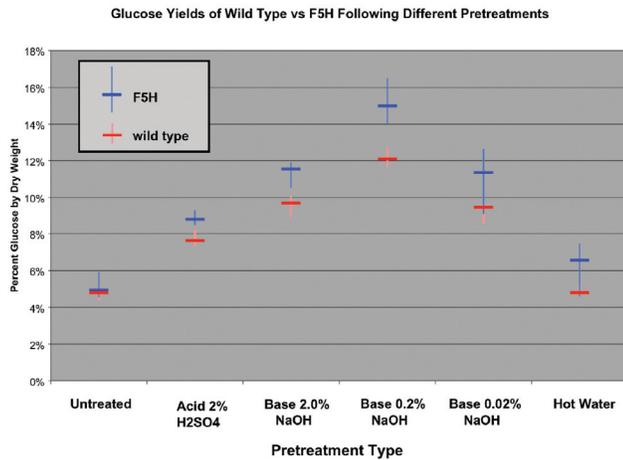


Figure 1. Comparison of saponification glucose yields following various pretreatments; except for untreated material, the F5H transgenic produced significantly higher yields.

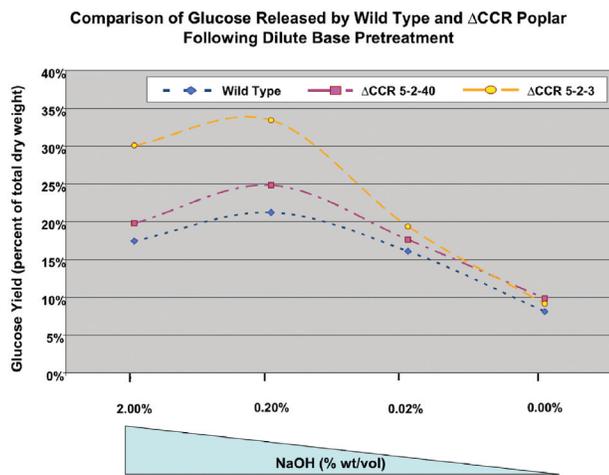


Figure 2. Glucose release for two CCR-deficient transgenics and their controls, following base pretreatment. The most CCR-deficient line has markedly elevated glucose release. Note: in both cases, non-exhaustive saccharification conditions were used allow differences to be meaningfully represented.

5 Biomass Trait Screening in *Brachypodium* Accessions and Mutant Populations

Cynthia Cass,^{1,2} Chris Schwartz,^{1,3} Laura Severs,^{1,2} Samantha Sakol,^{1,2} Michael Steinwand,⁴ John Vogel,⁴ Richard Amasino,^{1,3} and John Sedbrook^{1,2*} (jcsedbr@ilstu.edu)

¹DOE Great Lakes Bioenergy Research Center (GLBRC), Madison, Wis. (<http://www.greatlakesbioenergy.org/>); ²School of Biological Sciences, Illinois State University, Normal; ³Dept. of Biochemistry, University of Wisconsin, Madison; and ⁴Western Regional Research Center, U.S. Department of Agriculture, Albany, Calif.

Project Goals: Our short-term goal is to identify genes and gene variants affecting plant biomass quality and quantity, using *Brachypodium distachyon* as a model system. Our long-term goal is to use these data to improve bioenergy crops such as switchgrass, *Miscanthus*, and poplar through breeding and gene transformation.

The grass *Brachypodium distachyon* is emerging as an important model system for bioenergy crop grasses such as switchgrass and *Miscanthus* owing to its small genome size (~300Mbp), small stature, short generation time, transformability, and self-fertilization. The DOE Joint Genome Institute (JGI) has sequenced the *Brachypodium* genome and a large number of *Brachypodium* Expressed Sequence Tags (ESTs). These data are easily accessible and searchable at www.Brachypodium.org. We are taking a two-pronged approach to identify novel biomass trait genes and allele variants in *Brachypodium*. We are surveying a genetically diverse collection of wild type *Brachypodium* accessions for a variety of traits relevant to biomass production including cell wall hydrolytic enzyme digestibility, cell wall composition and structure, and flowering time, the last of which has a profound affect on biomass production. Not only are these data relevant for identifying gene variants that could be employed to improve bioenergy crops, they are also essential for understanding how different *Brachypodium* genetic backgrounds could modify the phenotypes of novel mutations. In that regard, we are screening through large collections of *Brachypodium* EMS mutagenized plants using both Near Infrared (NIR) spectroscopy (generates a fingerprint of cell wall composition) and an HPLC based cell wall digestibility assay (detects differences in the release of glucose and xylose from cell wall polysaccharides). We will present data on some promising mutants we have identified and discuss how the DOE Great Lakes Bioenergy Research Center (GLBRC) facilitates phenotypic characterization and gene discovery.

6

EST-SSR Markers Discriminate Switchgrass Ecotypes

M.D. Casler^{1*} (michael.casler@ars.usda.gov),
J.E. Zalapa,² S.M. Kaeppler,² N. de Leon,² and
C.M. Tobias³

¹USDA-ARS, U.S. Dairy Forage Research Center, Madison, Wisconsin; ²Dept. of Agronomy, University of Wisconsin, Madison; and ³USDA-ARS, Western Regional Research Center, Albany, Calif.

<http://www.greatlakesbioenergy.org>

Project Goals: Our objective was to use EST-SSR markers to discriminate upland from lowland ecotypes, using broader set of cultivars and individuals per cultivar than has been previously investigated. Longer-term goals of these studies are: (1) to use DNA markers to assist in identifying and selecting parents for development of heterotic gene pools and hybrid cultivars in switchgrass, and (2) to identify hybrid or backcross genotypes of mixed upland and lowland parentage, both in breeding programs and in natural populations.

Switchgrass (*Panicum virgatum*) is an important crop for bioenergy feedstock development. It is native to North America, ranging from Mexico to Canada east of the 100th meridian and adapted to a wide range of habitats (e.g., tallgrass prairie, savanna riparian habitats, etc.). Switchgrass has a range of ploidy from $2n=2x=18$ to $2n=12x=108$ and two main ecotypes: upland and lowland. The two ecotypes originate in different habitats: the upland type, originating on upland soils and the lowland type, originating in riverine and riparian habitats. Further, upland and lowland switchgrass have significantly different adaptations, with lowland types generally found south of 35°N latitude or USDA Hardiness Zones 6 and higher, while upland types tend to be more frequent at higher latitudes. Because there is a significant difference in morphology, growth pattern, and adaptation zones between the upland and lowland ecotypes, it is important to be able to easily discriminate between upland and lowland ecotypes. The presence of hybrid vigor in upland x lowland crosses and the possibility that the two ecotypes may act as natural heterotic groups creates a further need for accurate and efficient discrimination between the two ecotypes. Previous studies to identify DNA markers to discriminate between upland and lowland ecotypes have utilized a very small number of lowland ecotypes, typically only the two most common cultivars, Alamo and Kanlow.

Our objective was to use EST-SSR markers to discriminate upland from lowland ecotypes, using broader set of cultivars and individuals per cultivar than has been previously investigated. Longer-term goals of these studies are: (1) to use DNA markers to assist in identifying and selecting parents for development of heterotic gene pools and hybrid cultivars in switchgrass, and (2) to identify hybrid or backcross

genotypes of mixed upland and lowland parentage, both in breeding programs and in natural populations.

Seven lowland cultivars [Alamo (n=16), Kanlow (n=16), Miami (n=2), SG5 (n=9), Stuart (n=2), Timber (n=9), and Wabasso (n=6)] and 11 upland cultivars [Blackwell (n=16), Carthage (n=8), Cave-in-Rock (n=16), Dacotah (n=16), Forestburg (n=17), Pathfinder (n=9), Shawnee (n=4), Shelter (n=5), Summer (n=16), Sunburst (n=9), and Trailblazer (n=8)] were included in this study, with number of plants for each cultivar shown in parentheses. Additionally, in order to determine their origin, 8 unique switchgrass plants from our USDA-ARS breeding program were included in the study. We used 40 EST-SSR loci (381 alleles) from a total of 2351 that gave optimal amplification (i.e., were highly repeatable, showed correct segregation according to the expected ploidy level of each cultivar, and did not produce artifactual peaks). Relationships among the 18 switchgrass cultivars were investigated using NTSYS-pc version 2.01 based on a molecular binary data set obtained for each individual within cultivars. The resulting binary data was analyzed using the SIMQUAL routine to generate Dice similarity coefficients. Dice similarity coefficients were then used to construct a phenogram employing the SAHN procedure based on the Unweighted Pair-Group Method of the Arithmetic Average (UPGMA). The binary data were also subjected to principal component analysis (PCA) using the EIGEN routine of NTSYS-pc version 2.01.

Plants classified as upland or lowland, based on origin and phenotype, were completely separated by EST-SSR markers, with only two exceptions (Figure 1). Those two exceptions were plants that originated in the USDA-ARS breeding program at Madison, Wisconsin. The plants had been classified as lowland based on phenotype, but their phenotypic is actually intermediate between the extreme upland and lowland phenotypes: intermediate heading date, intermediate height, intermediate color, and intermediate tiller size and density. Although these two individuals were distinct from both upland and lowland phenotypes, based on marker data these clustered with the upland ecotypes. In summary, the EST-SSR markers used in this study were extremely effective at discriminating between upland and lowland ecotypes and at identifying the genetic origin for two plants of unknown origin.

There was a moderate degree of differentiation between upland 4x and upland 8x groups, with about 70% of the individuals with each group discriminated from each other. Ploidy forms an effective reproductive barrier in switchgrass, so it is not surprising to see some level of genetic differentiation between tetraploid and octaploid individuals.

Research sponsored and funded by the Great Lakes Bioenergy Research Center Grant DE-FC02-07ER64494.

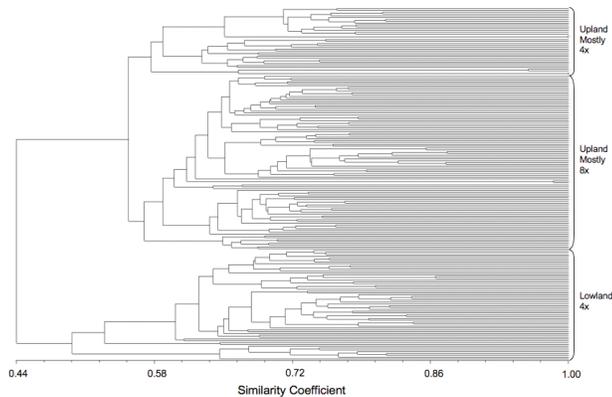


Figure 1. Cluster dendrogram of 192 switchgrass plants of known origin (upland 4x, upland 8x, and lowland 4x) grouped by similarity coefficients based on 381 EST-SSR markers.

7

Characterizing the Microbiome of Leaf-Cutter Ant Fungus Gardens

Garret Suen^{1,2*} (gsuen@wisc.edu), Jarrod J. Scott,^{1,2,3} Frank O. Aylward,^{1,2} Sandra M. Adams,^{1,2} Susannah G. Tringe,⁴ Adrián A. Pinto-Tomás,^{5,6} Clifton E. Foster,^{1,7} Markus Pauly,^{1,7,8} Paul J. Weimer,⁹ Kerrie W. Barry,⁴ Lynne A. Goodwin,¹⁰ Pascal Bouffard,¹¹ Lewyn Li,¹¹ Jolene Osterberger,¹² Timothy T. Harkins,¹² Steven C. Slater,¹ Timothy J. Donohue,^{1,2} and **Cameron R. Currie**^{1,2,3}

¹U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center and ²Dept. of Bacteriology, University of Wisconsin, Madison; ³Smithsonian Tropical Research Institute, Balboa, Ancon, Republic of Panama; ⁴DOE Joint Genome Institute, Walnut Creek, Calif.; ⁵Departamento de Bioquímica, Facultad de Medicina and ⁶Centro de Investigaciones en Estructuras Microscópicas, Universidad de Costa Rica, Ciudad Universitaria Rodrigo Facio, San José; ⁷Dept. of Biochemistry and Molecular Biology and ⁸DOE Plant Research Laboratory, Michigan State University, East Lansing; ⁹Dairy Forage Research Center, U.S. Department of Agriculture–Agricultural Research Services, Madison, Wis.; ¹⁰DOE Joint Genome Institute, Los Alamos National Laboratory, Los Alamos, New Mexico; ¹¹454 Life Sciences, a Roche Company, Branford, Conn.; and ¹²Roche Diagnostics, Roche Applied Science, Indianapolis, Ind.

Project Goals: Understanding efficient plant biomass degradation is a critical step toward the technological goal of ethanol production. Microbes are the primary organisms on earth capable of deconstructing lignocellulose, and herbivores can gain access to organic carbon stored in plant cell walls by forming symbiotic relationships with lignocellulolytic microbes. Our project aims to characterize an insect herbivore, leaf-cutter ants, which have been farming a fungus for food for ~10 million years. These

ants, one of the most widespread and dominant herbivores in the Neotropics, are capable of harvesting a tremendous amount of leaf-material to grow their fungus. They form massive colonies consisting of hundreds of fungus garden chambers supporting millions of workers. We are utilizing a combination of culture-independent and culture-dependent approaches to characterize the microbiome of leaf-cutter ant fungus gardens, and thereby understand how this microbial community synergistically degrades plant biomass.

For ~10 million years, leaf-cutter ants have been farming fungus for food. The ant–fungus system is one of the most complex described symbioses in nature, consisting of at least four mutualists and two pathogens. These ants, which are one of the most widespread insects in the Neotropics, can have massive colonies containing hundreds of fungus garden chambers and millions of workers that forage for hundreds of Kg (dry weight) of leaf material each year. The success of the leaf-cutter ants can be attributed, in part, to their ability to convert plant biomass into nutrients through their obligate mutualistic fungus. This fungus serves as the primary food source for the entire colony, and in return, the ants provide the fungus with a constant source of nutrients, protection from competitors, and dispersal through colony founding. As a result, the leaf-cutter ants serve as an excellent model for understanding how plant biomass degradation occurs in a highly-evolved, natural system.

Interestingly, very little is known about plant biomass degradation in leaf-cutter ant fungus gardens, even though this process likely plays a critical role in these colonies reaching immense sizes. For example, it is thought that the fungus the ants cultivate for food is responsible for the majority of plant biomass degradation in the garden, despite the fact that it is incapable of deconstructing lignocellulose. To explore the possibility that a largely uncharacterized microbial community is present and responsible for biomass deconstruction in leaf-cutter ant fungus gardens, we describe the fungus garden microbiome of the leaf-cutter ant *Atta colombica*.

We employ a combination of sugar composition analysis, 16S rDNA sequencing, community metagenomics, and whole-genome sequencing to demonstrate that lignocellulose is degraded within leaf-cutter ant fungus gardens, and that this microbiome is dominated by γ -proteobacteria in the family *Enterobacteriaceae*. Our analysis also identified a diversity of microbial genes predicted to encode for enzymes involved in cellulose and hemicellulose deconstruction, suggesting that a community of microbes is likely involved in plant biomass deconstruction. Comparative metagenomic analyses with 13 other microbiomes revealed that the leaf-cutter ant fungus garden exhibits a similar carbohydrate-degrading potential as bovine rumen, which is also capable of processing large amounts of plant biomass. Finally, genomic and physiological characterization of two dominant bacteria in the fungus garden provided evidence for their capacity to degrade lignocellulose, and suggests a potential mutualism, as these bacteria are known nitrogen-fixing symbionts of leaf-cutter ants. Our analysis of the leaf-cutter ant fungus

garden microbiome provides insight into how this microbial community synergistically deconstructs plant biomass.

8

Optimization of Enzymes for Alkaline-Pretreated Biomass Conversion

Shishir Chundawat,^{1,3} Dahai Gao,^{1,3} Venkatesh Balan,^{1,3} Goutami Banerjee,^{2,3} John Scott-Craig,^{2,3} Jonathan D. Walton,^{2,3} David Mead,⁴ Phil Brumm,⁴ and **Bruce Dale**^{1,3*} (bdale@egr.msu.edu)

¹Biomass Conversion Research Laboratory, Dept. of Chemical Engineering and Materials Science, ²DOE Plant Research Laboratory, and ³DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing; and ⁴Lucigen Corporation, Madison, Wisconsin.

Project Goals (Abstracts 8 and 9): The GLBRC's mission is grand, but simply stated: to perform the basic research that generates technology to convert cellulosic biomass to ethanol and other advanced biofuels. To accomplish its mission in a university research environment, the GLBRC will: apply cutting-edge research to help create a new generation of sustainable bioenergy feedstocks, processing technologies, and fuels; evaluate the economic and environmental impacts of these new technologies; use the results to guide research activities; bring technological advances to other academic scientists, the private sector, and the marketplace; balance mission-driven project management and evaluation with the creative milieu of its academic, private sector and national laboratory research partners; recruit broad segments of the academic, industrial, and national laboratory communities to develop and participate in relevant research programs; and provide a training program for future leaders of the biofuels industry.

Enzymes that release fermentable sugars from biomass feedstocks are one of the major costs in converting lignocellulose to ethanol. The major goal of this project is to build optimal enzyme mixtures for alkaline-pretreated biomass. This project has two stages of development. The first is to define an optimized "minimal enzyme set" composed of those enzymes that are almost certainly essential for lignocellulose degradation. This set includes exo-glucanases (cellobiohydrolases; CBHs), endo-glucanase (EG), endo-xylanase (EX), β -xylosidase (BX), and β -glucosidase (BG). The second part is to develop an "optimized enzyme mixture" in which additional enzymes, called here "accessory" enzymes, are added to the minimal set in order to create a superior mixture. By optimum we mean having the highest specific activity (lowest protein loading) to achieve a benchmark release (e.g., 85%) of glucose and xylose in a specified time (48 hr). We used Ammonia Fiber Expansion (AFEX) pretreated corn stover for this work. The enzymes for these experiments came from several sources. These include enzymes purified from commercial *Trichoderma reesei* preparations, *Trichoderma* genes expressed in *Pichia pastoris*, *Trichoderma*

proteins over expressed in *Trichoderma* itself and *A. nidulans* orthologs of the *Trichoderma* enzymes expressed in *Pichia pastoris*. Another important source of enzymes for our experiments are proteins, mainly bacterial, from other projects in the GLBRC. We have successfully completed the first stage of the project by defining the core enzymes needed to achieve the benchmark glucan and xylan conversions in 48 hr of hydrolysis. The second phase of the project using accessory enzymes along with the core enzymes is in progress. Preliminary results indicate that there is synergy operating between bacterial hemicellulases and fungal cellulases. When completed, the optimized mixture is predicted to have higher specific activity than current commercial mixtures, enabling a reduced enzyme load during hydrolysis.

9

Physicochemical Characterization of Alkali-Pretreated Lignocellulosic Biomass

Shishir Chundawat,^{1,3} Leonardo Sousa,^{1,3} James Humpala,^{1,3} Venkatesh Balan,^{1,3} **Bruce Dale**^{1,3*} (bdale@egr.msu.edu), Ramin Vismeh,^{2,3} A. Daniel Jones,^{2,3} Fachuang Lu,^{4,5} Lily Summer,^{4,5} Ali Azarpira,^{4,5} John Ralph,^{4,5} Chris Bianchetti,^{4,5} George Phillips,^{4,5} Bryon Donohoe,⁶ Michael Himmel,⁶ Umesh Agarwal,⁷ Paul Langan,⁸ S. Gnanakaran,⁹ and Albert Che¹⁰

¹Biomass Conversion Research Laboratory, Dept. of Chemical Engineering and Materials Science, ²Dept. of Biochemistry and Molecular Biology, and ³Great Lakes Bioenergy Research Center, Michigan State University, East Lansing; ⁴Dept. of Biochemistry and ⁵Great Lakes Bioenergy Research Center, University of Wisconsin, Madison; ⁶Chemical and Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado; ⁷Fiber and Chemical Sciences Research, Forest Products Laboratory, Madison, Wisconsin; ⁸Biosciences Division and ⁹Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico; and ¹⁰Environmental Science and Chemistry, American University, Washington, D.C.

Project Goals: See goals for abstract 8.

The development of an economically viable and environmentally sustainable bio-based chemical industry has been impeded by the native recalcitrance of plant cell walls to thermochemical and biological processing. Alkaline pretreatments like Ammonia Fiber Expansion (AFEX) enhance cell wall digestibility through certain ultra-structural and chemical modifications that are currently poorly understood, unlike other acidic pretreatments (e.g. dilute acid, steam explosion). Understanding the physicochemical mechanisms by which alkaline based pretreatments enhance cell wall enzymatic digestibility would result in development of improved pretreatment methodologies and reduction in cellulosic ethanol production costs. The goal of this project is to identify ultra-structural and chemical modifications incorporated within lignocellulosic cell walls during

alkaline based treatments (e.g. AFEX, alkaline peroxide). High resolution microscopic (SEM, TEM) and 3D-EM-Tomographic studies indicate ultra-structural alteration of AFEX treated cell walls via formation of nanoporous (5-500 nm) tunnel-like networks. Closer analysis (ESCA, AFM and confocal fluorescence microscopy) of outer cell wall surfaces reveals presence of heterogeneous deposits rich in AFEX degradation products and other cell wall extractives (e.g. lignin, arabinoxylan based oligomers, calcium). More than 50 alkaline based degradation products have been identified and quantified using LC-MS/MS and GC-MS, with ammonolysis based by-products (acetamide and phenolic amides) being the predominant ones. Pretreated biomass was characterized by NMR to elucidate modification of various cell wall components during AFEX. Raman and XRD analyses indicate allomorphic conversion of cellulose I to III by treatment with anhydrous liquid ammonia. Importantly, the cellulose III allomorph was found to have a significantly higher rate of enzymatic hydrolysis than untreated cellulose possibly due to differences of glucan chain packing within the cellulose crystal lattice.

10

Combinatorial Discovery of Enzymes for Biomass Deconstruction

Brian G. Fox* (bgfox@biochem.wisc.edu), Abolfazi Arabshahi, Allison Riederer, Nathaniel L. Elsen, Pirom Chenprakon, David Stephenson, Paul Weimer, and Yuriy Bukhman

Dept. of Biochemistry, Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

Project Goals: The goal of our work is to provide a new combinatorial paradigm for evaluation of enzymes from new environmental sources as well as synthesized genes and engineered enzymes.

The goal of our work is to provide a new combinatorial paradigm for evaluation of enzymes from new environmental sources as well as synthesized genes and engineered enzymes. The GLBRC bioenergy platform derives from work with cell-free translation at the NIH Protein Structure Initiative-funded Center for Eukaryotic Structural Genomics, where over 10,000 genes from various eukaryotic organisms have been cloned, tested for expression, and in the best performing cases, purified and subjected to structure determinations. An adaption of the modular vector design of this platform provides the basis for this new effort on genes and proteins contributing to cellulose destruction. Genes identified by bioinformatic evaluation of new genomes or by microarray evaluation of gene expression in cellulose utilizing organisms, and proteins identified by mass spectral proteomic studies can be targeted for high-throughput cloning and cell-free translation. Methods to prepare combinations of genes or to supplement existing mixtures of enzymes with new translation products are demonstrated to provide biomass deconstruction without need for purifica-

tion of translation products. Iterative substitution of gene variants can be used to identify protein homologs with better behavior in defined assays that can include diagnostic small molecules, model purified celluloses, or actual biomass substrates. Results of the application of these methods to characterization of the reactivity of various treated and untreated biomass materials will be presented.

This work is funded by DOE DE-FC02-07ER64494 as a subgrant from the Great Lakes Bioenergy Research Center to B.G.F.

11

Construction of Gram-Negative Consolidated Bioprocessors

David H. Keating* (dkeating@glbrc.wisc.edu), Jeffrey Gardner, Mary Tremaine, and Robert Landick

Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

Project Goals: Research within the Great Lakes Bioenergy Research Center (GLBRC) aims to generate an improved understanding of the bottlenecks associated with conversion of lignocellulose to ethanol. We seek to construct consolidated bioprocessing strains capable of the degradation of lignocellulose, and fermentation of the liberated sugars to ethanol. We will then use these consolidated bioprocessing strains to better understand the current enzymatic and metabolic bottlenecks in cellulosic ethanologenesis. These studies are currently focused on the gram-negative bacterium *Escherichia coli* due to its sophisticated genetics, well-understood physiology, and widespread use as an industrial microbe.

