USDA/DOE
Plant Feedstocks Genomics for Bioenergy Program
Project Director/Principal Investigator Meeting

Town and Country Resort and Convention Center
San Diego, CA

January 10, 2014

Table of Contents

AGENDA

USDA/DOE Plant Feedstocks Genomics for Bioenergy
Principal Investigator/Project Director Meeting
January 10, 2014, San Diego, CA

8:00 **Introduction** – Cathy Ronning, Ed Kaleikau

Plenary Talks, 2011 Awardees

8:10 Todd Mockler, Donald Danforth Plant Science Center
Modulation of Phytochrome Signaling Networks for Improved Biomass Accumulation Using a Bioenergy Crop Model

8:35 John McKay (Jan Leach), Colorado State University
An Integrated Approach to Improving Plant Biomass Production

9:00 David Braun, University of Missouri
Functional Genomics of Sugar Content in Sweet Sorghum Stems

9:25 Steve Kresovich, Clemson University
Genomic and Breeding Foundations for Bioenergy Sorghum Hybrids

9:50 Jianming Yu, Iowa State University
Genomic Selection of Biomass Traits in a Global Collection of 976 Sorghum Accessions

10:15 **BREAK**

10:35 Laura Bartley, University of Oklahoma
Association Mapping of Cell Wall Synthesis Regulatory Genes and Cell Wall Quality in Switchgrass

11:00 Erik Sacks, University of Illinois Urbana-Champaign
Quantifying Phenotypic and Genetic Diversity of Miscanthus sinensis as a Resource for Knowledge-Based Improvement of M. ×giganteus (M. sinensis × M. sacchariflorus)

11:25 Eric Beers, **Virginia** Polytechnic Institute and State University
Functional Interactomics: Determining the Roles Played by Members of the Poplar Biomass Protein-Protein Interactome

11:50 Luca Comai, University of California Davis
Creation and High-Precision Characterization of Novel Poplar Biomass Germplasm

12:15 **LUNCH** (on your own)

"Speed Talks," 2013 Awardees

1:15 Sarah Lebeis (Jeff Dangl), University of North Carolina Chapel Hill
Functional Manipulation of Root Endophyte Populations for Feedstock Improvement

1:20 Katrien Devos, University of Georgia
Unraveling the Genetics of Two Key Biomass Traits That Differentiate Upland and Lowland Tetraploid Switchgrass Ecotype, Colonization by Mycorrhizal Fungi and Frost Tolerance

1:25 Yiwei Jiang, Purdue University
Genetic Control of Flowering in Switchgrass

1:30 Matias Kirst, University of Florida
Accelerated Development of Optimal Pine Feedstocks for Bioenergy and Renewable Chemicals Using Genome-Wide Selection

1:35 Ray Ming, University of Illinois Urbana-Champaign
Pyramiding Genes and Alleles for Improving Energy Cane Biomass Yield

1:40 A.S.N. Reddy, Colorado State University
Global Analysis of Epigenetic Regulation of Gene Expression in Response to Drought Stress in Sorghum

1:45 Steven Strauss, Oregon State University
Structural Polymorphisms as Causes of Heterosis in Populus

1:50 **KBase Update**
Doreen Ware (Cold Spring Harbor Laboratory/USDA ARS)
Dave Weston (Oak Ridge National Laboratory)

2:00 **Roundtable Discussion**

3:00 **BREAK**

3:20 **Poster session**

5:00 **Adjourn**

SPEAKER ABSTRACTS (In Presentation Order)

Modulation of Phytochrome Associated Gene Networks for Improved Biomass Yield in the Bioenergy Crop Model Brachypodium

Todd Mockler

Donald Danforth Plant Science Center

The genesis of this project was our observation that biomass accumulation was greater in Brachypodium, Switchgrass, and rice plants grown in monochromatic red light than in blue or white light, despite the levels of photosynthetically active radiation (PAR) being the same in all conditions. Our studies suggest that the enhanced biomass productivity observed in monochromatic red light may be a consequence of phytochrome signaling that modulates shade-avoidance and photosynthesis. Modulation of photosynthesis by manipulating phytochrome signaling may open the door to engineering crops for enhanced *intrinsic yield*, which unlike *yield protection*, has not necessarily been a target for crop improvement efforts to date. We have cloned and validated a collection of Brachypodium transcription factor (TF) ORFs. These include a number of TFs specifically of interest in this project based on their predicted functions in phytochrome signaling networks. We have also generated a collection of 768 synthetic promoters for use in yeast one-hybrid assays with Brachypodium TFs. We have developed and deployed a robotics system for automating high-throughput yeast one-hybrid screening. We are currently screening the collection of synthetic promoters using selected TFs. We have also collaborated to deploy a LemnaTec-based high-throughput plant phenotyping facility for high-resolution digital phenotyping of model plants. The phenotyping platform's controlled-environment chamber enables precise control of environmental variables, including light, temperature, and humidity, and weighing and watering stations allow for precise control and quantification of water availability and usage. To facilitate our study of Brachypodium transcription factors implicated in phytochrome signaling networks, in collaboration with Dr. James Carrington's group (Danforth Center) we have developed new constructs and methods for amiRNA knockdown of target genes in Brachypodium. These constructs are being used in the Brachypodium transformation pipeline to knock down target transcription factors; and the resulting transgenic lines will be phenotyped on the LemnaTec platform. We also used Brachypodium whole-genome microarray based gene expression profiling datasets to generate a co-expression based gene network model. We identified a network module strongly over-represented with genes implicated in or associated with photosynthesis. Closer examination of this module and identification of TFs having connections to the photosynthesis genes in this module revealed a subnetwork of putative photosynthesis regulating Brachypodium TFs and their targets that includes several of the TFs we had originally prioritized for *in planta* perturbation in this project. We are focusing current yeast one-hybrid studies on interrogating the

promoters of the identified subnetwork of Brachypodium photosynthesis genes with the identified putative phytochrome-signaling associated TFs. We are also prioritizing these TFs for *in planta* overexpression and knockdowns.

An Integrated Approach to Improving Plant Biomass Production

John Mckay

Project Directors: John McKay, Daniel Bush[1], Andrew Kern[2], Hei Leung[3], Bingyu Zhao[4], and Jan E. Leach[1]

[1]Colorado State University; [2]Rutgers University; [3]International Rice Research Institute; [4]Virginia Tech University

Our goal is to translate functional genetic understanding of traits important for biomass production from rice to switchgrass. In this presentation, we report our progress in two