The conversion of *E. coli* to a consolidated bioprocessor requires the introduction of genes encoding lignocellulases, as well as a mechanism for their secretion from the cell. We developed liquid and solid media assays that facilitate the rapid identification of bacteria capable of biomass degradation. Using this media, we have categorized known cellulose-degrading organisms such as *Cellvibrio japonicus* for their ability to degrade key GLBRC biomass substrates, and to identify novel gram-negative cellulolytic strains. We have introduced cellulase genes from *C. japonicus* into *E. coli* and generated first generation strains capable of cellulose degradation. Using a genetic system developed for *C. japonicus*, we have obtained evidence that the majority of cellulolytic activity produced by *C. japonicus* is secreted via the Type II secretion system. We are currently engineering *E. coli* to express the *C. japonicus* Type II secretory apparatus in order to improve secretion of cellulolytic enzymes.

We are also engineering *E. coli* to efficiently produce and tolerate ethanol. To develop efficient *E. coli* ethanologens, we have constructed deletions that inactivate alternative pathways of electron flow, and introduced the *Zymomonas mobilis* pyruvate decarboxylase and alcohol dehydrogenase genes,

which allow efficient conversion of pyruvate to ethanol. Our studies demonstrated that this ethanologen was capable of efficient ethanol production under micro-aerobic conditions, but grew very poorly under strictly anaerobic conditions. To better understand the mechanism behind the poor anaerobic growth of this strain, we subjected it to thirteen rounds of sequential subculture under anaerobic conditions. We identified 32 mutants that grew well anaerobically with glucose as a carbon source, five of which exhibited productivities greater than that of strain KO11 (the current *E. coli* benchmark). We are currently subjecting these strains to re-sequencing and multiomic analysis to understand the genetic changes responsible for restoration of anaerobic growth.

Collectively, we expect that these approaches will allow for the isolation of lead organisms that can then be subjected to additional metabolic engineering and directed evolution, with the aim of improving lignocellulolytic ability, ethanol tolerance, and ethanol production.

12 Modeling and Summarizing Growth Curves from High-Throughput Screening Data

Yury V. Bukhman* (ybukhman@glbrc.wisc.edu), Jason Shao, Trey Sato, and Sydnor Withers

DOE Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

Project Goals: This project provides computational support for high-throughput screening of large collections of microbes grown on various media. The screening experiments at GLBRC aim to identify yeast strains that are most suited for growth in media that include potential feedstocks for industrial fermentation of plant biomass into ethanol. The experiments monitor cell density of yeast cultures grown in microtiter plates by measuring optical density at 600 nm at regular intervals. We aim to mathematically model the resultant growth curves and summarize them by computing a few essential characteristics of each curve.

We have developed R scripts to process and summarize our screening data. We chose to fit theoretical functions, instead of using numerical differentiation or local regression methods, in order to be able to work with limited number of data points, which maintains maximal robotic screening throughput. We have tested several known growth curve models and found that the Gompertz function^{1,2} gives the best results with our data. We have also developed methods for automatic generation of initial guesses of curve parameters and for dealing with growth curves that have anomalous shapes. Upon processing the screening data, our software generates a table of biologically meaningful growth curve characteristics, such as the fastest intrinsic growth rate, lag time, and total growth. These characteristics were used to select yeast strains with optimal growth properties in the examined

media. Although the development of this methodology was motivated by the needs of the yeast screening program at GLBRC, it has general utility for high-throughput screening studies of microbial strain collections.

References

1. Gompertz, B. "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies." *Phil. Trans. Roy. Soc. London* 123, 513-585, 1832.
2. Zwietering, M.H., et al. "Modeling of the Bacterial Growth Curve." *Applied and Environmental Microbiology* 56 (6), 1875-1881, 1990.

13 Exploiting Natural Diversity in Wild Yeast Strains to Dissect the Mechanisms of Ethanol and Thermotolerance

Jeffrey A. Lewis^{1,2,*} (jalewis4@wisc.edu), Shau Chee C. Chung,² and **Audrey P. Gasch**^{1,2}

¹Great Lakes Bioenergy Research Center, Madison, Wis.; and ²Laboratory of Genetics, University of Wisconsin, Madison

Project Goals: see below

Pretreatment of cellulosic biomass for enzymatic saccharification generates degradation products that impair microbial fermentation. Additional stressors that affect the final ethanol yield include osmotic stress, oxidative stress, heat, and ethanol itself. This project specifically focuses on ethanol and heat tolerance in *Saccharomyces cerevisiae*—stressors that are especially relevant for simultaneous saccharification and fermentation.

Acquired stress resistance is the phenomenon where cells exposed to a mild dose of a primary stress can survive an otherwise lethal dose of a subsequent stress. In the case of ethanol, acquired resistance likely reflects the adaptation required to survive the increasing ethanol concentrations that accumulate during fermentation. While studying acquired stress resistance in *S. cerevisiae*, we made a surprising discovery—our lab strain (S288c) could not acquire resistance to ethanol. Intriguingly, our lab strain could acquire thermotolerance, even though the mechanisms of ethanol tolerance and thermotolerance are thought to be shared.

By testing a wide panel of wild and industrial yeast isolates, we verified that the lab strain was indeed an outlier; acquired ethanol resistance is wide-spread in nature. We have compared the global transcriptional response to both ethanol and heat, in both the lab strain and two wild strains (with a natural ability to acquire ethanol tolerance). Stark differences existed in the transcriptional profile between the lab strain and the two wild strains. In particular, genes known to function in both ethanol and thermotolerance (i.e. genes involved in membrane metabolism, chaperones, and

trehalose metabolism) were differentially expressed. These differences are valuable clues for elucidating the regulatory circuits for both acquired ethanol resistance and acquired thermotolerance. Screening of mutants (chosen on the basis of the transcriptome data) has led to the discovery of novel genes of involved in ethanol resistance, highlighting the power of this approach.

14

Utilization of Directed Evolution, Resequencing, and Multiomics to Improve Ethanol Tolerance in *Escherichia coli*

Michael S. Schwalbach* (mschwalbach@glbrc.wisc.edu), Frances Tran, Mary T. Tremaine, Patricia J. Kiley, Robert Landick and **David H. Keating**

Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

Project Goals: The Great Lakes Bioenergy Research Center mission is grand, but simply stated; to perform the basic research that generates technology to convert cellulosic biomass to ethanol and other advanced biofuels. While the larger mission is pursued on multiple fronts by many researchers, the focus of this particular research project is to understand the bottlenecks associated with conversion of lignocellulose to ethanol. Ethanol toxicity represents one such bottleneck since elevated concentrations of ethanol inhibit fermentation yields and ultimately growth in ethanologenic bacteria.

Research within the Great Lakes Bioenergy Research Center aims to understand the bottlenecks associated with conversion of lignocellulose to ethanol. Ethanol toxicity represents one such bottleneck since it inhibits fermentation yields and ultimately growth in ethanologenic bacteria. However, ethanol tolerance capabilities vary amongst ethanologenic bacteria, with bacteria such as *Zymomonas mobilis* and *Lactobacillus buchneri* able to survive external ethanol concentrations approaching 15% (v/v) while others, such as *Escherichia coli*, are inhibited at concentrations greater than 3% (v/v). Presently the molecular mechanisms underlying differences in ethanol sensitivity are not wholly characterized nor are the cellular responses that mitigate the toxic effects of ethanol well understood. To understand better which traits contribute to improved ethanol tolerance, we conducted directed evolution experiments to incrementally increase ethanol tolerance of *E. coli* strain MG1655. Ethanol tolerant mutants and wild type cells were then exposed to 4% ethanol and the response was monitored via multiomics. From these growth experiments we collected transcriptomics, metabolomics, lipidomics and proteomic data from three time points: mid-log, ten minutes post ethanol addition, and one hour post ethanol addition. Global transcriptional analysis from cells collected ten minutes after ethanol addition indicated that wild type and the ethanol tolerant mutant exhibited mechanistically similar responses to initial ethanol challenge. Conversely, examination of the tran-

scriptional response of cells collected one hour after ethanol addition determined the initial response had diminished in the ethanol tolerant mutant, whereas wild type maintained a transcriptional profile similar to what was observed ten minutes after ethanol addition. Our results suggest that the ethanol tolerant mutant was able to mitigate the adverse effects of the ethanol challenge quickly whereas wild-type remained adversely affected by ethanol for a longer period of time. By comparing the complete multiomic responses and genomic differences in multiple strains evolved to tolerate varying concentrations of ethanol, we will identify the key cellular processes associated with ethanol tolerance.

15

Generation of a Computational Metabolic Network Representing the Pangenome of *Escherichia coli* and Construction and Validation of Six Strain-Specific Models

David J. Baumler^{1,2*} (dbaumler@wisc.edu), Roman G. Peplinski,² Jeremy D. Glasner,² Jennifer L. Reed,^{1,3} and Nicole T. Perna^{1,2,4}

¹Bringing Advanced Computational Techniques to Energy Research (BACTER) Institute; ²Genome Center of Wisconsin; ³Dept. of Chemical and Biological Engineering, and ⁴Dept. of Genetics, University of Wisconsin, Madison

Project Goals: BACTER is a research training initiative funded by the U.S. Department of Energy to promote the development of computational biology and bioinformatics techniques that are of direct value to energy science.

Within the genus *Escherichia*, only one strain, the laboratory strain *E. coli* K-12 MG1655, has had a genome-scale metabolic computational model constructed. This single model has proven useful for many applications, such as guiding the bioengineering of strains for increased production of desired end-products. We sought to enhance these efforts by constructing metabolic models for additional *E. coli* strains using a strategy based on analysis of the collective contents of all *E. coli* genomes, or the pangenome. We used 17 complete genome sequences (Table 1) to generate an *E. coli* pangenomic metabolic network consisting of the collective information from 76,080 ORFs. These ORFs were clustered into 17,647 orthologous groups. The 1,260 orthologous groups containing the ORFs used in the most recent metabolic network for *E. coli* K-12 were identified, and the gene to protein to reaction associations were propagated to the other *E. coli* strains. All remaining orthologous groups were surveyed for new metabolic reactions to add to the pangenomic network. This allowed us to update the metabolic model for *E. coli* K-12 MG1655 to account for 1,322 ORFs, and now includes a pathway for phenylacetate metabolism. A model was constructed for the very similar *E. coli* K-12 strain W3110 and differs by one metabolic reaction for galactitol transport/utilization. Genome-scale metabolic models were also constructed for enterohemorrhagic *E. coli*

O157:H7 strains EDL933 and Sakai, and uropathogenic *E. coli* strains CFT073 and UTI89 (Table 2). The metabolic networks for the pathogens contained numerous lineage-specific ORFs when compared to the K-12 models. All six *E. coli* models were used to simulate growth in different conditions. The results were compared to experimental data we collected for each strain including tests for the utilization of 76 different carbon sources in conditions with or without oxygen, and growth in batch culture. The experiments reveal metabolic differences between strains and the *in silico* results accurately predict some of these differences. Our findings demonstrate that use of the pangenomic metabolic network allows rapid construction of additional *E. coli* strain-specific models that can accurately predict strain-specific phenotypes and offers a larger suite of metabolic capabilities for engineering new *E. coli* strains.

Table 1. *E. coli* genomes used to construct the pangenome metabolic network

Strain	ORFs
<i>E. coli</i> K-12 MG1655	4,141
<i>E. coli</i> K-12 W3110	4,171
<i>E. coli</i> EDL933 (EHEC) ^a	5,196
<i>E. coli</i> Sakai (EHEC) ^a	5,253
<i>E. coli</i> CFT073 (UPEC) ^b	4,889
<i>E. coli</i> UTI89 (UPEC) ^b	4,944
<i>E. coli</i> 536 (UPEC) ^b	4,599
<i>E. coli</i> 53638 (EIEC) ^c	5,172
<i>E. coli</i> APEC O1 (APEC) ^d	5,045
<i>E. coli</i> ATCC 8739	4,236
<i>E. coli</i> E2348/69 (EPEC) ^e	4,652
<i>E. coli</i> E24377A (ETEC) ^f	4,953
<i>E. coli</i> EC4115 (EHEC) ^a	5,467
<i>E. coli</i> HS	4,393
<i>E. coli</i> K-12 DH10B	4,136
<i>E. coli</i> SE11	4,973
<i>E. coli</i> SMS-3-5	4,906

^aEnterohemorrhagic *E. coli* (EHEC)

^bUropathogenic *E. coli* (UPEC)

^cEnteroinvasive *E. coli* (EIEC)

^dAvian pathogenic *E. coli* (APEC)

^eEnteropathogenic *E. coli* (EPEC)

^fEnterotoxigenic *E. coli* (ETEC)

Table 2. *E. coli* strain-specific metabolic model information

Strain	Additions			Deletions		Essential Reactions ^a	Total in model	
	ORFs	Reactions	Isozymes	ORFs	Reactions		ORFs	Reactions
K-12 MG1655	-	-	-	-	-	-	1,322	2,282
K-12 W3110	0	0	0	1	1	0	1,321	2,281
O157:H7 EDL933	38	9	19	51	56	8	1,328	2,235
O157:H7 Sakai	36	8	23	52	57	8	1,329	2,333
UPEC CFT073	9	9	25	87	55	10	1,269	2,236
UPEC UTI89	8	8	26	71	55	6	1,285	2,235

^aReactions with no orthologous ORF (s), but are essential to the *in silico* model

This work was supported through the BACTER training program (PI: Julie C. Mitchell University of Wisconsin-Madison [http://www.bacter.wisc.edu/BACTER/Welcome.html]), funded by the DOE:GTL #: DE-FG02-04ER25627, BACTER is a research training initiative funded by the U.S. Department of Energy to promote the development of computational biology and bioinformatics techniques that are of direct value to energy science.

16 Constraint-Based Analysis of Microbial Regulatory and Metabolic Networks for Ethanol Production

Joonhoon Kim^{1,2*} (jkim92@wisc.edu), Dipak Barua,^{1,2} and Jennifer L. Reed^{1,2}

¹Dept. of Chemical and Biological Engineering and ²DOE Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

Project Goals: Design microbial strains for improved ethanol production using computational models of transcriptional regulation and metabolism.

Computational modeling and analysis of metabolic networks has been successful in metabolic engineering of microbial strains for valuable biochemical production. Limitations of currently available methods are that they are often based on reaction deletions rather than gene deletions and that they do not consider the regulatory networks that control metabolism. Thus, such methods may result in strategies that are not genetically feasible, or designed strains might not be able to grow due to the regulatory restrictions. To overcome these limitations, we developed an effective method to systematically integrate transcriptional regulatory networks and metabolic networks, which allows for the simulation of gene deletion and overexpression.

Using integrated transcriptional regulatory and metabolic models, we developed an automated approach (*Gene-Force*) for refining transcriptional regulatory rules against high-throughput growth phenotypic data. The developed approach was applied to well-curated transcriptional regulatory and metabolic models of *Escherichia coli* (Covert et al, 2003), and resulted in an overall ~ 10% improvement in model prediction accuracy for a large collection of mutant growth phenotypes (Glasner et al, 2003; Ito et al, 2005). An advantage of using an integrated model of metabolism and regulation is that an integrated model is better at predicting essential genes under a given condition, and hence it prevents gene deletions which are lethal from being included in the strategies. Accordingly, strains that are designed with regulatory considerations should grow better initially and may achieve the desired phenotype faster.

We have subsequently developed a new approach (OptORF) for identifying metabolic engineering strategies based on metabolic and transcription factor gene deletions, as well as gene overexpressions. This new approach uses integrated models of metabolism and regulation, and searches for the minimal metabolic and/or regulatory perturbations that couple biomass and biochemical production, thus proposing adaptive evolutionary strain designs. Using genome-scale models of *E. coli* (Covert et al, Nature 2003), we have implemented OptORF and compared its metabolic engineering strategies for ethanol production to those found using OptKnock (Burgard et al, 2003). The developed OptORF approach is general and can be applied to the production of different compounds in other biological systems.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

17

Identification of Stress-Tolerant *Saccharomyces cerevisiae* Strains for Fermentation of Lignocellulosic Feedstocks by High-Throughput Phenotypic Screening

Sydnor Withers,¹ Jason Shao,¹ Benjamin Bice,¹ Miguel Dominguez,¹ Venkatesh Balan,² Supratim Datta,⁴ Mark Davis,⁵ Tom Jeffries,^{5,6} Blake Simmons,⁴ Bruce Dale,² Wes Marner,¹ Yury Bukhman,¹ and **Trey Sato**^{1*} (tksato@glbrc.wisc.edu)

¹Great Lakes Bioenergy Research Center, Madison, Wis.; ²Michigan State University, East Lansing; ³Joint BioEnergy Institute, Emeryville, Calif.; ⁴BioEnergy Science Center and National Renewable Energy Laboratory, Golden, Colo.; ⁵University of Wisconsin, Madison; and ⁶USDA Forest Products Laboratory, Madison, Wis.

Project Goals: A major bottleneck in the conversion of lignocellulosic biomass into ethanol by the brewer's yeast *Saccharomyces cerevisiae* is the inhibition of fermentation resulting from cellular stress, which can be caused by degradation products generated during feedstock pretreatment. At the Great Lakes Bioenergy Research Center (GLBRC), we have sought to identify environmental and industrial *S. cerevisiae* isolates with greater tolerance to these stresses than laboratory strains, and understand the genetic, biochemical, and molecular traits contributing to these phenotypic differences. Results from these studies will provide insight for improving existing production strains, or developing new stress-tolerant strains for cellulosic ethanol production.

Although commonly used for the industrial production of grain ethanol, *S. cerevisiae* has a number of biological hurdles that currently prevent it from being widely utilized in the generation of fuel ethanol from lignocellulosic biomass.

Cellular and physiological stresses imposed by degradation products from feedstock pretreatment are known to limit the yield and efficiency of fermentation. Well-known degradation products include acetic acid, furfurals and lignin-derived phenolic compounds, all of which can vary in structure and concentration with different pretreatment processes.

Ongoing research at the GLBRC has compared environmental stress tolerance between laboratory and wild yeast strains, and discovered that genetic background is a significant determinant in the ability of strains to tolerate environmental stress. This also suggested that *S. cerevisiae* strains isolated from diverse natural or industrial environments, opposed to domesticated lab strains, may have traits that may allow for greater tolerance to the toxins that result from pretreatment of lignocellulosic biomass. To identify strains that may be tolerant to stresses imposed during cellulosic fermentation, we examined the growth properties of hundreds of diverse *S. cerevisiae* strains cultured in various lignocellulosic hydrolysates. These various hydrolysates were prepared from biomass pretreated by ammonium fiber expansion (AFEX), ionic liquid (IL), oxalic acid or dilute sulfuric acid. High throughput robotic screening and computational methods were developed and applied to identify the top performing strains in the various media conditions. One of the top strains included an environmental isolate that, in contrast to most other strains, grows well in AFEX-pretreated corn stover hydrolysate at elevated temperatures of 37 and 40°C. Characterization of these top performing strains for ethanol production and ability to ferment xylose is currently in progress and will be presented.

18

Efforts to Enhance Solar Hydrogen Production by Heterocyst-Forming Cyanobacteria

Eric L. Hegg^{1*} (EricHegg@msu.edu), Katrin Gaertner,¹ Adam Cornish,¹ Sigal Lechno-Yossef,² Xiaowen Kong,² C. Peter Wolk,² Hajime Masukawa,³ Robert P. Hausinger,^{1,3} Jeong-Jin Park,³ and Claire Vieille^{1,3}

DOE Great Lakes Bioenergy Research Center, ¹Dept. of Biochemistry and Molecular Biology, ²MSU-DOE Plant Research Laboratory, and ³Dept. of Microbiology and Molecular Genetics, Michigan State University, East Lansing

<http://www.greatlakesbioenergy.org/>

Project Goals: The overall goal of our project is to engineer *Anabaena* to enhance photobiological H₂ production to a commercially practical level. To accomplish this task we are using a number of different strategies including mutating nitrogenase to produce more H₂, expressing a native bidirectional [NiFe]-hydrogenase in heterocysts, and heterologously expressing an [FeFe]-hydrogenase in heterocysts. In addition, to guide genetic engineering and to enhance further H₂ production, we are elucidating the

pathways by which electrons are channeled to H₂ and the major pathways that compete for those electrons.

H₂ has the potential to become an important clean and renewable energy commodity, especially if it is generated by organisms such as cyanobacteria that use sunlight as the sole energy source and water as the ultimate electron donor. Hydrogenases (H₂ases) and nitrogenases (N₂ases) are the enzymes that produce H₂. Because these enzymes are also O₂-sensitive, oxygenic photosynthesis and H₂ production are normally separated temporally or spatially. *Anabaena* spp. and related cyanobacteria form specialized cells known as heterocysts in which N₂ases and H₂ases are protected from O₂ by inactivation of O₂-producing PSII, accelerated respiration, and synthesis of a thick envelope of glycolipids and polysaccharides that impedes O₂ penetration. Reductant required for N₂ fixation and H₂ production is generated by photosynthesis in vegetative cells and is transported to heterocysts as sugar. Our ultimate goal is to engineer a Hup⁻ strain of *Anabaena* (in which the uptake hydrogenase is inactive) to increase H₂ production to a commercially practicable level.

To increase H₂ production by N₂ase in *Anabaena*, we mutated residues near the active site, generating a total of 49 single variants. Several mutants exhibited significantly higher H₂ production rates in the presence of N₂ than did the parental Hup⁻ strain. Nineteen additional single and double site-directed mutants designed to obstruct a putative channel connecting the active site to the protein surface did not exhibit significantly higher H₂ production rates than did the parental Hup⁻ strain.

Two different strategies are being pursued to express H₂ases in *Anabaena* heterocysts. In the first strategy, efforts are under way to overexpress the native, bidirectional [NiFe]-H₂ase (Hox) genes in heterocysts using the strong *nif* promoter. Genes required for the maturation of Hox are being expressed on a replicating plasmid from the heterocyst-specific promoters of *coxBII* and *patB*. In the second strategy, [FeFe]-H₂ases and the proteins needed for their maturation are being heterologously expressed in *Anabaena* heterocysts using both integrating and replicating plasmids driven by heterocyst-specific promoters. To obtain optimal expression and activity, we are testing different combinations of [FeFe]-H₂ases and maturation proteins from a variety of organisms including *Chlamydomonas reinhardtii*, *Clostridium acetobutylicum*, and *Shewanella oneidensis*.

To elucidate how electrons are channeled to H₂ production, we compared gene expression in phototrophic, photoheterotrophic, and heterotrophic *Anabaena* cultures using RNA extracted separately from vegetative cells, heterocysts, and whole filaments. Principal component analysis of gene expression data confirmed that gene expression patterns in the vegetative cells differ from those in the heterocysts. 19%, 17%, and 16% of the genes are differentially expressed (at least 2X difference, p<0.01) between vegetative cells and heterocysts in phototrophic, photoheterotrophic, and heterotrophic growth conditions, respectively. In addition, gene expression patterns in vegetative cells vary in the different growth conditions, consistent with the fact that vegetative

cells are solely responsible for carbon uptake. Surprisingly, Hup genes and several N₂ase maturation genes are expressed in the vegetative cells in both phototrophic and photoheterotrophic cultures. These and other results will be discussed.

19 Genome-Enabled Analysis of Partitioning of Reducing Power During Fuel Production by the Photosynthetic Bacterium *Rhodobacter sphaeroides*

Wayne S. Kontur^{1*} (wskontur@wisc.edu), Eva C. Ziegelhoffer,¹ Melanie A. Spero,¹ L. Safak Yilmaz,² Daniel R. Noguera^{2*} (noguera@engr.wisc.edu), and Timothy J. Donohue^{1*} (tdonohue@bact.wisc.edu)

¹Dept. of Bacteriology, ²Dept. of Civil and Environmental Engineering, University of Wisconsin, Madison

Project Goals: We seek to understand and improve light- and feedstock-powered production of renewable fuels by the photosynthetic bacterium *Rhodobacter sphaeroides*. We have initially chosen to investigate hydrogen (H₂) production, both to optimize the production of H₂ itself and to serve as a model for partitioning of intracellular reducing potential (high-energy electron equivalents) utilized in microbial biofuel production in general. We are characterizing the intracellular pathways involved in distributing electrons throughout the cell via genome-enabled assays, such as microarray and proteomics analyses, coupled with biomass, polyhydroxybutyrate (PHB) and chemical oxygen demand (COD) analyses. We are using mutant strains to determine the effects of deleting genes predicted by the global gene expression assays to impact the intracellular reducing potential available for fuel production.

Rhodobacter sphaeroides is a photosynthetic purple non-sulfur bacterium that can accumulate a large pool of intracellular reducing potential (high energy electrons) during photoheterotrophic growth on various carbon sources, such as organic acids and biomass-derived sugars. We are studying the distribution of this reducing potential in the interest of increasing the fraction that goes toward fuel production. We are initially focusing on H₂ production, for its own utility as a fuel and in industrial purposes, and as a model to understand how electrons needed for production of other fuels will be portioned to other pathways.

We have obtained global transcript levels in cultures with various levels of H₂-production and compared them to those in non-H₂ producing control cultures to determine candidate gene products that contribute electrons to or siphon reductant from H₂ production. We also compared transcript levels in H₂-producing cultures grown on various carbon sources to characterize the relative expression levels of the genes involved in cellular pathways that act as electron sinks. In addition, we are assaying these same cultures for the

distribution of electrons between the cellular end-products, PHB, other biomass, H₂, and soluble microbial products.

From prior knowledge, for reasons outlined below, we expected four systems would impact the amount of intracellular reducing potential available for fuel production: nitrogenase, hydrogenase, the carbon dioxide fixation pathways, and PHB synthesis.

- Nitrogenase is the primary (or sole) source of H₂, produced as a byproduct of nitrogen fixation. As expected, expression of the nitrogenase structural genes is increased in all cultures that produce detectable H₂. We are testing whether expression level correlates to total H₂ production amounts; preliminary evidence suggests that cellular nitrogenase enzymatic activity does correlate with the amount of H₂ produced.
- Hydrogenase is expected to oxidize H₂ and recycle reducing potential back into cellular metabolism. We find that expression of the hydrogenase structural genes varies inversely with net H₂ production; cultures that produce relatively low amounts of H₂ (those using xylose, glucose or glycerol) have relatively high expression of hydrogenase genes, while cultures that produce relatively high amounts of H₂ (those using lactate or succinate) have relatively low expression of hydrogenase genes, which suggests that the presence of functional hydrogenase may contribute to the disparities in the amounts of H₂ produced from different carbon sources. Experiments are in progress to determine the level of functional hydrogenase enzyme made in these cultures and how the reducing potential is distributed in defined hydrogenase mutants.
- Carbon dioxide fixation consumes reducing power, so it has the potential to siphon electrons from fuel production. We find that expression of carbon fixation genes is also inversely proportional to H₂ production (as is the case for hydrogenase genes), which suggests that carbon dioxide fixation may contribute to the disparities in H₂ amounts from cultures grown on different carbon sources. *R. sphaeroides* has two sets of carbon fixation genes, and we are investigating mutants which only contain one gene from each set (*cfxA*- and *cfxB*-) to determine the relative importance of the individual pathways on H₂ production.
- *R. sphaeroides* can accumulate the polymer PHB as an energy and carbon storage compound. We find little difference in expression of PHB synthase genes in H₂-producing cultures grown on different carbon sources. However, we find that the amounts of PHB produced differ between such cultures, and we are testing the effect of deleting the PHB synthase gene on H₂ production.

20

Use of Proteomics Technologies for the Characterization of Proteins, Microbes, and Microbial Communities Important for Bioenergy Production

Mary S. Lipton* (mary.lipton@pnl.gov), Stephen J. Callister,¹ Thomas Taverner,¹ Kim K. Hixson,² Samuel O. Purvine,² Kristin E. Burnum,¹ Roslyn N. Brown,¹ Angela D. Norbeck,¹ Janani Shutthanandan,¹ Matthew E. Monroe,¹ Ronald J. Moore,¹ Therese R. Clauss,¹ Carrie D. Nicora,¹ Shishir Chundawat,³ Dahai Gao,³ Venkatesh Balan,³ Bruce Dale,³ Tim Donohue,⁴ Eva Ziegelhoffer,⁴ Audrey Gasch,⁴ Jeff Lewis,⁴ Robert Landick,⁴ Patricia Kiley,⁴ Daniel Noguera,⁴ Cameron R. Currie,⁴ Garrett Suen,⁴ and Jarrod J. Scott⁴

¹Biological Sciences Division and ²Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Wash.; ³Michigan State University, East Lansing; and ⁴University of Wisconsin, Madison

Project Goals: This project employs the state of the art proteomics facilities resident at the Pacific Northwest National Laboratory for the rapid global determination of protein identification, and protein expression patterns in microbes or plants. Ranging from protein preparations purified from fungal and bacterial sources that demonstrate the ability to degrade lignocellulosic material to the quantitative proteomic profiling of microbes and microbial communities, the proteomics capability at PNNL is employed at producing data to further the understanding of systems important to the GLBRC.

Microbial processing for substrates to biofuels, whether the conversion of lignocellulosic material to ethanol or microbial biorefineries to produce hydrogen or electricity is a central part of the GLBRC mission. Inherent in the use of microbes for these purposes is the characterization of the fundamental machines of the cell, the proteins, and how these proteins dictate microbial function. The quantitative determination of protein expression patterns and how these patterns change with changing cell state is critical for the GLBRC to remove bottlenecks in the biofuels pipeline. Additionally, accurate measurements of protein levels and modifications will provide more extensive insights into both the plants and the microbes in the bioenergy pipeline. These analyses include temporal profiling through the analysis of time course studies, characterization of posttranslational modifications, and determination of subcellular localization of proteins. Extension of proteomic capabilities to community profiling will enable deeper understandings of how microbes interact with each other in environmental settings.

In the past year, the proteomics facility has supported the GLBRC in four aspects. The characterization of cellulolytic and hemicellulolytic rich enzyme cocktails found that the dominant cellulases were CBH I, Xyloglucanase, CBH

II, EG I, EG II, EG III, β -glucosidase. The dominant hemicellulases found in most enzyme cocktails were Endoxylanase (GH 11), β -xylosidase, arabinofuranosidase (GH 62 & 54) and Glucuronidase (GH 67). This data will help determine critical classes of cellulases and hemicellulases necessary for hydrolyzing lignocellulosic biomass and are currently absent in commercially available mixtures. Supplementation of a minimalist and optimal enzyme set for hydrolyzing ammonia fiber expansion (AFEX) treated biomass (i.e. corn stover) will help reduce the total number and amount (mg protein/gm substrate) of enzymes required for hydrolysis.

Quantitative analysis of protein abundance in cells is one method for illustrating the manner in which cells perform function or adapt to their environment. The proteomics facility is applying label free proteomics for the quantitative characterization of many bacterial and fungal systems to understand ethanol tolerance, ethanol production and hydrogen production. In *E. coli*, analysis of aerobic and anaerobic cell cultures yielded a combined total of 1697 proteins identified from the two cultures using strict cross correlation and cutoff values for the peptides and the requirement of two peptides per proteins. Of the 1697 proteins identified, 46 proteins were found in the anaerobic cultures only, 30 peptides were found in the aerobic cultures only, and 1621 proteins were identified in both cultures. Quantitative analysis of 1254 proteins between both of the samples showed 86 proteins showed at least a two-fold increase in abundance in the aerobic sample and 56 showed at least a two-fold increase in abundance in the anaerobic sample. The qualitative and quantitative characterization of *Rhodobacter sphaeroides* proteome showed an increase in abundance of nitrogen fixation genes when the organism was grown to stationary phase when compared with log phase. In yeast, differential protein expression patterns are being used to determine the mechanism behind ethanol tolerance.

Expansion of the proteomics from microbes into microbial communities can provide insights into many of the biological mechanisms present in the communities as well as information about strain heterogeneity depending on the size and quality of the sequenced genome. We are employing proteomic analyses to gain insight into the symbiotic relationship between leafcutter ants, fungi, and bacteria. This fungus digests the cellulose in the leaves and uses it to fuel its own growth. The leaf-cutter ants then feed on the fungus. The bacterial community found in this system is believed to protect the fungus from parasites and aid in cellulose digestion. The metagenome of the fungus garden, the genome of the leafcutter ant, and the genome of the fungus have been sequenced by the DOE Joint Genome Institute and the proteomics capabilities at PNNL will be used to identify protein expression within this system. Following high sensitivity and mass accuracy capillary LC-MS/MS measurements of trypsin-digested proteomes, uninterpreted tandem MS spectra will be compared to potential bacterial, fungus and ant protein sequences using Sequest. We believe our metaproteomic analyses could identify novel proteins with applications toward biofuel and antibiotic development.

This research is supported by the Office of Biological and Environmental Research of the U.S. Department of Energy. Portions of this research were performed in the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the DOE's Office of Biological and Environmental Research. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute through Contract No. DE-AC05-76RLO 1830.

21

Production of Lower Viscosity Oils as Biofuels in Transgenic Plants: Deep Transcriptional Profiling Reveals a Novel Acetyl-CoA Diacylglycerol Acetyltransferase from *Euonymus alatus*

John Ohlrogge^{1*} (Ohlrogge@msu.edu), Timothy Durrett,¹ Adrian Troncosco,¹ Jillian Fan,¹ Xia Cao,¹ Nick Thrower,¹ Erika Lindquist,² Christa Pennacchio,² Curtis Wilkerson,¹ and Mike Pollard¹

¹Dept. of Plant Biology and DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing and ²DOE Joint Genome Institute, Walnut Creek, Calif.
<http://www.greatlakesbioenergy.org/>

Project Goals: Design of improved plant oil structure for fuel use. The high viscosity of plant oils causes problems with standard diesel engines. Therefore, most biodiesel is produced by conversion of triacylglycerol to methyl or ethyl esters. Acetyl-glycerides are abundant in some plant species and represent a form of triacylglycerol with acetyl rather than acyl groups at sn-3. This structure results in oil with predicted lower viscosity that should directly replace diesel fuel #4 without need for transesterification. The goal of this study is to identify genes involved in biosynthesis of this valuable type of plant oil.

Unlike microarrays, EST sequencing provides a method of transcript analysis that allows quantitative comparisons between genes and between different plant species. In order to identify genes involved in plant oil biosynthesis and the transcription factors and other regulatory systems that control oil accumulation, Michigan State, together with JGI has sequenced over 10 million ESTs from a variety of oil-seeds and other oil rich tissues. Why do we need millions of ESTs? Key enzymes of lipid metabolism (e.g. acyltransferases, phospholipases, thioesterases) are very low abundance and can be difficult to detect by conventional EST sequencing. Deep EST sequencing using 454 pyrosequencing provides a large increase in EST sequence information and allows us to accurately quantify low level expression. By sequencing libraries from multiple species we obtain information on what similarities and differences distinguish oil synthesis in seeds producing unusual fatty acids and in seeds compared to other tissues such as mesocarp that produce high oil levels. Replicate analysis of samples (including cDNA synthesis and PCR) gave a 0.99 correlation coefficient between #

reads per gene. Therefore, 454 sequencing is technically and biologically reproducible and provides an accurate measure of gene expression. We have observed that core enzymes of fatty acid biosynthesis are, in general, expressed in consistent stoichiometric ratios in a number of different oilseeds and tissues. Therefore, those genes that fall outside the usual stoichiometry offer insight into unique metabolism. For example, we observe very low expression of the FatB thioesterase that controls saturated fatty acid production in castor, which agrees with the fact that castor is an oilseed with extremely low saturated fatty acid content.

Endosperm tissue from *Euonymus alatus* (Burning Bush) accumulates high levels of 3-acetyl-1,2-diacyl-*sn*-glycerols (ac-TAGs) as the major storage lipids. Ac-TAGs are unusual triacylglycerols (TAGs) with an *sn*-3 acetate group instead of a long-chain fatty acid and have added value applications in direct use as biodiesel and lubricant oil feedstocks. In addition to producing ac-TAGs, *Euonymus* fruit also synthesizes normal, long-chain TAGs (lc-TAGs) in their aril tissue. The close developmental coordination and spatial proximity of two tissues with the ability to produce different TAGs presents a unique opportunity to understand the accumulation of unusual TAGs in plants. By sequencing ESTs from these tissues we have identified candidate genes involved in ac-TAG biosynthesis. One such candidate, subsequently named EaDacT (*Euonymus alatus* diacylglycerol acetyl-transferase) was highly expressed in the endosperm and absent from the aril. Expression of EaDacT in yeast resulted in the accumulation of ac-TAGs, but not lc-TAGs. *In vitro* assays with microsomes from yeast expressing EaDacT demonstrated that the enzyme possesses acetyl-CoA diacylglycerol acetyltransferase activity, but not long chain acyl-CoA diacylglycerol acyltransferase activity. Expression of EaDacT in *Arabidopsis* seed caused the accumulation of up approximately 40% of ac-TAGs in the seed oil. These results demonstrate that EaDacT synthesizes ac-TAGs in *Euonymus* endosperm and illustrate the utility of deep transcriptional profiling as a gene discovery platform for modifying the seed oil properties of plants.

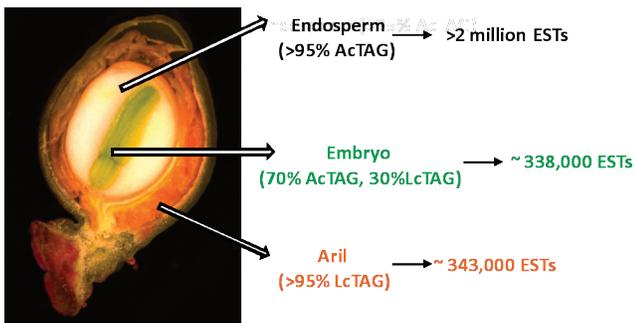


Fig 1. *Euonymus* fruit produce ac-TAGs in a tissue specific manner. The endosperm and embryo tissues of the *Euonymus* seed accumulate high levels of ac-TAGs whereas the aril tissue surrounding the seed synthesizes only lc-TAGs. Transcript profiles were obtained of these different tissues to isolate the enzyme(s) responsible for the synthesis of ac-TAGs.

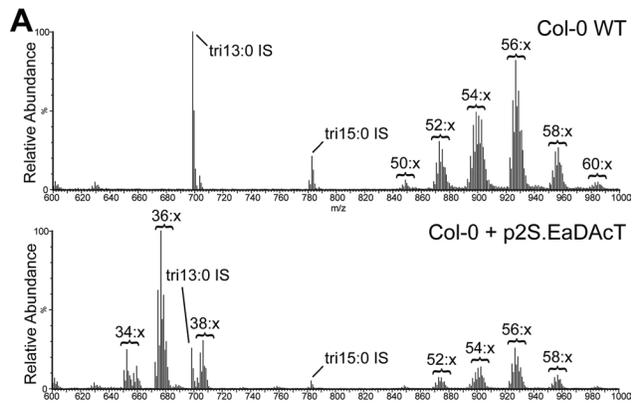


Fig 2. *Arabidopsis* seeds expressing EaDacT accumulate ac-TAGs. Positive-ion ESI mass spectra of neutral lipid extracts from Col-0 wildtype seed or T₃ seed from a representative Col-0 plant expressing EaDacT. Peaks correspond to *m/z* values of the [M + NH₄]⁺ adduct. Tridecanoin (tri13:0) and tripentadecanoin (tri15:0) were added as internal standards. The number of acyl carbons in each series of TAG molecules is indicated.

22 Sustainable Production of Fatty Acid Derived Fuels and Chemicals in Engineered Microorganisms

Rebecca M. Lennen,^{1,2} Spencer Hoover,^{1,2} Tyler Youngquist,^{1,2} and Brian F. Pfleger^{1,2*} (pfleger@engr.wisc.edu)

¹Dept. of Chemical and Biological Engineering, University of Wisconsin, Madison; ²U.S. Department of Energy Great Lakes Bioenergy Research Center, Madison, Wisconsin

Project Goals: (1) Develop and demonstrate metabolic engineering strategies for assembling fatty acid overproducing microorganisms. (2) Identify barriers that limit the production of fatty acids using processed biomass as a substrate. (3) Develop strategies to convert fatty acids into useful products, including fuels and chemicals.

The development of renewable alternatives to diesel and jet fuels is highly desirable for the heavy transportation sector, and would offer benefits over the production and use of short-chain alcohols for personal transportation. Here we report the development of a metabolically engineered strain of *Escherichia coli* that overproduces medium-chain length fatty acids via three basic modifications: elimination of β-oxidation, overexpression of the four subunits of acetyl-CoA carboxylase, and expression of a plant acyl-acyl carrier protein (ACP) thioesterase from *Umbellularia californica* (BTE). The expression level of BTE was optimized by comparing fatty acid production from strains harboring BTE on plasmids with four different copy numbers. Expression of BTE from low copy number plasmids resulted in the highest fatty acid production. Up to a seven-fold increase in total

fatty acid production was observed in engineered strains over a negative control strain (lacking β -oxidation), with a composition dominated by C12 and C14 saturated and unsaturated fatty acids. Next, a strategy for producing undecane via a combination of biotechnology and heterogeneous catalysis is demonstrated. Fatty acids were extracted from a culture of an overproducing strain into an alkane phase and fed to a Pd/C plug flow reactor, where the extracted fatty acids were decarboxylated to saturated alkanes. The result is an enriched alkane stream that can be recycled for continuous extractions. Complete conversion of C12 fatty acids extracted from culture to alkanes has been demonstrated yielding a concentration of 0.44 g L⁻¹ (culture volume) undecane.

23

Sustainably Filling the Field to Fuel Pipeline: A GLBRC Research Priority

P. Jasrotia,¹ **G.P. Robertson**¹ (robertson@kbs.msu.edu), J.M. Tiedje^{2*}, K.D. Thelen,² S.K. Hamilton,¹ R. Jackson,³ D.A. Landis,² D.W. Schemske,² B.E. Dale,² S.M. Swinton,² R.C. Izaurralde,⁴ and P.J. Meier³

¹W.K. Kellogg Biological Station (KBS), Hickory Corners, Michigan; ²Michigan State University, East Lansing; ³University of Wisconsin, Madison; and ⁴University of Maryland, College Park

Project Goals: GLBRC research to improve biofuels sustainability overall objective: Support the biomass-to-bioenergy pipeline by developing ecological, agricultural, and life cycle practices that are economically viable and environmentally responsive. Evaluate and improve for different biofuel crops. 1. carbon neutrality and net greenhouse gas mitigation across the entire biofuel life cycle at multiple scales. 2. ecosystem services in biofuel landscapes (e.g., water quality, biodiversity, pest suppression).

Biofuels are attractive for economic, environmental, and strategic reasons. Reducing our dependency on foreign oil is a key national security issue. Although corn and soybeans are excellent sources of biofuels, cellulosic feedstocks clearly are the wave of the future. Poplar, switchgrass, *Miscanthus*, and even managed prairie ecosystems are potential sources of cellulosic biofuels. If we can solve the problems of effectively transforming this biomass into cellulosic ethanol or other fuels, there will be significant environmental positive gains in reducing emissions of greenhouse gases and other pollutants. However, concerns have been raised about the sustainability of cellulosic cropping systems. Biofuel production systems based on annual grains are the most straightforward for growers but may not be the most productive or sustainable in the long term. Understanding the basis for sustainable biofuel production systems is crucial for the long-term success of these systems. The rapid growth, low mineral content and high biomass yield of cellulosic crops make them a favored feedstock choice. But the important question remains: can we produce enough biomass to sustain

the biofuel industry without compromising environmental security. The biomass production potential of these crops is directly linked to concerns related to land-use change, environmental degradation, and food security. To answer these questions, field experiments are underway at several scales in Michigan and Wisconsin. Eight model cropping systems are studied for productivity, carbon balance, and biodiversity impacts. A major goal of this initiative is to test and develop biofuel crops that provide both high yields and environmental benefits such as greenhouse gas mitigation, clean water, and pest protection. Our aim is to provide a comprehensive portrait of the production potential of different cellulosic biofuel crops as well as their environmental impacts and benefits. We use field results to parameterize and test quantitative models that can then be used to simulate crop growth across larger regions.

24

Metagenomics of Bacterial Communities from the Rhizosphere of Switchgrass

Ederson da Conceição Jesus, Endang Susilawati, Benli Chai, Qiong Wang, Jorge Rodrigues, and **James M. Tiedje*** (tiedje@msu.edu)

Michigan State University, East Lansing

Project Goals: The specific goal of our group is to assess the structure and functional diversity of the rhizosphere community associated with biofuel energy crops by applying metagenomic approaches.

Switchgrass is one of the main plants being studied for biofuel production in the United States. Besides being indigenous and producing a large amount of biomass, it is able to grow in marginal lands, a feature that may be associated not only with plant genotype, but also with their association to beneficial microbes. Managing these microbes can help to improve biomass production and decrease production costs. For this reason, we are studying bacterial communities in the rhizosphere of switchgrass through a metagenomic approach, which allows us to assess not only the taxonomic composition of the bacterial communities in the rhizosphere, but also their functional composition. Additionally, we aimed to validate the application of metagenomics to study the rhizosphere of switchgrass. Roots of the cultivar Cave-in-Rock grown at the MSU experimental farm were sampled in 2008 from the 0-20 cm depth. After removing the excess soil, the rhizosphere soil was recovered and used for total DNA extraction with the method developed by Zhou et al. (1996). The DNA was analyzed by 454-Titanium pyrosequencing at JGI, generating 291 Mb of sequence in 642441 reads with an average length of 453 bp. Potential artificial duplicates were removed (Gomez-Alvarez et al., 2009) and the assembled data set was used for analysis with the MG-RAST pipeline. The data set without duplicates contains 487,660 contigs totaling 237,422,518 basepairs with an average fragment length of 486.86 bp. A total of 297,143 sequences (60.93%) could be

matched to proteins in SEED subsystems (using an e-value cut-off of $1e^{-5}$). Most of the reads were assigned to essential functions for cell maintenance such as amino acids and carbohydrate metabolism. The reads were assigned mainly to Proteobacteria (74%), followed by Acidobacteria (4%), Actinobacteria (3%) and Bacteroidetes (2%). Pseudomonadaceae assignments accounted for 54% of all reads. In fact, 140,697 fragments of the metagenome map to 5,233 of 5,858 features from the *Pseudomonas fluorescens* PfO-1 genome and the total base pair length of sequences matching this genome results in approximately 2.6X coverage. One lane of Illumina sequence has also recently been provided by JGI for the same DNA sample, and is being used to evaluate both the enhancements by this technology, as well as deeper insight into the gene biology of the rhizosphere. Functions commonly associated with *Pseudomonas* such as stress resistance and iron scavenging were also found. *Pseudomonas* was also found in large numbers by cultivation from the same rhizosphere sample, providing cultured models for further study. We conclude that potentially beneficial microbes, such as *Pseudomonas* are present in the rhizosphere of switchgrass and that the metagenomic approach is allowing us to obtain information about the functionality of rhizosphere communities.

25

Education and Outreach Opportunities Linked with the Research and Development of Sustainable Cellulosic Biofuels

John M. Greenler^{1*} (jgreenler@glbrc.wisc.edu), Sara J. Krauskopf,¹ Steven C. Slater,¹ Timothy J. Donohue,^{1,2} Richard M. Amasino^{1,3}

¹U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center (GLBRC), ²Dept. of Bacteriology, and ³Dept. of Biochemistry, University of Wisconsin, Madison

<http://www.greatlakesbioenergy.org/education/>

Project Goals: Develop education modules and outreach programs for K-16 communities on energy, carbon chemistry, sustainability (environmental, economic and social) issues and other areas related to biofuel production and use; offer bioenergy-focused programming for students and educators; develop informative materials to raise awareness of biofuels and related sustainability issues among the general public, extension staff, policy makers and industry representatives; provide a venue to collaborate on the development of bioenergy education and outreach materials; and present general information about the GLBRC's mission and accomplishments to interested parties.

There is currently a significant research effort to develop sustainable biofuels from cellulosic plant materials. To be broadly sustainable, this research utilizes a diverse array of modern scientific methods, and is very interdisciplinary

and collaborative in nature. Given that this work is in the public eye, there are significant opportunities to engage learners in both the details of bioenergy; and importantly, into the underlying scientific principles of biogeochemical and energy systems. Assessments of student understanding, however, reveal many consistent misconceptions that hinder students' ability to comprehend these systems.

A significant focus of our work is to develop 1) a further categorization of the range of understandings related to carbon cycles and energy flow, and 2) K-16 educational materials that will use biofuels as an entrée to engage learners in a more accurate comprehension of these basic scientific concepts. Specific content at the base of these materials includes life-cycle assessments and systems thinking. The effects of biofuel production on biodiversity, ecosystem services, climate change, and global energy dynamics are considered as well. A number of persistent difficulties are seen in assessments of student understanding at all levels. Many students, for instance, do not connect the decomposition of plants and animals to CO₂ in the atmosphere. In response to this and related patterns, we are designing a series of activities to make connections to fundamental biological concepts with which they are more familiar. Associated with the development of these educational materials are summer research experiences for undergraduates and teachers.

submitted post-press

Functional Annotation of *Fibrobacter succinogenes* Carbohydrate Active Enzymes

David Mead (dmead@lucigen.com), Julie Boyum, Colleen Drinkwater, Krishne Gowda, David Stevenson, Paul Weimer, and Phil Brumm

Lucigen and C56 Technologies, Middleton, Wisc.; Great Lakes Bioenergy Research Center, University of Wisconsin, Madison; and USDA-Agricultural Research Service, Madison, Wisc.

Project Goals: Eliminate bottlenecks in plant cell wall deconstruction.

Fibrobacter succinogenes is a predominant cellulolytic bacterium that degrades plant cell wall biomass in ruminant animals, and is among the most rapidly fibrolytic of all mesophilic bacteria. This gram-negative, strictly anaerobic bacterium is also of interest because it does not utilize either freely secreted enzymes or cellulosomal-type structures to digest cellulose. A dozen cellulolytic enzymes have been expressed and characterized previously, and an outdated partial genome sequence indicates that there are at least 33 unique glycosyl hydrolases encoded by *F. succinogenes*. In order to better understand plant cell wall degradation we have developed new tools to capture, express and identify many of the carbohydrate active enzymes (CAZymes) from this microbe. The complete genome sequence of *Fsu* was finished by the DOE Joint Genome Institute in late 2009, contributing to the growing database of cellulolytic

microbes. Preliminary analysis indicates that *F. succinogenes* contains ~ 133 glycosyl hydrolase and 63 CBM-containing genes, the most of any microbe when expressed as a percent of the total gene number.

Based on the genomic sequencing results, the number of *F. succinogenes* genes annotated as CAZymes far exceed those that have been experimentally determined by conventional enzymatic approaches. One of the goals of this work is to functionally characterize all the putative glycosyl hydrolase genes from *Fsu*, as bioinformatic analysis is an inadequate proxy for actual activity results. Before the genome sequence was available we developed a robust method to enzymatically capture functionally active CAZymes in *E. coli*. Using new expression tools developed at Lucigen and C5-6 Technologies and a multi-substrate screen for xylosidase, xylanase, β -glucosidase and cellulase activities, we generated and screened 5760 random shotgun expression clones for these activities. This represents ~ 2 X genome expression coverage. 169 positive hits were recorded and 33 were unambiguously identified by sequence analysis of the inserts. Eliminating duplicates, 24 unique CAZyme genes were found by functional screening, or 40% of the ~60 genes present in this genome potentially detectable by the multiplex assay. Several previously uncharacterized enzymes were discovered using this approach. With the full genome sequence available we will attempt to express and characterize all of the recognizable CAZymes, as well as the CBM-containing genes for actual enzyme activity. The active enzymes will also be sent to other partners in the GLBRC to assess their ability to deconstruct plant biomass.

Joint BioEnergy Institute (JBEI)

26

The Joint BioEnergy Institute: Addressing the Challenges of Converting Biomass to Fuels

J.D. Keasling* (JDKeasling@lbl.gov), H.W. Blanch, P.D. Adams, P.C. Ronald, and B.A. Simmons

Joint BioEnergy Institute, Emeryville, Calif.

Project Goals (Abstracts 26-54): In the San Francisco Bay Area, three national laboratories, major public and private universities, industry, and federal agencies have joined together to create the Joint BioEnergy Institute (JBEI). This institute will develop the basic science and technology to create an array of environmentally friendly biofuels using plant biomass and microbes. JBEI will focus its scientific effort in three key areas: feedstock production, deconstruction, and fuels synthesis. JBEI will employ an opportunistic "start-up company" approach, partnering with industry, to develop new science and technologies that address the most challenging steps in industrial

bioenergy processing. Crosscutting technologies in computational tools, systems and synthetic biology tools, and advanced imaging will be applied in a multi-pronged approach for biomass-to-biofuel solutions in addition to discovery-driven benefits for biohydrogen research, solar-to-fuel initiatives, and broader DOE programs.

In the San Francisco Bay Area, three national laboratories, major public and private universities, industry, and federal agencies have joined together to create the Joint BioEnergy Institute (JBEI). This institute is designed to address the mission of the DOE Bioenergy Research Center program: "to produce fundamental scientific discoveries and major technological advances to enable the development of cost-effective, energy-efficient, and commercially viable processes for large-scale conversion of lignocellulosic biomass into fuels." This institute is developing the basic science and technology to produce fuels from plant biomass by microbial routes. JBEI is focusing its scientific effort in three key areas: elucidating cell wall biosynthesis, lignocellulose deconstruction, and fuels biosynthesis, while employing cross-cutting technologies throughout all of its research.

There are key challenges in each of the institute's scientific divisions. The Feedstocks Division is developing an understanding of hemicellulose biosynthesis and analyzing the recalcitrance of plant cell walls to deconstruction. By modifying lignin and reducing acetate and ferulate content, crops will be better suited to biofuels production. In the Deconstruction Division, new approaches to biomass pretreatment are being developed that result in solubilization and separation of plant cell wall components. Enzymatic hydrolysis of pretreated biomass represents a significant cost in formation of saccharide monomers; efforts are underway to identify and produce new enzymes from unique environments that are more effective.

In the Fuels Synthesis, Division, a challenge is to produce advanced biofuels needed for diesel and jet engines. Many fuel-producing organisms incompletely metabolize C₅ (xylose, arabinose) sugars, while inhibitors (e.g., acetate) released from biomass pretreatment can limit fuels production. Introducing new metabolic routes to advanced biofuels relies on Synthetic Biology approaches. Few tools are available for biomass conversion research. The Technologies Division is developing high-throughput 'omics and microfluidic approaches for many aspects of bioenergy research, as well as high-throughput biochemical, spectroscopic, and imaging methods for the rapid characterization of biomass.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

27

Exploiting Natural Variation in *Arabidopsis thaliana* to Understand Cell Wall Biosynthesis and Composition

Katy M. Christiansen* (KMChristiansen@lbl.gov),
A. Michelle Smith, Dominique Loque, and **Joshua L. Heazlewood**

Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Project Goals: See goals for Abstract 26.

Genetic variation in different accessions of *Arabidopsis thaliana* has occurred through thousands of years of geographic spread and adaptation. As a result of this diversity, disparate accessions have phenotypic differences that can be used to identify genes that contribute to the biosynthesis and composition of cell walls. Forward and reverse genetic screens to identify single mutants often prove difficult for detection of genetic differences that may lead to more subtle phenotypes. Exploiting the inherent genetic variation in *Arabidopsis* accessions through quantitative trait analysis will allow for the detection of variation in cell wall biosynthesis and composition. We selected two methods for screening parental accessions: measuring monosaccharide composition by HPAEC and structural changes by Near Infrared (NIR) spectroscopy. From these analyses, we selected Ri-0 as the most different accession from the reference accession Col-0. Recombinant inbred lines (RILs) derived from Col-0 and Ri-0 parents were used to determine quantitative trait loci (QTL) that contribute to the differences observed in monosaccharide and pectin content and NIR spectroscopy. Two putative candidate genes encoding enzymes involved in nucleotide sugar conversion have been identified. We are currently determining differences between the parental versions of the genes and how any differences contribute to the observed phenotypes. Additionally, we have re-sequenced two *Arabidopsis* accessions, Bay-0 and Shahdara, in collaboration with the Joint Genome Institute. This effort has resulted in a collection of SNPs between these accessions and the reference *Arabidopsis* accession, Col-0. Resequencing efforts are being expanded to other accessions, beginning with Ri-0. Information obtained from resequencing will aid in QTL analysis and be of service to the general *Arabidopsis* community.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

28

Towards the Plant Golgi Proteome

Harriet T. Parsons,¹ Christopher J. Petzold,¹ Alyssa Redding-Johanson,¹ Tanveer Batth,¹ and **Joshua L. Heazlewood*** (JLHeazlewood@lbl.gov)

¹Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

The plant cell wall is comprised of complex sugar polymers including cellulose, hemicellulose and pectin. The Golgi apparatus within the plant cell produces a large proportion of these polysaccharides prior to their incorporation into the cell wall. The details of how these complex sugars are synthesized and delivered are currently poorly understood. Our current knowledge of the protein constituent that comprise the plant Golgi is relatively poor when compared to other subcellular components within the cell. In order to further understand the role of this organelle in cell wall biosynthesis we are characterizing this subcellular compartment using proteomics. We are employing an orthogonal approach which utilizes density centrifugation followed by charge based separation of the organelle on a Free Flow Electrophoresis system. Analysis of Golgi purified fractions from *Arabidopsis* cell culture by mass spectrometry after FFE separation indicates the method is suitable for isolation of this organelle from plants. We have identified 300 to 400 proteins from these fractions and found approximately 50 glycosyl transferases likely involved in matrix polysaccharide biosynthesis. Overall approximately 50% of the proteins in this list are of known or likely Golgi in origin; approximately 35% are unknown or are derived from the endosomal system and 15% appear to be contaminants from other organelles and membranes. Fluorescent protein constructs are being designed to confirm localizations of novel and ambiguous proteins. This technique will enable us to commence in-depth comparative cell wall proteomics focusing on protein function and changes and in the Golgi apparatus.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

29

Synthesis of Phenylpropanoid-Esters and -Amides in *Arabidopsis thaliana* to Engineer a Cleavable Lignin

Aymerick Eudes, Prajakta Pradhan, Fan Yang, and
Dominique Loqué* (DLoque@lbl.gov)

Feedstocks Division, Joint BioEnergy Institute,
 Emeryville, Calif.

Project Goals: See goals for Abstract 26.

The development of alternative transportation fuels that can meet future demands while reducing global warming is critical to the national, environmental, and economic security of the United States. Currently, biofuels are produced largely from starch, but there is a large, untapped resource (more than a billion tons per year) of plant biomass that could be utilized as a renewable, domestic source of carbon-neutral, liquid fuels. However, significant roadblocks hamper the development of cost-effective and energy-efficient processes to convert lignocellulose biomass into fuels. Lignin is a very strong phenolic polymer, which embeds cellulose and hemicellulose, and its recalcitrance to chemical and biological degradations inhibits the conversion of cell wall polysaccharides (cellulose and hemicellulose) into fermentable sugars. Unfortunately, lignin provides such compressive resistance to plant cells that it cannot simply be genetically removed without incurring deleterious consequences on plant productivity. Alternative strategies to significantly reduce lignin recalcitrance would be modifying its composition and deposition. We are currently developing an alternative strategy, which is focusing on the partial replacement of the “hard bonds” (e.g. ether, carbon bonds) in the lignin polymer with “easily cleavable” ones (e.g., amide or ester bonds). For this propose, we are rerouting part of the lignin biosynthesis towards the synthesis of phenylpropanoid-derived molecules such as hydroxycinnamic acid amides and esters in order to partially replace conventional lignin monomers in the cell wall. Biosynthetic pathways and preliminary data for de novo synthesis in *Arabidopsis* of selected phenylpropanoid-derived compounds are presented.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

30

Elucidating Switchgrass Genome Structure and Function of Cell Wall-Related Enzymes

M.K. Sharma, R. Sharma* (RASharma@lbl.gov),
 L.E. Bartley, M.D. Zorn, P. Cao, E. Marvinney, and
P.C. Ronald

Joint BioEnergy Institute, Lawrence Berkeley National
 Laboratory, Berkeley, Calif.

Project Goals: See goals for Abstract 26.

JBEI, in collaboration with other DOE centers, is organizing and supporting the sequencing the genome of switchgrass, *Panicum virgatum* L. clone AP13. Our goal is to elucidate the genome content to facilitate comparative functional genomics studies of this promising biofuel feedstock. Whole genome shotgun sequencing using 454-Titanium technology is underway at the DOE-Joint Genome Institute (JGI). The first phase of switchgrass sequencing will be used to produce approximately 4x sampling of the AP13 genome in 400 bp 454 reads and approximately 30x sampling in 75 bp Illumina reads. We are also using various approaches that include sequencing of hypomethylated restriction libraries, and BAC (Bacterial Artificial Chromosome) and fosmid end sequencing to efficiently cover gene space and establish long-range connectivity among assembled shotgun sequence contigs, respectively. One AP13 BAC library has been synthesized and characterized for mitochondrial/chloroplast DNA contamination as well as coverage by high-density filter hybridizations. We are using available information about genes associated with cell wall biosynthesis, degradation, biomass production and stress tolerance from other grass species to identify homologous ESTs in switchgrass and screen corresponding BACs for in depth characterization. The first ten BACs have been selected and are being sequenced at JGI. We have also selected BACs containing single copy genes to make direct comparisons between homeologous chromosomes. Due to the high level of colinearity among grass genomes, we can use sorghum and foxtail millet as reference genomes in assembling the switchgrass sequence data. With the completion of whole genome sequence, cell wall related genes will be computationally identified and comprehensive phylogenomic analysis with other grasses and dicots will be carried out to identify grass-specific and switchgrass-specific genes involved in cell wall metabolism. Already, Initial phylogenomic analysis with switchgrass ESTs has shown that the CslG family previously considered as dicot-specific, are present in the switchgrass genome. In coordination with JGI we are developing a high quality annotation pipeline and databases of grass cellwall-osome sequences to support comparative functional genomics in grasses. Our own and others' 454 sequence ESTs (400 bp reads) and approximately 500,000 Sanger ESTs (800 bp reads) in public repositories will assist us in producing gene inventories and gene annotation.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy,

Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

31

Identification of Genes Involved in Acetylation of Cell Wall Polysaccharides in *Arabidopsis thaliana*

Yuzuki Manabe^{1*} (YManabe@lbl.gov), Yves Verherbruggen,¹ Majse Nafisi,² Chithra Manisseri,¹ Peter Benke,¹ Aindrila Mukhopadhyay,¹ Caroline Orfila,² Yumiko Sakuragi,² and **Henrik Vibe Scheller**¹

¹Joint BioEnergy Institute, Emeryville, Calif. and

²Dept. of Plant Biology and Biotechnology, University of Copenhagen, Denmark

Project Goals: See goals for Abstract 26.

Acetylation of cell wall polysaccharides has long been observed in various plant species; however, the enzymes involved in the acetylation have thus far not been identified. Both pectins and hemicelluloses are acetylated to various degrees. While the *in vivo* role of polysaccharide acetylation is still unclear, it is known to affect biofuel yield from lignocellulosic biomass due to inhibition of enzymatic degradation of the polysaccharides and fermentation by microorganisms. Therefore, a decreased level of acetate esters in lignocellulosic biomass may increase the efficiency of biofuel production. JBEI is therefore investigating the mechanism of polysaccharide acetylation and assessing the possibility of modifying acetylation level *in planta*.

We have analyzed four *Arabidopsis* homologues of a protein known to be involved in polysaccharide acetylation in a fungus. *Arabidopsis* mutants with insertional mutagenesis in the respective genes were identified, and we found that at least one of the mutants, designated *reduced wall acetylation* (*rwa1*, *rwa2*, *rwa3* and *rwa4*) had decreased levels of acetylated cell wall polymers. Two independent alleles of *rwa2* mutants were examined by analyzing alcohol insoluble residues extracted from leaves. Extracts treated with 0.1M NaOH released about 20% lower amounts of acetic acid when compared to wildtype. Interestingly, the monosaccharide composition of the cell wall polysaccharides in *rwa2* was not altered. Current efforts are aimed at determining which acetylated polysaccharides are affected in the *rwa2* mutants. There was no apparent visible difference Scheller observed between wildtype and either allele of mutants at any developmental stages. However, both alleles of *rwa2* have displayed increased resistance toward *Botrytis cinerea*, a necrotrophic fungus. The other mutants, *rwa1*, *rwa3*, and *rwa4*, did not have detectable changes in acetylation, presumably due to genetic redundancy. Double, triple and quadruple mutants are currently being investigated.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

32

Forward Genetic Screen to Identify Rice Mutants with Changes in Cell Wall Composition and Saccharification Efficiency

Miguel Vega-Sánchez* (MEVegasanchez@ucdavis.edu), Mawsheng Chern, Wing Hoi Sze-To, Laura Bartley, A. Michelle Smith, Bradley Holmes, Rajiv Bharadwaj, **Blake Simmons**, and **Pamela Ronald**

Feedstocks, Deconstruction, and Technology Divisions, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Understanding plant cell wall biosynthesis is crucial for the development of the next generation of biofuels derived from lignocellulosic material. Current limitations in the harvest of fermentable sugars from cellulose derive from the inherent recalcitrance of plant cell walls. Basic knowledge of how the structure and composition of the cell wall can be modified to obtain biomass suitable for efficient and economically viable biofuel production is needed. We are using a forward genetics approach to identify genes responsible for cell wall characteristics affecting cell wall composition and deconstruction. By means of fast neutron mutagenesis, we have generated a rice mutant population consisting of 6,500 M0 lines and harvested more than 100,000 M1 seed from approximately 4,000 M0 plants. Leaves and stems from these lines have been collected and we are in the process of screening them for alterations in saccharification efficiency and cell wall composition. To screen for changes in fermentable sugar release from biomass, we have optimized a protocol using either hot water or dilute acid pre-treatment followed by enzymatic saccharification for adaptation into a 96 well format. In addition, we are standardizing a high throughput microfluidics platform for analyzing alterations in the C5/C6 monosaccharide ratios of total sugar extracts from leaves to identify mutants with changes in cell wall composition. We have also validated a method for pre-screening intact dried leaf tissue using Near Infrared spectroscopy to identify outliers in the mutant population that will be then analyzed using the microfluidics system. Once cell wall mutants are confirmed, we will extract DNA from wild type and highly prioritized mutant candidates and then carry out whole genome comparative hybridization on rice tiling arrays. This approach will allow us to identify genes in deleted region responsible for the mutant phenotypes. Mutants will be complemented with candidate genes using transgenic analysis and assayed for restoration of the cell wall phenotypes.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

33

Glycosyltransferases (GTs) from the *Arabidopsis* CAZy Family: High-Throughput Cloning of a Library of GT and GT-Related Genes

Peter McInerney,¹ Stephanie Morrison,¹ Hou Cheng Chu,¹ Ai Oikawa,² Henrik Scheller,² **Pam Ronald,² Paul Adams,¹** and Masood Hadi^{1*} (MZHadi@lbl.gov)

¹Technology and ²Feedstocks Divisions, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

The CAZy (Carbohydrate Active EnZyme) family of enzymes includes glycosyltransferases (GTs), glycosylhydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterase (CEs) and carbohydrate binding modules (CBMs). Many enzymes in this family are involved in various aspects of plant cell wall metabolism. The GTs represent one of the most diverse CAZy groups, with 91 separate protein families (not including non-classified sequences) that are assigned based on 3D protein structure, catalytic mechanism and donor/acceptor substrate requirements. In the simplest terms, GTs catalyze the transfer of sugar molecules from a donor molecule to an acceptor. However, the seemingly limitless combination of specific sugar, donor and acceptor molecules underscores the necessity for a large number of enzymes of this family. Our group is undertaking an effort to clone all 455 GTs in the CAZy database from *Arabidopsis thaliana* as well as the 90 GT-like proteins identified from other bioinformatic analyses. This library of GT clones will be a valuable resource at JBEI for a wide range of applications. At the level of biofuels research, results applicable to cell wall engineering are expected, based on the observation that a significant proportion of GT genes play roles (or are proposed to play roles) in cell wall metabolism. Furthermore, from a basic science standpoint, a great deal of new information should result from the study of these genes, since many of these genes are hypothetical or have unknown functions. Our approach relies heavily on automation, for informatics steps such as PCR primer generation and DNA sequence analysis, in addition to laboratory robotics, for assembly of enzymatic reactions and purification steps. To date clones for 80% of the targets have gone through the pipeline and are being sequence verified. Current efforts are centered on maximizing the number of clones that perfectly match the target DNA sequence by optimizing our cloning workflow. Once production of sequence-validated clones is complete, we will transfer genes to vectors suited to specific needs (e.g. expression for biochemical analyses or crystallography trials, fluorescence localization studies, etc.).

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

34

Screening Glycosyltransferases for Enzymatic Activity

Anongpat Suttangkakul^{1*} (ASuttangkakul@lbl.gov), Ai Oikawa,¹ Carsten Rautengarten,¹ Jesper Harholt,² Yves Verherbruggen,¹ Michelle Truong,¹ Peter McInerney,¹ Masood Hadi,¹ Peter Benke,¹ Naomi Geshi,² Jacob K. Jensen,² Eva Knoch,^{1,2} and **Henrik Vibe Scheller¹**

¹Joint BioEnergy Institute, Emeryville, Calif. and ²Dept. of Plant Biology and Biotechnology, University of Copenhagen, Denmark

Project Goals: See goals for Abstract 26.

Declining sources of fossil fuels, global warming and political instability in oil producing regions have led many countries to develop strategies for alternative energy. Plant biomass is a convenient way to harness solar energy and photosynthesis, and biomass is already an important supplement to fossil fuels. However, the energy efficiency of biofuel production is low, and environmental impact can be high. There is a great need to develop new technologies that can provide fuels, especially liquid fuels for transportation, in an efficient and environmentally friendly way.

Plant cell walls are composed mainly of polysaccharides and production of biofuels from biomass requires decomposition of the polymers. Many of the polymers are recalcitrant to degradation and some degradation products cannot be converted efficiently into fuels or may even be inhibitory. Better understanding of the biosynthesis of the cell wall polysaccharides may enable development of crops with improved properties as biofuels feedstocks. Despite rather detailed information on the structure of the cell wall polysaccharides, little is known about their biosynthesis. The key enzymes are glycosyltransferases (GTs) and plants need a large number of GTs to synthesize the complex polysaccharides present in the walls. In *Arabidopsis*, approximately 450 GT genes have been identified; however, in spite of a significant effort, only few GTs have had their activity determined.

We have expressed many of the *Arabidopsis* GTs in tobacco and *E. coli* and developed assays to determine their activity. The activity of a given GT can be determined by assumption-free assays where a range of different substrates are tested, e.g. combinations of nucleotide sugars and monosaccharide. In other cases, mutant analysis or phylogeny strongly suggests a particular role for a GT, which can then

be tested with more specific assays. Examples of the use of these strategies to determine activity of GTs involved in pectin, hemicellulose, and arabinogalactan biosynthesis will be presented.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

35

Microfluidic Technology for Biofuels Applications

Rajiv Bharadwaj* (Rbharadwaj@lbl.gov), Aarthi Chandrasekaran, April Wong, **Blake A. Simmons**, **Paul D. Adams**, and **Anup K. Singh**

Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

We are developing a suite of microfluidic technologies to address the throughput limitations of conventional approaches for biomass analysis. We have developed a microfluidic electrophoretic assay for rapid (< 1 min) and multiplexed analysis of lignocellulosic biomass samples. To address the low throughput of conventional enzyme screening approaches, we have developed a rapid (1-2 hr) integrated microscale platform for cell-free expression and activity analysis of thermophilic cellulases. We have also developed a high-throughput approach for saccharification studies of ionic-liquid pretreated solid biomass. Currently, we are integrating the above technologies to develop optimized cellulase cocktails for cost-effective production of biofuels from lignocellulosic feedstocks.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

36

A Microscale Platform for Integrated Cell-Free Expression and Screening of Cellulase Activity

Aarthi Chandrasekaran,^{1*} Rajiv Bharadwaj,¹ Joshua I. Park,² Masood Z. Hadi,¹ Rajat Sapra,² **Blake A. Simmons**² (BASimmons@lbl.gov), **Paul D. Adams**,¹ and **Anup K. Singh**¹

¹Technology Division, ²Deconstruction Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

A critical step in the efficient conversion of lignocellulosic biomass to fuel is the deconstruction of the biomass to fermentable sugars. Several efforts are therefore focused on identification, expression and characterization of novel enzymes that hydrolyze lignocellulosic biomass. High-throughput enzyme assays that enable rapid screening of these enzymes can greatly accelerate the current enzyme engineering efforts for biofuels development. In this study, we have developed a miniaturized high-throughput, fluorescence-based screening platform for rapid activity profiling of thermophilic cellulases at elevated temperatures (>80°C). This platform integrates cell-free expression and functional characterization of the cellulases in microwell arrays with volumes as low as two microliters. Herein, we demonstrate the use of this approach to express and screen a panel of thermophilic β -glucosidases and cellobiohydrolases.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

37

Ionic Liquid Pretreatment of Biomass: Dynamic studies with Light Scattering, GC-MS and FTIR

Patanjali Varanasi* (Pvaranasi@lbl.gov), Manfred Auer, **Blake Simmons**, and Seema Singh

Deconstruction and Technology Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Among the many choices for alternative energy, biofuels promise to have the most potential for clean and renewable energy. Cellulose and hemi-cellulose present in the biomass can be converted to simple sugars through enzymatic hydrolysis and further to advanced biofuels downstream. But lignin present in the biomass hinders the enzyme accessibility to cellulose and thus enzyme efficiency and total

biofuel yield. At JBEI, research efforts are focused on Ionic liquid (IL) pretreatment to overcome biomass recalcitrance. However, biomass degradation during IL pretreatment and its effect on microbial growth is not understood. Hence, for efficient biofuel production it is important to fundamentally understand the deconstruction of biomass and the compounds produced from biomass de-polymerization during IL pretreatment process for process optimization. To understand the depolymerization of biomass, avicel (model cellulose), model lignin monomers, dimmers and lignin polymer (kraft lignin and low sulfatealkali lignin), and different biomass (switchgrass, pine and eucalyptus) were pretreated with IL at 120°C and 160°C for different time periods (1, 3, 6, 12h). The resultant compounds from the pretreatment process were analyzed using FTIR and GC-MS. Light scattering was used to find if the pretreatment caused lignin dissolution or is effective in any depolymerization of lignin. The extent of lignin depolymerization was found to be temperature dependent. The results from these experiments show that treating biomass at 160°C for 12 hrs may be the best route to degrade biomass if depolymerization is desired.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

38

Understanding Ionic Liquid Pretreatment of Lignocellulosic Biomass by Hyperspectral Raman Imaging

Lan Sun* (LanSun@lbl.gov), Chenlin Li, **Blake A. Simmons**, and Seema Singh

Pretreatment Group, Deconstruction Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Pretreatment of lignocellulosic biomass is essential for breaking down the highly interwoven matrix of cellulose, lignin and hemicellulose, which facilitates enzyme accessibility and adsorption to cellulose for efficient saccharification. Ionic liquids are a new class of non-volatile solvents exhibiting excellent solvating properties and have shown great promise for lignocellulosic biomass pretreatment with easy recovery of cellulose by rapid precipitation with anti-solvents. Ionic liquids have been demonstrated to be very effective in cellulose solubilization in bulk, and have shown to swell cell walls perhaps by breaking inter and intra chain hydrogen bonding. However, to date, molecular level understanding of ionic liquid pretreatment on lignin and its impact on different tissue and cell types of biomasses is lacking. The aim of this research is to develop a fundamental understanding of ionic liquid pretreatment by monitoring the compositional changes during the pretreatment

process. Raman microscopy based on molecular vibrational spectroscopy is a label-free imaging technique capable of real-time and noninvasive examination of plant cell walls with chemical selectivity. In this research, we employed hyperspectral Raman imaging to study tissue and cell type specific distribution of cell wall components and the impact of ionic liquid pretreatment on various cell types of corn stover to identify signatures for predicting deconstruction-ability and understand pretreatment dynamics. The Raman mapping results have shown that the distribution of lignin and cellulose varies significantly across different tissue and cell types in the following order: sclerenchyma cells and tracheids > epidermal cells > bundle sheath cells > parenchyma cells. Lignin content decreases rapidly in tracheids and sclerenchyma cells and slowly in parenchyma cells during ionic liquid pre treatment. Significant cell wall swelling of various cell types during ionic liquid pretreatment was revealed by confocal fluorescence microscopy.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

39

Characterization of a Hyperthermophilic Cellobiohydrolase from *Caldicellulosiruptor saccharolyticus*: Enzymatic Hydrolysis of Cellulose Mediated by Substrate Binding

Joshua I. Park^{1,2*} (JIPark@lbl.gov), Michael S. Kent,^{1,3} Supratim Datta,^{1,2} Bradley M. Holmes,^{1,2} Dean C. Dibble,^{1,2} Zhaohua Huang,⁴ **Blake A. Simmons**,^{1,2} and Rajat Sapra^{1,2}

¹Deconstruction Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California; ²Dept. of Biomass Science and Conversion Technology, Sandia National Laboratories, Livermore, Calif.; ³Biodefense and Materials Science Center, Sandia National Laboratories, Albuquerque, N.M.; ⁴Depts. of Bioengineering, Therapeutic Sciences, and Pharmaceutical Chemistry, University of California, San Francisco

Project Goals: See goals for Abstract 26.

We cloned, expressed, purified, and characterized a recombinant cellobiohydrolase (EC 3.2.1.91) domain from celB, a modular cellulolytic gene from *Caldicellulosiruptor saccharolyticus* that contains glycoside hydrolase family 10 (GH10), carbohydrate binding module family 3 (CBM3), and GH5 domains. The deletion analysis of *celB* confirmed that the constructs containing the GH5 domain were able to hydrolyze the soluble substrates carboxymethyl-cellulose (CMC) and *p*-nitrophenyl- β -D-cellobioside (*p*NPC). Therefore, we focused our study on the recombinant CBM3-GH5 and GH5. The recombinant proteins were expressed in *E. coli*,

and purified to homogeneity by affinity and ion-exchange chromatography methods. The functional stability and melting temperature measurements demonstrated that both CBM3-GH5 and GH5 are highly stable up to 80°C at pH 5.5. CBM3-GH5 and GH5 were also able to hydrolyze microcrystalline cellulose (Avicel), ionic liquid (IL)-pretreated cellulose, and IL-pretreated corn stover to produce cellobiose; CBM3-GH5 produced more cellobiose than GH5 did from these insoluble substrates. We employed fluorescence confocal microscopy and total internal reflective fluorescence (TIRF) methods to investigate whether the binding interaction between the enzyme and substrate was attributed to the product yield from the insoluble substrates. We observed stronger binding interaction between CBM3-GH5 and cellulose (both microcrystalline and amorphous) than that between GH5 and cellulose. Thus, the higher product yields from the enzymatic hydrolysis of microcrystalline cellulose, IL-pretreated cellulose, and IL-pretreated corn stover by CBM3-GH5 were possibly mediated by the interaction between the CBM3 domain and the substrates. The recombinant CBM3-GH5 is a thermostable and active cellobiohydrolase that could be used with other types of cellulolytic enzymes for degradation of IL-pretreated biomass to produce fermentable sugars.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

40 Targeted Enzyme Discovery in Feedstock-Adapted Microbial Communities

John M. Gladden* (JMGladden@lbl.gov), Martin Allgaier, Amitha M. Reddy, Jean S. VanderGheynst, Terry C. Hazen, **Blake A. Simmons**, Phillip Hugenholtz, and **Steven W. Singer**

Microbial Communities Group, Deconstruction Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Efficient saccharification of pre-treated feedstocks is essential to produce inexpensive biofuels derived from biomass. The enzyme cocktails used for this step need to be robust and able to withstand inhibitors produced during pretreatment. To identify enzymes suitable for saccharification on an industrial scale, we have focused on adapting microbial communities to specific feedstocks and pretreatment conditions (i.e. high temperatures, ionic liquid contamination, etc.), followed by functional characterization of secreted glycosyl hydrolases. To select for thermophilic enzymes, microbial communities derived from compost were adapted to biomass feedstocks at 60°C. Phylogenetic profiling of these communities show that each consists of a simple mixed con-

sortia with just two or three species in high abundance. The secretomes obtained from these consortia have biomass-deconstructing enzymatic activity that is both thermostable and active in high concentrations of ionic liquid, two highly desirable characteristics for industrial enzymes. Zymography was used to investigate the complement of glycosyl hydrolase enzymes expressed by the consortia, detecting at least a dozen active enzymes for multiple polysaccharide substrates. Currently, proteins within individual zymogram bands are being studied by MS-based proteomics. Candidate glycosyl hydrolases will be identified by comparing measured peptide masses to predicted protein sequences from Carbohydrate Active enZYme (CAZy) database and genome sequences of reference organisms related to consortial members. Future work will utilize metagenomic and single-cell genomic sequencing to document comprehensively the glycosyl hydrolases secreted by these highly active microbial communities. Cataloguing the glycosyl hydrolases in these secretomes will enable us to design thermophilic enzyme cocktails for biomass deconstruction that function under the conditions required for industrial conversion of biomass to biofuels.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

41 Ionic Liquid Cation Influence on the Dissolution of Isolated Lignins and Biomass for Ethanol Production

Anthe George^{1*} (AGeorge@lbl.gov), Peter Benke,² Kim Tran,¹ **Blake Simmons**,¹ and Bradley Holmes¹

¹Pretreatment Group, Deconstruction Division and ²Technology Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

For economically viable fermentation-based lignocellulosic biofuel production, pretreatment of the biomass is required. Ionic liquids are proving a promising pretreatment medium. However, debate exists as to the extent of biomass dissolution by ionic liquids and the mechanisms by which this solubility occurs. In this study the degree of biomass dissolution as a function of ionic liquid cation was investigated. Four chloride anion based ionic liquids were studied for their ability to dissolve either organosolv, klason, indulin AT, and milled-wood lignins along with eucalyptus and pine biomass. Imidazolium, phosphonium, ammonium and pyridinium anions were screened. Size exclusion chromatography and MALDI-TOF methods were developed to study the degree of dissolution and polydispersity as a function of solvent cation. The comparison between isolated lignins

and biomasses deduced not only the degree of dissolution, but mechanistic information on the component of biomass dissolved by ionic liquids.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

42

Trapping Lignin Degrading Microbes in Tropical Forest Soil

Kristen M. DeAngelis^{1,2*} (KMDeangelis@lbl.gov), Martin Allgaier,^{2,3} Whendee L. Silver,^{1,4} Yaucin Chavarria,¹ Julian Fortney,¹ Phillip Hugenholtz,^{1,3} **Blake Simmons**,^{2,5} Kerry Sublette,⁶ and Terry C. Hazen^{1,2}

¹Ecology Dept., Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, Calif.; ²Microbial Communities Group, Deconstruction Division, Joint BioEnergy Institute, Emeryville Calif.; ³DOE Joint Genome Institute, Walnut Creek, Calif.; ⁴Dept. of Environmental Science, Policy and Management, University of California, Berkeley; ⁵Sandia National Laboratories; and ⁶University of Tennessee

Project Goals: See goals for Abstract 26.

Lignin in feedstock plant material represents a barrier to more efficient plant-to-biomass conversion and can also hinder enzymatic access to cellulose. For this reason, it is critical to develop a way to degrade recovered lignin for next generation feedstock-derived biofuels. While the best-known ligninases are fungal, bacteria are more amenable to emerging cellulosic biofuels technologies. Tropical rain forest soils in Puerto Rico are likely dominated by bacterial decomposers because of the frequent anoxic conditions and fluctuating redox characteristic of these soils, so we focused here to search for novel bacterial lignase producers. To do this, we buried bug traps containing lignin-amended and unamended biosep beads in the soil and incubated them for 1, 4, 13 and 30 weeks. At each time point, phenol oxidase and peroxidase enzyme activity was found to be elevated in the lignin-amended versus the unamended beads, while cellulolytic enzyme activities were significantly depressed in lignin-amended beads. Quantitative PCR of bacterial communities showed more colonization in the lignin-amended compared to the unamended beads after one and four weeks, which attenuated over the course of the incubation. The microbial community was analyzed by microarray (PhyloChip) and by pyrotag sequencing of the community 16S ribosomal RNA genes. Community trends were strongly driven by time but also lignin-amendment to the beads. These techniques also allow us to identify which taxa were increased in lignin-amended compared to unamended beads,

which included representatives from the phyla Actinobacteria, Firmicutes, Acidobacterial and Proteobacteria.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

43

High-Resolution Electron Microscopy Imaging of Plants and Pretreated Biomass

B. Knierim, L. Prak, M. Zemla, D. Jorgens, P. Varanasi, C. Li, Ö. Persil Cetinkol, L. Sun, K. Tran, Y. Verhertbruggen, H.V. Scheller, D.C. Dibble, B.M. Holmes, **B.A. Simmons**, **P.D. Adams**, S. Singh, and M. Auer* (Mauer@lbl.gov)

Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Using sophisticated sample preparation followed by scanning and transmission electron microscopy imaging we have studied a variety of plants and pretreated biomass samples, with a focus on the effects of ionic liquid pretreatment on Switchgrass plant cell walls. We have compared ionic liquid pretreatment to other pretreatment techniques such as acid pretreatment and ammonia fiber expansion, and find ionic liquids to exhibit a much larger effect on the biomass. By imaging plant material being pretreated with ionic liquids from different time points, we found that the cell wall material shows significant changes within less than 30 minutes of exposure to ionic liquids. We utilized widefield TEM to cover a statistically significant number of cell walls in different tissues and compared their precise dimension as a function of exposure to ionic liquids. We found that only the secondary cell wall of sclerenchyma cells undergoes dramatic changes, increasing to twice its original dimension over time, whereas primary cell walls were not found to expand significantly.

Interestingly, the effect of the ionic liquid pretreatment is very different for various plant feedstocks: While switchgrass and corn stover get dissolved easily there is only a minor effect on Eucalyptus arboretum. Other efforts currently underway at JBEI include cell wall characterization of *Arabidopsis* mutants both via optical and EM imaging as well as mechanical stress testing, and in addition imaging of lignocellulose digesting microbial communities.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

44

Biobrick Vectors and Datasheets: A Synthetic Biology Platform for Metabolic Engineering

Taek Soon Lee* (TSLee@lbl.gov), Rachel A. Krupa, Meghdad Hajimorad, Nilu Prasad, Sung Kuk Lee, and **Jay D. Keasling**

Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Synthetic biology has been explored as a powerful tool in metabolic engineering. Standardization of biological parts is one of the most fundamental grounds to provide the foundation of designing synthetic biological systems, and there have been several attempts of establishing standard assembly strategy. For example, researchers at MIT had established BioBrick™ standard biological part strategy using XbaI and SpeI restriction enzymes and started the Registry of Standard Biological Parts. Here, using a similar strategy, we present a new standard using more robust BamHI and BglII restriction enzymes for the construction of novel plasmids with gene expression devices. We have designed and constructed 96 biobrick-compatible plasmids with a various combination of replication origin, antibiotic resistance, and transcriptional promoter. With these plasmids, we have collected protein expression data in various culture conditions using fluorescent protein as a reporter and documented them as a format of datasheet. This biobrick vector datasheet will be a useful source of information for designing and engineering metabolic pathways toward biofuel production.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

45

Flux Analysis of Biodiesel-Producing *E. coli*

H.G. Martin* (HGMartin@lbl.gov), J. Gin, E. Steen, P. Benke, H. Burd, A. Mukhopadhyay, and **J.D. Keasling**

Fuels Synthesis Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Biofuel yields can be limited below their theoretical maximum because of pathway characteristics (e.g. lack of enzyme activity or unbalanced gene expression) or a lack of flux directed towards the synthesizing pathway. The goal of the Host Engineering directorate is to prevent the latter. A way to achieve this is to knock genes out in such a way that carbon flux gets channeled towards the desired pathway. In

order to do that in a rational manner it is desirable to know the internal metabolic fluxes and have a way to predict the outcome of knockout experiments. Here we present a internal metabolic flux profiles for a biodiesel-producing *E. coli* measured through ¹³C metabolic flux analysis (¹³C MFA) and knockout suggestions obtained through Flux Balance Analysis (FBA) constrained by ¹³C MFA flux measurements.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

46

Identification of Genes Essential to Long-Chain Alkene Biosynthesis in *Micrococcus luteus*

Ee-Been Goh, **Jay D. Keasling**, and Harry R. Beller* (HRBeller@lbl.gov)

Fuels Synthesis Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Aliphatic hydrocarbons are appealing targets for advanced cellulosic biofuels, as they are predominant components of petroleum-based gasoline and diesel fuels and thus would be compatible with existing engines and fuel distribution systems. We have studied alkene biosynthesis in *Micrococcus luteus*, a close relative of *Sarcina lutea* (now *Kocuria rhizophila*), which was previously reported to biosynthesize *iso*- and *anteiso*-branched, long-chain alkenes. The underlying biochemistry and genetics of alkene biosynthesis were not elucidated in those studies. We show here that heterologous expression of a three-gene cluster from *M. luteus* (Mlut_13230-13250) in a fatty-acid overproducing *E. coli* strain resulted in production of long-chain alkenes, predominantly 27:3 and 29:3 (no. carbon atoms: no. C=C bonds). Heterologous expression of Mlut_13230 (*oleA*) alone produced no long-chain alkenes but unsaturated aliphatic monoketones, predominantly 27:2, and *in vitro* studies with the purified Mlut_13230 protein and tetradecanoyl-CoA produced the same C₂₇ monoketone. Gas chromatography-time of flight mass spectrometry confirmed the elemental composition of all detected long-chain alkenes and monoketones (putative intermediates of alkene biosynthesis). Negative controls demonstrated that the *M. luteus* genes were responsible for production of these metabolites. Studies with wild-type *M. luteus* showed that the expression of Mlut_13230-13250 and 29:1 alkene biosynthesis both corresponded with bacterial population over time. We propose a metabolic pathway for alkene biosynthesis starting with acyl-CoA (or -ACP) thioesters and involving decarboxylative Claisen condensation as a key step, which we believe is catalyzed by OleA. Such activity is consistent with our data

and with the homology of Mlut_13230 (OleA) to FabH (β -ketoacyl-ACP synthase III), which catalyzes decarboxylative Claisen condensation during fatty acid biosynthesis.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

47

Towards Automated Assembly of Biological Parts

N.J. Hillson* (NJHillson@lbl.gov), J.W. Thorne, M.Z. Hadi, and **J.D. Keasling**

Fuels Synthesis and Technology Divisions, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

The production of clean renewable biofuels from cellulosic starting material requires concerted feedstock engineering, deconstruction of plant matter into simple sugars, and microbial fermentation of the sugars into biofuel. These three efforts share significant molecular biological challenges, including the construction of large enzymatic libraries (e.g. vast collections of glycosyl transferases, cellulases, and efflux pumps), the generation of combinatorial libraries (e.g. multi-functional enzyme domain fusions; variations in copy number, promoter and ribosomal binding site strength), and the concurrent assembly of multiple biological parts (e.g. the incorporation of an entire metabolic pathway into a single target vector). With these challenges in mind, we are developing hybrid multi-part assembly methodologies and translating them to robotics-driven protocols. Given a target library to construct, our vision is that the high-throughput methodology will provide automated oligo and optimal assembly process design, and robotic control of the PCR and multi-part assembly reactions. The beneficial output of this work will include reagents and resources for, and collaborations with, members of the JBEI and larger life sciences communities, reducing the time, effort and cost of large scale cloning and assembly tasks, as well as enabling research scales otherwise not feasible without the assistance of computer-aided design tools and robotics.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

48

Improving Biofuel Production by Using Efflux Pumps to Limit Solvent Toxicity

Mary J. Dunlop, Masood Hadi, Mario Ouellet, **Paul D. Adams**, **Jay D. Keasling**, and Aindrila Mukhopadhyay* (Amukhopadhyay@lbl.gov)

Fuels and Technology Divisions, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Biofuels can be produced by microbes that break down plant matter or sugars to make fuel. However, biofuel-producing microbes are limited by the intrinsic toxicity of the solvent like biofuels they are trying to produce. The more fuel the cell produces, the more toxic the surrounding environment becomes. RND efflux pumps are a class of membrane transporters that confer resistance to a wide variety of toxins, including solvents. We focus on investigating the role of native, as well as heterologously expressed, efflux pumps in *E. coli*. Targeted studies focus on the well-characterized *E. coli* AcrAB-TolC system, and efflux pumps from solvent resistant bacteria such as *Pseudomonas putida*. Because efflux pumps are likely to be specific to certain fuel molecules and stressors, a wider range of native and heterologous efflux pump systems must be tested against different fuel compound exposure, growth conditions, and in different engineered hosts. To address our broad goal of improving solvent resistance using efflux pumps, a high-throughput approach has been initiated to create a library of expression vectors representing all efflux pumps from *E. coli* as well from other organisms known to be naturally resistant to solvents.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

49

Quantitative Proteomics for Metabolically Engineered Biofuel Pathway Optimization

Christopher Petzold* (CJPetzold@lbl.gov), Alyssa Redding, Tanveer Bathth, Rossana Chan, Rachel Krupa, Taek Soon Lee, Aindrila Mukhopadhyay, **Paul Adams**, and **Jay Keasling**

Fuels Synthesis and Technology Divisions, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Analytical methods are fundamental to establishing the success of biofuel metabolic engineering efforts. Monitoring

and quantifying pathway intermediates facilitates identification of bottlenecks and where alterations offer the greatest potential impact on titer. However, many different factors may contribute to a bottleneck. High levels of protein production or high activity of an upstream enzyme will produce substrates at rates higher than can be consumed by later steps in the pathway. Analogously, downstream enzymes may have low activity or be poorly produced, and factors such as low solubility or rapid protease degradation can limit the amount of enzyme available. Consequently, the detection and quantification of the enzymes of interest is integral to optimizing engineered pathways. Quantitative proteomics analysis, consisting of liquid chromatographic separation coupled to mass spectrometry, is a rapid method to correlate protein expression levels with metabolite titers. At JBEI, both targeted and untargeted proteomics experiments are used to quantify proteins crucial for biofuel production. With these methods we are characterizing a variety of protein expression conditions for several biofuels pathways, including various promoters and plasmid systems, to identify bottlenecks and determine optimal protein levels for high biofuel titers.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

50

Strategies to Improve Resistance and Production Phenotypes of *E. coli*

Adrienne E. McKee, John Haliburton, Helcio Burd, **Jay Keasling**, and Swapnil Chhabra* (SRChhabra@lbl.gov)

Fuels Synthesis Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

Many of the microbial phenotypes of interest to metabolic engineers are complex in that multiple genes, pathways, and regulatory networks are involved in generating the targeted behavior. *A priori* prediction of additional changes that will further improve phenotypes can therefore prove difficult due to our incomplete understanding of the functions and connectivity of gene products far removed from the pathway of interest. One complement to rational approaches is to exploit the strength of mutation and selection or screening to obtain strains capable of improved resistance to pretreatment growth improved production titers in the case of production phenotypes. Towards these goals, we are employing and refining methods that rely on natural or augmented mutation rates or on directed protein evolution to improve *E. coli* phenotypes. Our three major avenues of investigation include selection for inhibitor resistance by continuous culture in chemostats, development of inducible

and temperature sensitive mutator plasmids, and generation and screening plasmid libraries of mutated gene regulators for enhanced phenotypic behavior. We are interested in both the genetic and regulatory alterations that underlie phenotypic improvements.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

51

Microbial Production of Fatty Acid–Derived Fuels and Chemicals in *Escherichia coli*

Eric J. Steen^{1,2} * (EJSteen@lbl.gov), Yisheng Kang,¹ Mario Ouellet,¹ Helcio Burd,¹ John Haliburton,¹ Michael Lee,^{1,2} Alyssa Redding,¹ Heather Szmids,¹ Jennifer Gin,¹ Nathan Hillson,¹ Harry R. Beller,¹ and **Jay D. Keasling**^{1,2,3,4,5}

¹Joint BioEnergy Institute, Emeryville, Calif.; ²Dept. of Bioengineering, ³QB3 Institute, and ⁵Dept. of Chemical Engineering, University of California, Berkeley; and ⁴Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Project Goals: See goals for Abstract 26.

Increasing energy costs and environmental concerns have emphasized the need to sustainably produce renewable fuels and chemicals. Major efforts to this end are focused on the microbial production of high-energy fuels through cost-effective “consolidated bioprocesses”. Fatty acids are composed of long alkyl chains and represent nature’s “petroleum,” being a primary metabolite class used by cells for both chemical and energy storage functions. These energy rich molecules are today isolated from plant and animal oils for a diverse set of products ranging from fuels to oleochemicals. A more scalable, controllable, and economic route to this important class of chemicals would be through the microbial conversion of renewable feedstocks, such as biomass-derived carbohydrates. Here we demonstrate the engineering of *E. coli* to produce structurally tailored fatty acid ethyl esters (biodiesel), fatty alcohols, and waxes directly from simple sugars and the further engineering of the biodiesel-producing cells to secrete hemicellulases, a step toward producing these compounds directly from hemicellulose, a major component of plant-derived biomass¹. Although this complete production scheme has been demonstrated, increases in titer, productivity, and yield are necessary for industrial transition. Strategies employed for increasing yields of biodiesel include balancing the enzymes in the pathway, condensing the pathway onto a triple-operon, single-plasmid system, and subsequent chromosomal integration. These efforts guided towards understanding fermentation scalability, pathway stability, and balancing pathway enzymes for biodiesel production have resulted in higher yields.

Reference

1. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, del Cardayre SB, Keasling JD. "Microbial Production of Fatty Acid-Derived Fuels and Chemicals from Plant Biomass," *Nature* (2009) accepted for publication.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

52

JBEI Electronic Laboratory Notebook System

D. Pletcher* (DCPletcher@lbl.gov), J. Gladden, B. Holmes, J. Keasling, S. Lane, W. Morrell, A. Herrera Panganiban, J. Petersen, F.W. Townsend-Merino, J. Wildman, and **P.D. Adams**

Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

JBEI is pursuing a novel approach to implement an electronic laboratory notebook (ELN) system. Ease of use, performance and cost are primary factors in the design of the system. Commercially successful, inexpensive software products are combined into a full-featured, hybrid system using integration software developed in-house. Powerful desktop note-taking software on a commodity tablet computer enables scientists to capture and organize notes, sketches, images, documents and other digital data files. All notebook information is periodically copied to a central, web-accessible repository for search and review by other authorized users. Integration with a secure digital time-stamping service ensures that intellectual property can be defended in court. Pilot users have exhibited a high level of satisfaction; new users are adopting the system willingly based on peer recommendations. The completed system will facilitate online discovery and scientific collaboration, and will serve as a core component of our GTL Knowledgebase.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

53

JBEI Computational Biology Core

Dylan Chivian^{1,2,3*} (DCChivian@lbl.gov), Greg D. Friedland,^{1,2,4} John T. Bates,^{1,2,3} Paramvir S. Dehal,^{2,3} Marcin P. Joachimiak,^{2,3} Keith Keller,^{2,3} Morgan N. Price,^{2,3} Jason K. Baumohl,^{2,3} and **Adam P. Arkin**^{1,2,3,5*}

¹DOE Joint BioEnergy Institute, Berkeley, Calif.; ²Virtual Institute for Microbial Stress and Survival, Berkeley, Calif.; ³Lawrence Berkeley National Laboratory, Berkeley, Calif.; ⁴Sandia National Laboratories, Livermore, Calif.; and ⁵University of California, Berkeley

Project Goals: See goals for Abstract 26.

The Computational Biology Core Group in the Technology Division of the Joint BioEnergy Institute (JBEI) is responsible for data integration and comparative, evolutionary, and functional genomic analysis for the purpose of engineering microbes for biofuel production. Leveraging the VIMSS MicrobesOnline web resource (<http://www.microbesonline.org>) for comparative and evolutionary genomics and analysis of microarray, proteomic, and metabolomic data sets, we are extending the supported microorganisms to include bioenergy-relevant fungi and algae as well as integrating capabilities to allow for pursuit of questions specific to bioenergy challenges.

Computational analysis and biological engineering requires an understanding of the biology at several scales. The efforts of the Computational Biology Core are therefore geared towards creating tools to facilitate our research at the level of components, systems, cells, and communities. For example, biological degradation of plant cells walls is accomplished by enzymes containing multiple domains that in combination confer specificity and activity. We are studying the combinations nature has employed to allow for making our own combinations. At the atomic scale, we are computationally designing the structures of proteins for increased stability under industrial conditions. At the system level, we are working to discover metabolic pathways for biofuel production as well as the genetic factors involved in tolerance to harsh biomass pretreatment conditions and resistance to biofuel toxicity. To aid efforts in synthetic pathway engineering, we are building a framework for analyzing functional data in a metabolic network context. At the cellular level, we are using evolutionary studies with phenotype data and genetic analysis to engineer cell lines that are better suited to industrial conditions or have superior yields of the desired biofuel. Finally, at the community level, discovery of genes from environmental samples will expand the repertoire of enzymes we can engineer for biomass degradation under varying conditions. Our approach is to analyze the environmental genomic data in a phylogenetic context to allow for a higher-resolution annotation of the role of each enzyme, as well as to obtain an evolutionary picture of the key functional genes and organisms in each ecosystem.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy,

Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

54

Techno-Economic Modeling of Cellulosic Biorefineries

Daniel Klein-Marcuschamer* (DKlein@lbl.gov), Piotr Oleskowicz-Popiel, **Blake A. Simmons**, and **Harvey W. Blanch**

Deconstruction Division, Joint BioEnergy Institute, Emeryville, Calif.

Project Goals: See goals for Abstract 26.

The realization of biofuels at a commercial scale will depend on processes that are energetically, environmentally, and economically sustainable. One main thrust of biofuels research, therefore, has been in techno-economic analysis of biofuel alternatives, which facilitates process design, optimization, and performance evaluation. These studies usually rely on experimentally-derived or assumed parameters to estimate process performance values such as capital and operating costs, GHG emissions, biofuel yield on feedstock, among others. Naturally, the results of the simulations strongly depend on the parameter choices or scenario maps that are considered, which limits the information that the community can extract from the results if only a few instances are outlined. Here, we present a techno-economic model of lignocellulosic ethanol production that is open and transparent and that uses assumptions that reflect technologies that are currently available. Using this model, we have studied how feedstock, enzyme, and strain engineering efforts could potentially affect the economic and performance attributes of the process. The results indicate that pretreatment and hydrolysis yield, inhibition during fermentation, and energy consumption are major factors impacting the economic viability of the process.

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

BioEnergy Science Center (BESC)

55

The BESC Knowledgebase and Public Web Portal for Bioenergy-Related Organisms

Morey Parang,^{1,2} Denise Schmoyer,^{1,2} Tatiana Karpinets^{1,2*} (karpinetsv@ornl.gov), Mustafa Syed,^{1,2} Mike Leuze,^{1,2} Byung Park,^{1,2} Doug Hyatt,^{1,2} **Ed Uberbacher**,^{1,2} and Martin Keller^{1,2}

¹BioEnergy Science Center, Oak Ridge, Tenn. and ²Oak Ridge National Laboratory, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 55-56): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC researchers in characterization, modeling, education and data management seek to provide integrating data and understanding across the center. We use advanced technologies to analyze chemical and structural changes within biomass. We have developed a coordinated characterization pipeline to provide composition, pretreatment, and enzyme digestibility data from native and modified plant samples and to store this data within a data management system. We are developing knowledge environments to allow access to improved data analysis and modeling tools.

Combining experimental data with biological models will help define the genomic and physical basis of plant cell-wall recalcitrance and deconstruction. Laboratory Information Management Systems and knowledgebase environments have separate key roles.

The objective of the BESC Knowledge base project is to develop and implement an infrastructure to represent systems level knowledge, data and information about key plants, microbes, and molecules in BESC and provide views of these data to users. The Knowledgebase (KB) is designed to aid investigators in the comprehensive evaluation of organisms for bioenergy associated strategies. Specifically, the Knowledgebase provides annotations of sequenced bioenergy related organisms, their phenotypic characteristics, pathway genome databases (PGDBs) for each organism, orthology information, a set of tools for annotation and analysis of carbohydrate active enzymes (CAZymes tool kit) and a diverse set of product annotations from different public resources, like InterPro, CDD, COG, PFAM, SMART, and others. The user environment has (i) genomic tools and gene cards, (ii) BioCyc pathway tools including capabilities

to search and browse PGDBs, (iii) a comparative analysis environment for overlaying the metabolic maps with BESC or user provided experimental 'omics' data, (iv) a knowledge mining environment that allows phenotype comparison in terms of CAZy Families/ Pathways/ Enzymes and sequence annotation by CAZy families using PFAM domains, and (v) integration of investigator-derived or external data (External Experimental Data (GEO/ArrayExpress) with product annotation from diverse set of databases including CAZy. The collected information has been employed in the analysis of experimental data produced by BESC and for comparative analysis of phenotypic characteristics of the BESC targeted organisms in terms of their genomic, metabolic and cellular characteristics.

We have constructed a public portal to the BESC knowledgebase to provide the larger community with information about bioenergy organisms. The portal provides integrated views of information available from a variety of different public resources or produced by BESC (public domain) and presents this knowledge in a systematic and unified way. The portal environment allows users to search for different types of annotations for each organism or across organisms; download the annotations and results of queries in user friendly formats; compare phenotypic characteristics of organisms in terms of their genomic, metabolic and cellular characteristics; visualize and link the annotations to the experimental data for their further analysis.

The KB is linked to the BESC Laboratory Information Management System (LIMS), which is used for tracking samples and capturing experiment metadata. The combined KB and LIMS projects have developed ORNL standard data models for experimental workflow from the project level down to individual experiments, aliquots, and instrument metadata. We intend to leverage these ongoing efforts for the combined GTL data sharing effort and global GTL knowledge base developments.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

56 Carbohydrate-Active Enzyme Annotation Tools (CAT) in the BESC Knowledgebase Portal

Mustafa Syed^{1,2*} (syedmh@ornl.gov), Byung-Hoon Park,^{1,2} Tatiana Karpinetz,^{1,2} Michael Leuze,^{1,2} Ed Uberbacher,^{1,2} and Martin Keller^{1,2}

¹BioEnergy Science Center, Oak Ridge, Tenn. and ²Oak Ridge National Laboratory, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 55.

The Carbohydrate-Active Enzyme (CAZy) database is a manually curated repository of experimentally characterized enzymes that build and breakdown complex carbohydrates. The enzymes, also referred to as CAZymes, implement many important functions in the cell and of great interest of the biofuel research because of their involvement in biosynthesis and the degradation of the complex polysaccharides of the plant cell wall. Despite rich and invaluable information stored in the database, software tools utilizing this information for analysis and annotation of newly sequenced genomes by CAZy families are limited, partly because of limited capabilities of the CAZy database for searching and downloading enzyme sequences.

We have addressed this problem by developing a set of tools to search the database and to annotate a new sequence or a set of sequences with CAZy families. We used Perl scripts to scrape HTML web pages for each family in the CAZy database and have organized the downloaded information in a local MySQL database. The collected information was supplemented by protein sequences downloaded from Genbank and by sequence associated information, such as Genbank accession numbers, enzyme names, EC numbers, predicted protein families from the Pfams database and the organism taxonomy. We then developed a set of tools and a user friendly web interface to search the collected information by the enzymes name or keyword, CAZy family, organism taxonomy, name or keyword. Because many CAZymes, especially from Glycosyl Hydrolase (GH) families, have a complex modular architecture and are annotated by two and more families, an interface was developed to search such associations across the entire CAZy database.

Two complementary approaches were provided in the toolkit to annotate sequences with CAZy families. The first approach is based on a similarity search of a protein sequence or a set of sequences in a given genome against the entire non-redundant sequences of the CAZy database using uni-directional or bi-directional Blast. The second approach is based on associations between protein domains and CAZy families. In addition to links between Pfam domains and CAZy families available in the CAZy database, we have derived a set of new links by applying the association rule learning algorithm to the collected data. The augmented set of Pfam-to-CAZy family associations has significantly improved sensitivity of the annotation. The evaluation of the approaches using the manually curated genomes of *Clostridium thermocellum* ATCC 27405 and *Saccharophagus degradans* 2-40 indicated that in combination they can provide a high degree of specificity and sensitivity for predicting CAZymes in the newly sequenced organisms.

The developed tools were employed to predict CAZymes in a recent submission of the *Populus trichocarpa* genome (Jul 30 2009) and in the *Escherichia coli* genome. By using the discovered confident associations between CAZy families and pfam domains we assigned functional activity, taxonomic groups and CAZy families to several unknown domains including DUF2029, DUF297, DUF303, and DUF847. We have also revealed some conserved associa-

tions between CAZy families that are characteristics of specific taxonomic groups, like plants, fungi or bacteria.

The CAZy toolkit and environment is accessible as part of the BESC public portal at <http://cricket.ornl.gov/cgi-bin/cat.cgi>.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

57 Technology and Transgenics for Unparalleled Improvements in Switchgrass Biomass Quality

D.G.J. Mann^{1,5*} (dmann1@utk.edu), L.L. Abercrombie,^{1,5} H. Baxter,^{1,5} K. Berk,^{1,5} J.N. Burris,^{1,5} F. Chen,^{2,5} X. Chen,^{1,5} C. Fu,^{2,5} M. Halter,^{1,5} J. Han,^{1,5} F. Hardin,^{2,5} J.S. Hawkins,^{3,5} T. Hernandez,^{2,5} H. Hisano,^{2,5} Z. King,^{3,5} P.R. LaFayette,^{3,5} Y. Lee,^{2,5} M. Mazarei,^{1,5} J.R. Mielenz,^{4,5} R.J. Percifield,^{3,5} C. Poovaiah,^{1,5} H. Ramanna,^{2,5} M. Rudis,^{1,5} H. Shen,^{2,5} I. Torres-Jerez,^{2,5} X. Xiao,^{2,5} J. Zhang,^{2,5} J.L. Bennetzen,^{3,5} R.A. Dixon,^{2,5} R. Nelson,^{2,5} W.A. Parrott,^{3,5} M.K. Udvardi,^{2,5} Z.-Y. Wang,^{2,5} J. Zale,^{1,5} **C.N. Stewart Jr.**,^{1,5} and M. Keller^{4,5}

¹University of Tennessee, Knoxville; ²The Samuel Roberts Noble Foundation, Ardmore, Okla.; ³University of Georgia, Athens; ⁴Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁵BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 57-59): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and *Populus*) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generation of a large number of natural and modified plant samples as well as detailed studies into poorly understood cell wall biosynthesis pathways.

The Switchgrass activity is pursuing improved transformation methods, association studies, and targeted pathways such as lignin synthesis.

Switchgrass (*Panicum virgatum*) has become a leading candidate feedstock for biofuels in the U.S. and is a crucial model feedstock component of BESC. Biotechnology of switchgrass is important in screening potential cell wall

biosynthesis genes, and is being performed by six laboratories in three institutions within BESC. Accomplishments include altering lignin biosynthesis, improved tissue culture and transformation systems, optimization of a virus-induced gene silencing (VIGS) system and a new vector set for monocot transformation. A transgene pipeline committee was established to identify genes of interest to be evaluated in stably transformed switchgrass (through overexpression or knockdown technologies). To evaluate genes with unknown or poorly understood functions prior to stable transformation we are utilizing VIGS, which more rapidly (within 2 months) provides results through transient knockdown of target gene expression. Our VIGS system uses a *Brome mosaic virus*-based vector to silence genes in foxtail millet (*Setaria italica*), a closely related species to switchgrass which serves as a simpler model system. If a recalcitrance-altering phenotype is observed through VIGS, stable switchgrass transformants will be produced for that gene. One example of success of our research path is the downregulation of a gene from the lignin biosynthesis pathway. The transgenic switchgrass for this gene yields 25% more ethanol than its non-transgenic isolate. In order to coordinate gene expression and to facilitate more rapid screening of genes, we have developed a Gateway-compatible monocot transformation vector set (pANIC) for overexpression and RNAi with visual and selectable markers. BESC has facilitated the coordination of scientific expertise and research in switchgrass biotechnology that would have been otherwise impossible in any individual laboratory.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

58 Enhanced Quality, Value, Yield, Carbon Capture, and Sustainability of Switchgrass Biomass by the Improvement of Root, Microbe, and Soil Interactions

Sita Ghimire^{1,4*} (srghimire@noble.org), **K.D. Craven**,^{1,4} Srinivasa Chaluvadi,^{2,4} Jeff Bennetzen,^{2,4} Eric Worley,^{1,4} Jiading Yang,^{1,4} Michael Udvardi,^{1,4} and Martin Keller^{3,4}

¹The Samuel Roberts Noble Foundation, Ardmore, Okla.; ²University of Georgia, Athens; ³Oak Ridge National Laboratory, Oak Ridge, Tenn.; ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 57.

Nutrient management in biomass production systems serves to maximize yield and minimize production costs and environmental impact. Sustainable production of switchgrass for biofuel will depend, in part, on maximizing nutrient acquisition and assimilation throughout the growing season as well as minimizing nutrient loss at harvest. Nutrient acquisition

and uptake by plants can be enhanced by beneficial soil microbes as well as those existing endophytically within the roots of the switchgrass host. We have undertaken a comprehensive characterization of the microbes associated with the rhizosphere of planted switchgrass cultivars as well as those found within the healthy, surface-sterilized root systems of natural plants found in their native habitat. High levels of microbial biodiversity were detected for both fungi and bacteria, and several strains have been isolated for evaluation of fitness effects on elite switchgrass cultivars. Dramatic differences in rhizosphere and endophyte microbial populations have been found to be a function of host genetics by analysis of different switchgrass cultivars, and mapping studies are now initiated to identify the host genes that determine microbial composition in and around the root. Nutrient use efficiency in plants is a consequence of both the frugality of utilization in the field as well as the effective recycling of those nutrients from aerial tissues to the root system at or before senescence. To facilitate the breeding of varieties that are conservative in their use of soil nutrients to produce biomass, we assessed the natural variation in nutrient-use and remobilization efficiencies of 31 accessions of *Panicum virgatum* by measuring the concentration of 20 elements (N, P, K, Li, B, Na, Mg, Ca, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Mo and Cd) in shoots of field-grown plants harvested at two different stages of development. We detected significant differences between accessions for elemental composition at maturity and after senescence. The accessions/cultivars with the greatest nutrient-use efficiency (smallest loss of nutrient per unit biomass) were BN-14668-65, Kanlow, and Caddo from the point of view of N content, and Kanlow, Cave-in-Rock, and Blackwell from the point of view of P content in senescent shoots. These data will allow a holistic nutrient management strategy to be employed for maximizing yield and sustainability of this important bioenergy crop.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

59

Creating the Genetic and Genomic Foundations to Improve Bioenergy Production from Switchgrass

Katrien M. Devos^{1,4*} (kdevos@uga.edu), Ananta Acharya,^{1,4} Bochra Bahri,^{1,4} Wesley Dean,^{1,4} Matt Estep,^{1,4} Ansuya Jogi,^{1,4} Jennifer Hawkins,^{1,4} Yi-Ching Lee,^{2,4} Ryan Percifield,^{1,4} Ivone Torres-Jerez,^{2,4} Xueweng Wang,^{1,4} Limin Wu,^{1,4} Jiyi Zhang,^{2,4} Qihui Zhu,^{1,4} Joe Bouton,^{2,4} E. Charles Brummer,^{1,4} Malay C. Saha,^{2,4} Shavannor M. Smith,^{1,4} Yuhong Tang,^{2,4} Michael Udvardi,^{2,4} Jeffrey L. Bennetzen,^{1,4} and Martin Keller^{3,4}

¹University of Georgia, Athens; ²The Samuel Roberts Noble Foundation, Ardmore, Okla.; ³Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 57.

Switchgrass (*Panicum virgatum*), a grass native to U.S. prairies, has agronomic characteristics that make it an exceptionally promising feedstock for bioenergy production. There is great potential for improving the biomass composition and biomass yield of switchgrass using genomics-based breeding. We are using a multifaceted approach to create the genetic, genomic and breeding resources that are required for building a successful switchgrass biofuel production program. A comprehensive genetic map consisting of SSR and DArT markers is under construction in a lowland Alamo AP13 x upland Summer VS16 cross. The map will be used for trait mapping, as well as anchoring of the genomic sequence that is currently being produced from genotype AP13 by our collaborators at JGI. An ~6X coverage fosmid library has been generated from nuclear DNA of switchgrass cultivar Alamo, and >30 fosmids containing genes involved in switchgrass cell wall synthesis/composition have been selected, sequenced and annotated. Seven million ESTs from AP13 and VS16 have been generated and placed in a searchable database. The ESTs are derived from mRNA isolated from roots and shoots at three different stages of development and from mRNA enriched for secondary cell wall biosynthesis using laser capture microscopy. The sequence data will be used for SNP development and linkage disequilibrium studies, and will assist in the annotation of the switchgrass genomic sequence and will help in understanding and engineering improved cell wall production in switchgrass. A large biodiversity study consisting of 384 mostly lowland genotypes is under way to assess natural variation for components that affect recalcitrance to ethanol production, biomass yield and disease resistance. Genotyping and phenotyping will be combined to identify trait-marker associations. This information is fed into the breeding program to enhance the development of switchgrass cultivars with desirable cell wall composition and increased yield for commercialization. In summary, our

project provides the essential resources to conduct structural and functional analyses in switchgrass.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

60

Application of an Integrated High-Throughput Pretreatment and Enzymatic Hydrolysis (HTPPH) Screening Tool to Identify Key Biomass Features and Processing Conditions

Jaclyn DeMartini^{1,3*} (jaclyn.demartini@email.ucr.edu), Michael Studer,^{1,3} Charles E. Wyman,^{1,3} and Martin Keller^{2,3}

¹University of California, Riverside; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 60-64): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC researchers in characterization, modeling, education and data management seek to provide integrating data and understanding across the center. We use advanced technologies to analyze chemical and structural changes within biomass. We have developed a coordinated characterization pipeline to provide composition, pretreatment, and enzyme digestibility data from native and modified plant samples and to store this data within a data management system. We are developing knowledge environments to allow access to improved data analysis and modeling tools.

Knowledge gained by thoroughly characterizing biomass chemistry and structure will drive coordinated development of improved plant biomass and degradation microbes. Native, genetically modified, and partially deconstructed lignocellulosic samples will be analyzed.

A high throughput pretreatment and enzymatic hydrolysis (HTPPH) method has been developed as part of the BioEnergy Science Center (BESC) to screen natural and genetically modified biomass types for those with lower recalcitrance to sugar release, define pretreatment conditions, and screen enzyme formulations. The high throughput system has been shown to mimic conventional pretreatment and enzymatic hydrolysis laboratory methods but has the advantage of being able to quickly screen hundreds of

samples. Thus far, the HTPPH system has been successfully applied to a set of 47 natural *Populus trichocarpa* samples in BESC's Poplar Association Study to define trends in sugar release behavior and identify outliers that warrant further study. In addition, the HTPPH system has enabled screening of individual annual rings from a cross section of *Populus tremuloides* to investigate radial variation of sugar release and the importance of sampling technique. Results will be shown from these studies to demonstrate the power of the HTPPH system for screening biomass samples for sugar release and identifying those with reduced recalcitrance. In support of the system, a scaled-down method was also developed to determine biomass composition that speeds compositional analysis, increases accuracy, and reduces labor demands.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

61

Chemical Extraction of Down-Regulated C3H and HCT Alfalfa Reveals Structural Differences in Lignin

Angela Ziebell,^{1,6} Kristen Reichel,^{1,6} Rui Katahira,¹ Fang Chen,^{3,6} Yunqiao Pu,^{3,6} David Johnson,⁴ Robert Sykes,^{1,6} Art Ragauskas,^{3,6} Rick Dixon,^{2,6} Mark Davis^{1,6*} (mark.davis@nrel.gov), and Martin Keller^{5,6}

¹National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colo.; ²Samuel Robert Noble Foundation, Ardmore, Okla.; ³Institute of Paper Science and Technology, Georgia Institute of Technology, Atlanta; ⁴Chemical and Biosciences Center, National Renewable Energy Laboratory, Golden, Colo.; ⁵Oak Ridge National Laboratory, Oak Ridge, Tenn.; ⁶BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 60.

The composition of biomass with regards to lignin content can impact the ease and cost associated with biomass processing and lignin reduction through breeding and genetic modification therefore has potential for reducing costs in biomass processing industries. The compositional changes of two low lignin Alfalfa (*M. sativa*) lines are investigated; antisense down regulated *p*-coumarate 3-hydroxylase (C3H) and hydroxycinnamoyl transferase (HCT). Sequential base extraction readily reduced the lignin content of the transgenic lines leaving a residual H lignin component equal in all lines. We will show that the differences in the lignin extraction of the different lines can be related to the differences in the reactivity of the H monomer versus the normally dominant G and S monomers during lignin formation.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

62

Structural Characterization of the Xylan Oligosaccharides by Mass Spectrometry

Koushik Mazumder^{1,3*} (koushik@ccrc.uga.edu), Ameya Kulkarni,^{1,3} **William York**,^{1,3} and Martin Keller^{2,3}

¹University of Georgia, Athens; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 60.

The plant cell wall is a highly organized composite of many different polysaccharides, protein and aromatic substances. In plants the three main polysaccharides of the cell wall are cellulose, pectin and hemicellulose. Xylans play a very important role in establishing and maintaining the structural integrity of the plant's secondary cell walls, which are the most abundant components of terrestrial biomass. It is likely that differences in the molecular features of xylans (degree of branching and spatial arrangement along the xylan backbone) can alter the properties of the cell wall, including its recalcitrance to enzyme-catalyzed saccharification.

Structural characterization of carbohydrates is achieved usually using a combination of different techniques, which include NMR, GC-MS, mass spectrometry (ESI and MALDI-MS), specific chemical reactions or specific glycosidases. A major advantage of mass spectral analysis (ESI and MALDI-MS) over the other analytical techniques is its extremely high sensitivity (i.e., at the nanogram scale). We have developed highly sensitive methods for the analysis of xylan oligosaccharides using a combination of electrospray ionization (ESI) and multiple-stage mass spectrometry (MSⁿ) with a linear ion trap spectrometer. This approach provides specific information regarding the glycosyl sequences and branching patterns of these oligosaccharides. We will describe the application of these methods for the detailed structural characterization of the neutral and acidic xylan oligosaccharides obtained from the cell walls of various plant species.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

63

Understanding Cellulase Activity Using Single Molecule Spectroscopy

Yu-San Liu,^{1,2} Yining Zeng,^{1,2} Yonghua Luo,¹ Michael E. Himmel,^{1,2} **Shi-You Ding**^{1,2*} (Shi.you.Ding@nrel.gov), and Martin Keller^{2,3}

¹Biosciences Center, National Renewable Energy Laboratory, Golden, Colo.; ²BioEnergy Science Center, Oak Ridge, Tenn.; and ³Oak Ridge National Laboratory, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 60.

To develop more cost-effective approaches to liberate fermentable sugars from recalcitrant biomass, the enzyme cocktail used for saccharification must be improved. We have developed a single-molecule technique based on fluorescence imaging and atomic force microscopy to characterize the cellulose morphology changes and to track the binding orientation and the motion of cellulase components with spatial resolution at the nanometer scale. We used single molecule spectroscopy to study the surface morphology of crystalline cellulose, as well as the real-time behavior of enzymes while bound to cellulose crystals. Preliminary results have revealed a confined nanometer-scale movement of the cellulase components bound to cellulose with preferred binding orientation. Cellulose crystals have also been imaged in real-time showing surface roughness changes, sharpening, and peeling effects by enzyme hydrolysis. The single molecule approach used here offers new opportunities to guide us toward a fundamental understanding of cellulase function, especially the mechanism of the "processivity" of exoglucanase.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

64

Ultrastructure of Lignocellulose "Native-Pretreated-Deconstructed" by Advanced Solid-State NMR

Marcus Foston,^{1,3} Yunqiao Pu,^{1,3} **Arthur J. Ragauskas**^{1,3*} (arthur.ragauskas@ipst.gatech.edu), and Martin Keller^{2,3}

¹Institute of Paper Science and Technology, Georgia Institute of Technology, Atlanta; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 60.

Solid state nuclear magnetic resonance (NMR) methods can provide not only chemical information but atomistic structural details that are not easily accessible by other non-destructive high-resolution structural techniques. This makes solid state NMR methodology particularly useful when studying structural problems in biological systems such as the changes occurring in the ultrastructure and supramolecular structure of biomass. Native, pretreated and enzymatically deconstructed biomass was subjected to advanced ^{13}C , ^1H and ^2H 1D and 2D solid state NMR techniques, in which those results were then complemented with carbohydrate and Klason lignin analysis, enzymatic cellulose digestibility and gel permeation chromatography. Unique information about changes in key substrate characteristics, such as crystallinity index, microfibril/microfibril aggregate dimensions, pore size distribution, pore tortuosity, and possible macromolecular connectivities were studied and used to gain insight into the nature of recalcitrance, mechanisms of pretreatment and optimization of biofuel production.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

65

Elucidating the Mechanism of Xylan Biosynthesis: A Biochemical Approach

Maria J. Peña^{1,4*} (mpena@ccrc.uga.edu), Breeanna Urbanowicz,^{1,4} Jason Backe,^{1,4} Alan G. Darvill,^{1,2,4} Malcolm O'Neill,^{1,4} **William S. York**,^{1,2,4} and Martin Keller^{3,4}

¹Complex Carbohydrate Research Center and ²Dept. of Biochemistry and Molecular Biology, University of Georgia, Athens; ³Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 65–67): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and *Populus*) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generation of large number of natural and modified plant samples as well as detailed studies into poorly understood cell wall biosynthesis pathways.

Biosynthesis research provides detailed studies in these and other plant models on key biosynthetic pathways for

lignin, pectin, xylan and cellulose. The coordinated transformation of target genes for validation is in progress.

Xylan is the most abundant hemicellulose in biomass of poplar and switchgrass. This key polymer in cell wall architecture has been proposed to bind cellulose microfibrils and lignin. Modification of xylan structure has the potential to alter cross-linking between wall components, increasing their accessibility by hydrolytic enzymes and decreasing recalcitrance. Despite the biological and economic importance of xylan, its biosynthesis and subsequent incorporation into the wall are still poorly understood.

Our previous studies of *Arabidopsis* mutants showed that several genes that encoded putative glycosyl transferases are required for normal xylan biosynthesis (Peña et al., *Plant Cell* 19:549–563 2007). However none of the proteins predicted to be involved in xylan synthesis have been biochemically characterized. We are using a multifaceted approach to provide insight into mechanisms of xylan synthesis. A series of *Arabidopsis* mutants have been isolated corresponding to genes that encode putative glycosyl transferases, which are highly expressed during secondary growth. The plants have been analyzed to determine the effects of the mutations on plant fitness as well as how the amount, structure, and extractability of xylan synthesized by these plants is altered. Through the use of this approach, we have identified several good candidate genes for the improvement of biomass recalcitrance that influence the structure and quantity of xylan, yet are not critical to plant development. We are now extending this knowledge to alter cell wall structure in *Populus*, a model energy crop.

Furthermore, we have developed an *in vitro* xylan biosynthesis assay using microsomal membranes, fluorescence-labeled oligosaccharide acceptors, and UDP-sugar donors. In-depth structural characterization by NMR spectroscopy of the reaction products allowed us to detect and confirm several glycosyltransferase activities related to xylan synthesis. We are currently using this *in vitro* assay to identify and evaluate biosynthetic chemotypes in our series of *Arabidopsis* mutants with defective secondary wall formation. This system has also been applied to microsomes isolated from *Populus* trees and switchgrass, allowing the mechanisms of xylan synthesis to be compared to what has been observed in *Arabidopsis*. Switchgrass microsomes contain enzymes that, in the presence of UDP-Xyl, also extend the labeled acceptors. Our results suggest that the xylan backbone is extended by a comparable mechanism in monocots and dicots. The results of these studies will be presented their implications for the mechanism of xylan biosynthesis will be discussed.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

66

Immunolocalization of Cell Wall Carbohydrate Epitopes in Switchgrass (*Panicum virgatum*)

Utku Avci^{1,3*} (uavci@ccrc.uga.edu), Michael G. Hahn,^{1,3} and Martin Keller^{2,3}

¹Complex Carbohydrate Research Center, University of Georgia, Athens; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 65.

Plant cells are surrounded by a dynamic extracellular matrix called the cell wall that undergoes modifications in composition and configuration depending on the developmental stage and external factors such as environmental stress or interactions with pathogenic and symbiotic microbes. In order to better understand the cell wall architecture, it is necessary to visualize the components of the cell wall in situ. Cell wall-directed specific monoclonal antibodies are useful tools to examine the distribution of wall polymers in plant cell walls at the whole plant, tissue, cell, and sub-cellular levels. These antibodies can also be used to monitor changes in cell wall composition as a function of plant development and in response to mutational and environmental influences.

A toolkit of monoclonal antibodies (~170) against diverse cell wall polysaccharides structures now exists that includes antibodies recognizing epitopes present in most major classes of wall polymers. These antibodies are available to the research community largely through antibody stock centers: (http://www.crcr.uga.edu/~carbosource/CSS_home.html and <http://www.plantprobes.net/>).

The monocot, switchgrass (*Panicum virgatum*), is currently of interest as a potential source of biomass for biofuel production. Here we show that the antibody toolkit can be used to localize wall polysaccharide epitopes in switchgrass leaf blade, leaf sheath, and stem. The epitope localization patterns observed switchgrass are similarly complex to those observed in dicots, such as *Arabidopsis*, although the patterns are different for several antibodies. For example, xylan-directed antibodies label a much broader diversity of cell types in switchgrass than has been observed in *Arabidopsis*. In contrast, homogalacturonan-directed antibodies label only cell corners and cell-cell interfaces in switchgrass, in contrast to *Arabidopsis* where these antibodies label whole cell walls. A detailed understanding of polysaccharide localization patterns has the potential to identify potential targets for either modification or degradation that will lead to more efficient deconstruction of biomass into fermentable monosaccharides for biofuel production.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

67

Changes in Cell Wall Composition and Structure of Alfalfa Reduced Lignin Lines That Might Influence Biomass Recalcitrance

Sivakumar Pattathil^{1,2*} (siva@ccrc.uga.edu), Trina Saffold,^{1,2} Malcolm O'Neill,^{1,2} Jadel Wasson,^{1,2} Jeff Miller,¹ Fang Chen,^{1,3} Richard A. Dixon,^{1,3} William S. York,^{1,2} Michael G. Hahn,^{1,2} and Martin Keller^{1,4}

¹BioEnergy Science Center, Oak Ridge, Tenn.; ²University of Georgia, Athens; ³Samuel Roberts Noble Foundation, Ardmore, Oklahoma; and ⁴Oak Ridge National Laboratory, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 65.

Genetic modification of biomass crops is a widely used strategy for obtaining less recalcitrant plants with suitably altered cell walls that can be efficiently used for biofuel production. Lignin is one of the plant cell wall components that significantly contributes to biomass recalcitrance. Hence, efforts are underway to obtain biomass crops with reduced levels of lignin in their cell walls. In this regard, two alfalfa (*Medicago sativa*) lines were developed in which two important genes encoding the lignin biosynthetic enzymes, 4-coumarate 3-hydroxylase (C3H) and hydroxycinnamoyl-CoA shikimate/quinic transferase (HCT), were silenced. In depth characterization of the cell wall polysaccharides in these mutants has been carried out using chemical and immunological approaches and compared with walls from wild-type plants. The cell wall preparations were sequentially extracted with a series of solutions of increasing alkalinity and the amount of material extracted in each step was quantified by gravimetric and various colorimetric techniques. The glycosyl residue compositions of the solubilized materials were determined by GLC analysis of the alditol acetate derivatives and by HPAEC-PAD of the underivatized glycoses. The types of polysaccharide that were present in each fraction were also studied using NMR spectroscopic analysis of the per-O-acetylated material. In parallel to these chemical studies, ELISA analyses (Glycome Profiling) of the extracted materials were carried out using a library of cell wall glycan-directed monoclonal antibodies to identify glycan classes eluting in each cell wall extract. Glycome profiles showed that the cell walls from lignin-reduced lines are altered in their extractability when compared to the walls of wild-type plants. For example, chlorite did not release any glucuronoarabinoxylan epitopes from HCT-knock-down cell walls compared to wild-type walls. In contrast, oxalate extraction released glucuronoarabinoxylan epitopes from walls of both reduced-lignin lines that are not released from w.t. walls. Overall, our studies indicate that reduction in lignin biosynthesis leads to a loosening of the wall leading to more facile release of polysaccharides. Cell walls that are more loosely held together may be less recalcitrant to deconstruction for the purposes of biofuel production.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

68

QTLs and Bioinformatically Identified Candidate Genes Underlying Lignin Content and Cell Wall Constituents Are Differentially Expressed in Stem and Root Tissues of *Populus*

G.A. Tuskan^{1,3*} (tuskanga@ornl.gov), R. Priya,^{1,3} T.M. Yin,^{1,3} M. Davis,^{1,2,3} U. Kalluri,^{1,3} X. Yang,^{1,3} S. Jawdy,^{1,3} L. Gunter,^{1,3} X.Y. Zhang,^{1,3} and M. Keller^{1,3}

¹Oak Ridge National Laboratory, Oak Ridge, Tenn.; ²National Renewable Energy Laboratory, Golden, Colo.; and ³BioEnergy Science Center, Oak Ridge Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 68-70): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and *Populus*) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generation of large number of natural and modified plant samples as well as detailed studies into poorly understood cell wall biosynthesis pathways.

The *Populus* activity tests existing genetic resources (in association and activation tag studies) as well as the direct transformation of *Populus* for several hundred genes of interest.

Quantitative trait loci (QTL) studies are an integral part of plant research and are used to characterize the genetic basis of phenotypic variation observed in structured populations and inform marker-assisted breeding efforts. These QTL intervals can span large physical regions on a chromosome comprising hundreds of genes, thereby hampering candidate gene identification. Genome history, evolution, and expression evidence can be used to narrow the genes in the interval to a smaller list that is manageable for detailed downstream functional genomics characterization. As a first step in this study, we analyzed the lignin content of 29 cell wall constituents both in stem and root in an inter-specific three generation hybrid poplar pedigree. By establishing a high density genetic map for this pedigree, QTL analyses were conducted to explore the underlying genetic loci. The major findings are as follows; a) Lignin content in stem is significantly higher than that in root, 2) Cell wall constituents

can be classified into four groups, with strong correlation only observed within organs (stem or root), 3) Pleiotropic QTLs are common and 4) QTLs are differentially expressed in stem and root. Our second motivation for the present study was to address the need for a research methodology that identifies candidate genes within a broad QTL interval. Towards this end, a bioinformatics-based approach for subdividing candidate genes within QTL intervals into alternate candidate groups of decreasing probability. Application of this approach in the context of studying cell wall traits, specifically, lignin content and S/G ratios of stem and root in *Populus* plants, resulted in the identification of manageable sets of genes of both known and putative cell wall biosynthetic function. These results provide a roadmap for future experimental work leading to identification of new genes controlling cell wall recalcitrance, and ultimately, in the utility of plant biomass as an energy feedstock.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

69

Transcriptome and Metabolome Profiling of *Populus* Tension Stress Response

U.C. Kalluri^{1,3*} (kalluriudayc@ornl.gov), T. Tschaplinski,^{1,3} X. Yang,^{1,3} S. Jawdy,^{1,3} G. Tuskan,^{1,3} E. Lindquist,² Z. Wang,² and M. Keller^{1,3}

¹Oak Ridge National Laboratory, Oak Ridge, Tenn.; ²DOE Joint Genome Institute, Walnut Creek, Calif.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 68.

Tension wood is a special type of reaction wood that is formed on upper side of bent stems and branches of angiosperm trees. The wood type is characterized by properties, such as high cellulose content, low lignin content, higher xylem cell number and increased secondary cell wall thickness relative to normal wood. Since these characteristics also constitute desirable feedstock properties, we designed an integrated study to understand the differential molecular and phenotypic properties that underlie tension stress response in *Populus* stems. Here we report results from the GC-MS based metabolite profiling of xylem and phloem tissue samples collected from normal, tension, and opposite wood types. Our analysis reveals several significant quantitative differences in metabolites between wood types (normal, tension and opposite), although the nature of the metabolites present didn't differ appreciably. The nature of the metabolites present differed markedly between genotypes and among tissue types (xylem and phloem). We also report summaries from paired-end read transcriptome data generated using the Illumina platform. This constitutes a large dataset with up to ~5.5 million reads/ library mapped

to the annotated *Populus* transcriptome covering ~77% of predicted gene models in the *Populus* genome version 2.0. The metabolomic data is being assessed in the context of the transcriptome and LC-MS/MS proteome profiles generate plausible hypotheses on mechanisms of tension wood formation and identify new genes involved in enhanced cellulose biosynthesis.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

70

The Use of Metabolomics to Characterize Extreme Phenotypes in a *Populus* Activation-Tagged Population

T.J. Tschaplinski^{1,5*} (t2t@ornl.gov), G.A. Tuskan,^{1,5} U. Kalluri,^{1,5} T. Yin,^{1,5} S. Jawdy,^{1,5} L. Gunter,^{1,5} M. Martin,^{1,5} N. Engle,^{1,5} X. Yang,^{1,5} V. Busov,² S. Strauss,^{3,5} X. Ma,^{3,5} M. Davis,^{4,5} R. Sykes,^{4,5} and M. Keller^{1,5}

¹Oak Ridge National Laboratory, Oak Ridge, Tenn.; ²Michigan Technological University, Houghton; ³Oregon State University, Corvallis; ⁴National Renewable Energy Laboratory, Golden, Colo.; ⁵BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 68.

Bioenergy feedstock species provide the raw materials for biochemical conversion into sugars and ultimately liquid transportation fuels. The plant cell walls of feedstock plants contain three major polymers including cellulose (a complex chain of glucose molecules and the most abundant polymer on earth), hemicellulose (a complex mixture of five and six carbon sugars) and lignin (a polyphenolic matrix that protects the cell wall sugars). These three polymers form a tightly bound cell wall matrix that causes raw feedstock material to be highly recalcitrant.

Understanding and overcoming cell wall recalcitrance to improve the yields of fermentable sugars from feedstock plants is a key goal of BESC. More specifically, the *Populus* activity has been tasked with understanding and overcoming recalcitrance in *Populus*, a potential bioenergy feedstock plant. While we understand that complete removal of lignin would be ideal for sugar yields, it would be fatal to the plant. Thus, finding an optimum phenotype with a ratio of lignin to cellulose that results high amounts of sugar *and* viability in the plant is one of our goals. To help achieve this goal we are currently utilizing existing genetic resources including a population of activation-tagged poplar trees established at a field site in Oregon. Gas chromatography-mass spectrometry (GC-MS)-based metabolomics is being used to characterize the clones that exhibit extreme wood chemistry phenotypes.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

71

Visualizing Supramolecular Cell Wall Degrading Enzyme Complexes and Aggregates

Thomas Haas,^{1,3} Bryon Donohoe,^{1,3} Hui Wei,^{1,3} Yunfeng Yang,^{2,3} Martin Keller,^{2,3} **Michael E. Himmel**,^{1,3} and Shi-You Ding^{1,3*} (Shi.you.Ding@nrel.gov)

¹National Renewable Energy Laboratory, Golden, Colo.;

²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and

³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 71-74): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing [CBP]).

BESC research in biomass deconstruction and conversion targets CBP by studying model organisms (e.g., *Clostridium thermocellum*) to understand novel strategies and enzyme complexes for biomass deconstruction. We are also searching natural diversity for enzymes and microbes - in particular, thermophilic anaerobes.

BESC is studying and modeling the structure and activities of multi-enzyme complexes to design or screen for variants with better cell wall deconstruction capabilities.

Among the many biomass-digesting microorganisms, a number produce structured biomass-degrading enzyme complexes. These complexes, called cellulosomes, are known to contain a variety of biomass-degrading enzymes docked to structural proteins termed "scaffoldins," which also often contain carbohydrate binding domains. Cellulosomes and their structural and enzymatic components may play important roles in bioenergy production and in future bio- and nanotechnologies. Several cellulolytic members of the bacterial genus *Clostridium* produce cellulosomes. In order to understand the structure and organization of cellulosomes the biomass-degrading properties of these organisms, we have employed electron tomography of high-pressure frozen/freez substituted *C. cellulolyticum* cultures grown on native switchgrass to examine the complex 3-D ultrastructure of the whole, intact cell wall degrading system at 3-5 nm resolution. We also employed immunolabeling techniques and transmission electron microscopy (TEM) to detect two major cellulosome components *in situ*: A processive endocellulase, Cel48F, and a scaffoldin, CipC. Our

observations show that the cellulosome allows *Clostridium cellulolyticum* to employ cell wall deconstruction mechanisms that differ from the mode of action of fungal free-enzyme digestion. Electron tomography has also revealed structural details of the tethers that anchor cellulosomes to bacterial cells.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

72

High Throughput Pretreatment and Enzyme Hydrolysis: A Massively Parallel Approach

Stephen R. Decker* (steve.decker@nrel.gov), Michael J. Selig, Melvin P. Tucker, Mark Davis, Robert Sykes, Kristen Gracom, Michael E. Himmel,^{1,3} and Martin Keller^{2,3}

¹National Renewable Energy Laboratory, Golden, Colo.;

²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and

³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 71.

The complexities of plant cell walls exist on chemical, macrostructural, and ultrastructural levels, rendering a good understanding of recalcitrance a highly daunting undertaking. Understanding how this extensive heterogeneity contributes to cell wall recalcitrance, however, is one of the keystone efforts in the BioEnergy Science Center. Several approaches are in place to tease apart this convoluted puzzle, including efforts in plant breeding, natural variation screening, genetic mapping and sequencing, and *in planta* molecular gene manipulation. While looking for changes in a few mutants or natural variants is very straightforward, migrating the pretreatment, enzyme digestion, and analytical techniques to a massively parallel pipeline capable of handling thousands of small samples each month is exceedingly difficult. Despite the inherent challenges of this undertaking, developing and implementing a high throughput pipeline capable of screening huge libraries of plant variants is one of the key achievements of the BESC. Here we detail the problems and solutions of designing, engineering, building, and implementing the world's first massively parallel biomass pretreatment and enzyme hydrolysis pipeline to measure a realistic biomass recalcitrance phenotype.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

73

Critical Enzymes for Lignin Degradation

Simo Sarkanen^{1*} (sarka001@umn.edu), Yi-ru Chen,¹ Frank Summerfield,¹ Yun-Yan Wang,^{1,3} and Martin Keller^{2,3}

¹University of Minnesota, St. Paul; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 71.

The discovery of lignin depolymerase enzymes that effectively degrade lignins under ambient conditions will contribute to dramatic reductions in pretreatment costs for the saccharification and fermentation of lignocellulose to liquid fuels.

In the early 1980's, it was thought that the initial depolymerization of lignins at the hands of white-rot fungi is primarily caused by hydroxyl radicals. Certainly, hydroxyl radicals can oxidize and cleave lignin macromolecules but, from 1983 onwards, such matters received less emphasis owing to the claim that fungal peroxidases and laccases might exhibit ligninolytic properties. However, for mechanistic reasons that are well understood, these enzymes maintain a poise between cleaving and polymerizing lignin preparations, wherein depolymerization becomes more prominent only at low substrate concentrations. Thus, it has not yet been possible to cleave polymeric lignin preparations completely by enzymatic means *in vitro*, even though extensive degradation is readily achieved with hydroxyl radicals. Consequently, it is not surprising that attention has returned in recent years to the role of reactive oxygen species as agents of fungal lignocellulose degradation.

True lignin depolymerase activity. The availability of the genomes of two white-rot fungi and one brown-rot fungus has now made it possible to consider the protein models with predicted secretion signals from three closely related basidiomycetes. Such comparisons draw attention to particular kinds of enzymes that have not previously been directly implicated in ligninolysis. In comparing the functional secretomes of the white-rot fungi, *Trametes cingulata* and *Phanerochaete chrysosporium*, with that of the brown-rot fungus, *Postia placenta*, it must be remembered how the activities of these basidiomycetes differ toward lignin degradation. White-rot fungi are capable of degrading large proportions of the lignin in lignocellulose completely; however, brown-rot fungal activity results in a marked disappearance of the aliphatic side-chain structures between the aromatic rings that are not themselves cleaved. From this perspective, some new kinds of extracellular fungal enzymes have emerged as likely agents of lignin depolymerization. They are distinguished by their inability, from a mechanistic point of view, to polymerize lignin components, and thus they may be expected to play a prominent role in lignin degradation *in vivo*.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

74

Understanding the Cellulosome and Its Assembly: Towards Improving the CBP Process

Yannick J. Bomble,^{1,5*} (Yannick.Bomble@nrel.gov), **Michael F. Crowley**,^{1,5} Mark R. Nimlos,^{1,5} Markus Alahuhta,^{1,5} Qi Xu,^{1,5} Vladimir V. Lunin,^{1,5} Edward A. Bayer,² John W. Brady,^{3,5} Jiancong Xu,^{4,5} Moumita Sarahay,^{4,5} Jeremy C. Smith,^{4,5} Michael E. Himmel,^{1,5} and Martin Keller^{4,5}

¹National Renewable Energy Laboratory, Golden, Colo.; ²Weizmann Institute of Science, Rehovot, Israel; ³Cornell University, Ithaca, N.Y.; ⁴Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁵BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 71.

The mechanism by which *C. thermocellum* and other cel-lulosomal cell wall degrading enzymes assemble on the scaffoldin is not currently known. In our study, we focused on the cellulosome-integrating protein (Cip) A of *C. thermocellum* and key cel-lulosomal enzymes from families 5, 9 and 48. We developed the first coarse-grained model to study the formation and function of a cellulosome assembly within CHARMM. This study aims at understanding the mechanisms involved in the sequential binding of the cel-lulosomal enzymes to the CipA scaffold of *C. thermocellum*. Understanding this mechanism is essential in order to design efficient engineered cel-lulosomes. Also, individual subdomains acting on cellulose surfaces or with individual cellulose chains were studied using molecular dynamics and normal mode analysis. These domains include: catalytic domains, carbohydrate binding domains, fibronectins, and the immunoglobulin-like domain. We have studied extensively the complex, seven-domain family 9 enzyme from *C. thermocellum*, CbhA. From this work, several new protein structures were determined, including CBM4, FNIII2, FNIII3, and CBM3b. Taken *in toto*, these 7 subdomains have enabled novel computational studies providing new insights in the role of this enzyme and the function of its individual subdomains. This new understanding of cel-lulosome function will be combined with novel mutational strategies to modify the *C. thermocellum* cel-lulosome to yield superior cell wall degrading *C. thermocellum* strains.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

75

Application of Phylogenomic Techniques in Studying Glycosyltransferase and Glycoside Hydrolase Families

Yanbin Yin^{1*} (yinyb@csbl.bmb.uga.edu), Wen-Chi Chou,¹ **Ying Xu**,^{1,3} and Martin Keller^{2,3}

¹University of Georgia, Athens; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ³BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 75-77): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC researchers in characterization, modeling, education and data management seek to provide integrating data and understanding across the center. We use advanced technologies to analyze chemical and structural changes within biomass. We have developed a coordinated characterization pipeline to provide composition, pretreatment, and enzyme digestibility data from native and modified plant samples and to store this data within a data management system. We are developing knowledge environments to allow access to improved data analysis and modeling tools.

Combining experimental data with biological models will help define the genomic and physical basis of plant cell-wall recalcitrance and deconstruction.

Glycosyltransferase (GT) and glycoside hydrolase (GH) families are enzymes that build and degrade, respectively, carbohydrates and other glycosylated molecules. For example, genes of GT2 family are known to encode cellulose synthases and hemicellulose backbone synthases; and at least 11 GH families are cellulases or glucanases. More than 200 GT and GH families have been categorized by the carbohydrate active enzyme (CAZy) database, corresponding to ~100,000 NCBI-nr proteins. Some of the CAZy families are very huge, for example, the glycosyltransferase family 2 (GT2) consisting of more than ten thousand proteins from various organisms including animals, plants, fungi and bacteria. In addition, many GT and GH families are composed of enzymes with different biochemical functions. Thus it will be valuable to the carbohydrate research field to classify the large CAZy families into smaller subfamilies, ideally, each of them having distinct biochemical function.

We have populated the CAZy families by including homologs from metagenomes and fully sequenced plant genomes. For 211 out of 292 CAZy families which have a Pfam domain/family model to represent them, we have

identified 26,924 homologs from the JGI (Joint Genome Institute) metagenomes and 126,796 homologs from the CAMERA (Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis) marine metagenomes. These numbers together already exceed the total number CAZy proteins in the NCBI-nr database. We then combined CAZy proteins from NCBI-nr, metagenomes and fully sequenced plant genomes for each of the 211 CAZy families. The resulting data set of each family was further sub-classified based on sequence similarity and also based on phylogenetic topology; the resulting sub-families were further used to build hidden markov models (HMMs). Some key GT families that are responsible for plant cell wall polysaccharide biosynthesis were examined in further detail, namely the cellulose synthase superfamily (Csl) of GT2 which also includes hemicellulose backbone synthases, putative pectin and xylan synthases related GT8, GT43 and GT47 families etc.

Our study doubled the current CAZy database by including metagenomic CAZy homologs. Our sub-classification of CAZy families into subfamilies which are represented by HMM models provides a new tool to annotate newly sequenced genomes in terms of their CAZy compositions. Moreover numerous metagenome-specific subfamilies were found after the sub-classification, representing novel CAZy subfamilies that are not found in the NCBI-nr database but instead are found exclusively in the environmental metagenomes. These novel CAZy enzyme subfamilies may have new functions that are particularly interesting for bioenergy related researches, for instance, more robust and stronger biomass breaking hydrolases.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

76

Developing Proteogenomics in Plants: Analysis of Proteomics Data Suggests Hundreds of Gene Model Corrections in *Populus*

Andrey A. Gorin^{1,2*} (agor@ornl.gov), Xinran Liu,¹ Grace M. Prazniak,¹ Robert M. Day,¹ Nikita D. Arnold,¹ Tamah Fridman,¹ Udaya C. Kalluri,^{1,2} Gregory B. Hurst,^{1,2} Richard J. Giannone,^{1,2} Robert Hettich,^{1,2} Edward Uberbacher,^{1,2} **Ying Xu**,^{2,3} and Martin Keller^{1,2}

¹Oak Ridge National Laboratory, Oak Ridge, Tenn.;

²BioEnergy Science Center, Oak Ridge, Tenn.; and

³University of Georgia, Athens

Project Goals: See goals for abstract 75.

Mass-spectrometry (MS) based proteomics provides the most direct evidence of gene product presence in specific samples, but MS data interpretation requires sophisticated computational algorithms to uncover the wealth of the encoded information. Our research aims to develop more

robust and reliable identification algorithms for shotgun tandem MS (MS/MS) experiments. More comprehensive proteome characterizations will lead to more complete assessments of functional potential for collected plant samples, as well as, to more confident reconstruction of protein co-expression subnetworks, identification of regulatory signals, etc.

Here we present a study that merges genomics and proteomics bioinformatics methods in order to identify new genes and refine existing gene models in *Populus trichocarpa*—one of the most important target plants for the Bioenergy Science Center research program. The study combined three elements (a) protein database obtained by six-frame translation of the *Populus* genome; (b) a collection of over 900,000 MS/MS spectra obtained from various plant tissues; (c) a novel computational algorithm that at least 2-times outperforms standard database search techniques to reveal 470 novel peptides, which are present in the plant tissue samples, but cannot be explained by the existing gene models. Several layers of control were implemented to ensure complete reliability of the findings: distractor database identifications were checked both at the level of unique peptides and at the level of detected open reading frames.

Confidently identified novel peptides were mapped to the current genome annotation resulting in over 80 predictions of novel genes and over 200 gene models that should have corrections of gene boundaries. Importantly, hundreds of gene model corrections were obtained during re-analysis of the existing data collected for other purposes, at no additional cost for sample preparation or MS/MS runs. Based on this study, the scale of the BESC proteomics effort (millions of spectra per year) represents an unparalleled opportunity to obtain very precise maps of gene positions and the actual proteome of *Populus* species, contributing to both the discovery of important molecular mechanisms (such as factors contributing to recalcitrance) and an improvement of gene annotation algorithms in plants.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

77

Linking Genomic and Biochemical Information to Identify Cellulolytic Enzymes: The GH5 Family Test Case

Brian J. Cantwell^{1,2,4*} (cantwellbj@ornl.gov), Leonid O. Sukharnikov,^{1,2,4} Kathrine Xue,³ Alborz Bejnood,³ **Igor B. Zhulin**,^{1,2,4} and Martin Keller^{2,4}

¹University of Tennessee, Knoxville; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; ³Oak Ridge High School, Oak Ridge, Tenn.; ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 75.

Tremendous advances in genome sequences make it possible to search for novel biochemical functions beyond the current capabilities of experimental biology. One of the grand challenges in bioenergy research is find better ways of degrading cellulose for its conversion to a biofuel ethanol. Novel cellulose-degrading capabilities can be found in genomic and metagenomic data collected from cellulose-degrading organisms and environments. However, there is a significant problem. Cellulases, enzymes that break down the cellulose polymer into metabolizable substrates, can be found within different families of carbohydrate-hydrolyzing enzymes that represent distinct protein folds. On the other hand, within any cellulose-containing specific fold/family there are closely related enzymes that have different substrate specificity (e.g. mannosidases, xylanases, etc.). This makes it extremely difficult to identify potential cellulases in genomic and metagenomic datasets. The goals of this work is to permit the accurate identification of cellulose degrading enzymes from amino acid sequence information and the prediction of organisms that are likely to efficiently utilize cellulose. Currently, hidden Markov models (HMM) are able to identify glycoside hydrolase domains based on conserved secondary structure, but the exact substrate specificity of these proteins cannot be determined. By combining bioinformatic and phylogenetic techniques with available biochemical information, we aim to improve the classification of potential cellulose degrading enzymes, enable sequence based prediction of substrate use, and identify residues critical to substrate specificity. Here we present a test case: the Glycoside Hydrolase 5 family.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

78

Plant Biomass Deconstruction by Extremely Thermophilic Anaerobes of the Bacterial Genus *Caldicellulosiraptor*

Irina Kataeva^{1,4*} (kataeva@uga.edu), Sung-Jae Yang^{1,4}, Aleksandar Cvetkovic,¹ Farris L. Poole II,^{1,4} Janet Westpheling,^{1,4} Richard J. Giannone,^{2,4} Robert L. Hettich,^{2,4} Sara Blumer-Schuette,^{3,4} Derrick L. Lewis,^{3,4} Inci Ozdemir,^{3,4} Amy L. VanFossen,^{3,4} **Robert M. Kelly**,^{3,4} **Michael W.W. Adams**,^{1,4} and Martin Keller^{2,4}

¹University of Georgia, Athens; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; ³North Carolina State University, Raleigh; and ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals (Abstracts 78-82): The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance.

BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing, or CBP).

BESC research in biomass deconstruction and conversion targets CBP by studying model organisms (e.g., *Clostridium thermocellum*) to understand novel strategies and enzyme complexes for biomass deconstruction. We are also searching natural diversity for enzymes and microbes—in particular, thermophilic anaerobes.

Microbial research targets how the CBP microbe interacts with the features of the pretreated or native plant cell wall and testing strategies for using pure or mixed cultures for degradation.

Very few cultivated microorganisms can degrade lignocellulosic biomass without chemical pretreatment. We have shown that some species of *Caldicellulosiraptor* efficiently utilizes various types of untreated plant biomass, as well as crystalline cellulose and xylan. In the case of *C. bescii* (previously termed *Anaerocellum thermophilum*), which grows up to 90°C, growth substrates included insoluble plant biomass obtained after washing (at 75°C for 18 h) hardwoods such as poplar and high lignin grasses such as switchgrass. The predominant end products from all growth substrates were hydrogen, acetate and lactate. *C. bescii* also grew well on first- and second-spent biomass, where spent biomass is defined as the insoluble growth substrate recovered after the organism had reached late stationary phase. Electron microscopy and growth studies indicate that *C. bescii* attaches dynamically to the plant biomass. The organism has been grown in 600-liter cultures on both crystalline cellulose and on switchgrass and the nature of its extracellular proteins are being investigated, using proteomic and transcriptional analyses. Genome sequences for several *Caldicellulosiraptor* species have been completed, facilitating functional genomics studies aimed at identifying specific ABC transporters for biomass-derived sugars and novel multi-domain glycoside hydrolases that deconstruct cellulose and hemicellulose. Since biomass deconstruction varies considerably across the genome-sequenced members of the genus *Caldicellulosiraptor*, efforts are underway to determine the differentiating features of this novel group of bacteria from this perspective.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

79

Characterizing Cellulose Hydrolysis and Ethanol Production by the Extremely Thermophilic Cellulolytic Organism *Caldicellulosiruptor obsidiansis*

Zhiwu Wang* (wangz2@ornl.gov), Scott Hamilton-Brehm, Jennifer L. Morrell-Falvey, Jonathan Mielenz, Martin Keller, and **James G. Elkins**

Oak Ridge National Laboratory, Oak Ridge, Tenn. and BioEnergy Science Center, Oak Ridge Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 78.

Biofuel production from renewable lignocellulosic biomass is dependent on the efficient enzymatic hydrolysis of cellulose in order to release fermentable sugars. Cellulolytic microorganisms inhabiting geothermally heated terrestrial hot springs are known to possess multidomain/multifunctional cellulases and hemicellulases that display increased heat-stability. *C. obsidiansis* is a newly characterized extremely thermophilic anaerobe capable of hydrolyzing cellulose, xylan, and pretreated lignocellulosic biomass (switchgrass and *Populus*) while fermenting the resulting sugars into acetate, lactate, CO₂, H₂, and ethanol. In this study, we used confocal laser scanning microscopy and 3-D image reconstruction to monitor the spatial and temporal dynamics of colonization and degradation of cellulose by *C. obsidiansis*. These data indicate that *C. obsidiansis* forms colonies that spread horizontally and vertically after attachment on the cellulose membrane, forming depressions that punctured the substrate within 72 hours. The distance between the cells and the substrate surface played a crucial role in the conversion rate of the cellulose, and thus overall fermentation efficiency. In addition to these modeling efforts, we show that end-product profiles including ethanol can be influenced by applying different growth conditions. By understanding the mechanism of microbial colonization of recalcitrant polymerized sugars, it may be possible to manipulate the fermentation conditions and target genetic modifications to improve the utilization efficiency of substrate carbon to produce desired end-products such as ethanol.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

80

Advances in Organism Development for Consolidated Bioprocessing

David Hogsett^{1,4*} (dhogsett@mascoma.com), Vineet Rajgarhia,^{1,4} Kevin Wenger,^{1,4} Nicky Caiazza,^{1,4} John McBride,^{1,4} Allan Froehlich,^{1,4} William Kenealy,^{1,4} Chris Herring,^{1,4} Heidi Hau,^{1,4} Devin Currie,^{2,4} Adam Guss,^{2,4} Lee R. Lynd,^{1,2,4} and Martin Keller^{3,4}

¹Mascoma Corporation, Lebanon, N.H.; ²Thayer School of Engineering, Dartmouth College, Hanover, N.H.;

³Oak Ridge National Laboratory, Oak Ridge, Tenn.; and

⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 78.

Biomass recalcitrance - that is, the difficulty of converting solid biomass to reactive intermediates such as sugars - is the primary obstacle to cost-effective production of ethanol and other fuels from cellulosic biomass. One-step microbial fermentation of pretreated cellulosic biomass without added enzymes, referred to as consolidated bioprocessing (CBP) is increasingly recognized as a potentially game-changing approach by which to overcome biomass recalcitrance. Organism development for CBP can proceed by one of two strategies: improve ethanol production by cellulose-utilizing microbes, and improve cellulose-utilization in microbes that produce biofuels well.

Results will be presented on the development of both eukaryotic and prokaryotic microorganisms capable of directly converting lignocellulosic biomass to ethanol with little or no addition of traditional, commercial cellulases. Data will be presented on the development of yeast strains which express multiple cellulases and are capable of CBP of pretreated materials to ethanol. Additional data will be presented on the development of thermophilic anaerobes for use in CBP, including development of cellulase expression in the pentose-fermenting *Thermoanaerobacterium saccharolyticum* and metabolic engineering of naturally cellulolytic *Clostridium thermocellum* to increase ethanol yields.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

81

A Proteomic Approach to Quantifying the Mass Concentration of Cellulase Enzymes Produced by *Clostridium thermocellum*

Lois St. Brice^{1,4*} (Lois.A.St.Brice@dartmouth.edu), Andrew Dykstra,^{2,4} **Bob Hettich**,^{2,4} Miguel Rodriguez,^{2,4} Babu Raman,^{2,4} Joel Kreps,^{3,4} Lee R. Lynd,^{1,4} and Martin Keller^{2,4}

¹Thayer School of Engineering, Dartmouth College, Hanover, N.H.; ²Oak Ridge National Laboratory, Oak Ridge, Tenn.; ³Verenium, San Diego, Calif.; and ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 78.

To accurately assess rates of microbial cellulose utilization (MCU), it is essential to independently determine cell, cellulase and cellulose mass concentration. Such ability would enable us to directly address questions related to substrate utilization, allocation of cellular resources between cell and cellulase synthesis, cell- and cellulase-specific cellulose hydrolysis rates and bioenergetics. Current methodologies for cellulase determination involve work-intensive purification procedures. In this study we seek to develop and validate a method for cellulase determination which involves minimal manipulation of a fermentation sample. Using proteomic protein determination, we seek to reliably and robustly predict mass concentration of cellulase across varying growth conditions, substrates and cellulase types (cell free vs. cell associated). Our goal was to identify a core group of cellulosomal proteins from *Clostridium thermocellum* which can be assayed using proteomics to determine total cellulosomal protein. Ten proteins that comprise approximately 90% of total cellulosomal proteins have been identified in *Clostridium thermocellum* fermentations. From these ten proteins, 40 peptides have been selected for targeted analysis to determine cellulase mass concentration in cell digest samples. Good candidate peptide sequences were selected to use for quantification based on Mudpits, LTQ and triple quadrupole MS measurements. We analyzed the candidate proteins for variability in the fraction of total cellulosomal mass represented in samples from varying conditions. We optimized peptide selection for these protein components ensuring representative unique peptides with good signal quality for all proteins of interest were selected. We also examined the relationship between proteomic determination of total cellulase based on the peptides selected for analysis and protein determination using conventional protein measurement techniques like the Bradford assay. Once we established the core group, good peptides that could reproducibly be used to quantify them, and the correlation between the protein determination methods, a prediction capability curve was established to determine the total amount of cellulase in a sample.

The BioEnergy Science Center is a U.S. Department of Energy

Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

82

Deletion of Cel48S from *Clostridium thermocellum* and Its Affect on Cell Growth and Cellulosome Function

Daniel Olson^{1,2,4*} (dolson@mascoma.com), Shital Tripathi,² Richard Giannone,^{3,4} Nicky Caiazza,^{2,4} David Hogsett,^{2,4} **Lee Lynd**,^{1,4} and Martin Keller^{3,4}

¹Thayer School of Engineering, Dartmouth College, Hanover, N.H.; ²Mascoma Corporation, Lebanon, N.H.; ³Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁴BioEnergy Science Center, Oak Ridge, Tenn.

<http://bioenergycenter.org>

Project Goals: See goals for abstract 78.

A method was developed for making targeted gene deletions in *Clostridium thermocellum* and used to delete the *cel48S* gene. This gene was chosen as our initial target because it has been widely studied, it makes up a large portion of the cellulosome and it is believed to play a key role in cellulose solubilization. The resulting *cel48S* mutant strain was analyzed for its ability to grow on crystalline cellulose and was found to digest it as completely as its parent strain, a surprising result. In fact, the only difference was a slight decrease in the rate of growth, and a reduction in biomass production when grown on Avicel. The deletion of *cel48S* is the first report of targeted deletion of a component of the *C. thermocellum* cellulosome.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

83

BioEnergy Science Center Education and Outreach

S. Kral,^{1,4} J. Griffin,¹ H. Schulson,¹ P. Doney,^{2,4} S. Fowler,^{3,4} B. Davison,^{3,4} M. Keller,^{3,4} and **J. Westpheling**^{2,4*} (janwest@uga.edu)

¹Creative Discovery Museum, Chattanooga, Tenn.; ²University of Georgia, Athens; ³Oak Ridge National Laboratory, Oak Ridge, Tenn.; and ⁴BioEnergy Science Center

<http://bioenergycenter.org>

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach

to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitolerant microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing).

BESC researchers in characterization, modeling, education and data management seek to provide integrating data and understanding across the center. We use advanced technologies to analyze chemical and structural changes within biomass. We have developed a coordinated characterization pipeline to provide composition, pretreatment, and enzyme digestibility data from native and modified plant samples and to store this data within a data management system. We are developing knowledge environments to allow access to improved data analysis and modeling tools.

Education and outreach to the general public is critical in the acceptance and deployment of bioenergy. In addition to leveraging successful education and training programs already in place at our partner institutions, BESC has developed educational lessons and activities that target elementary and middle school children.

In addition to our efforts to prepare a new generation of scientists for the emerging fields of bioenergy through the interdisciplinary training of graduate students and postdocs, our center has taken a novel approach in that our education efforts begin with fifth graders. We have developed lesson plans aimed at 4th, 5th and 6th grades to educate and inform students about the basics of energy production and utilization. They include basic concepts such as the carbon cycle, lignocellulosic biomass as substrate for the production of biofuels as well as technical and economic obstacles to a biobased fuel economy. The hands-on activities and guided questions are also designed to meet educational objectives for these grades. These lessons have been piloted in a hundred classrooms in North Georgia and Tennessee and will be made available to schools nationwide in the fall of 2009. We have also begun to pilot interactive “science night” programs offered to students and the general public through local schools, museums and community centers. We will present details of some of the lessons and science night activities.

The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.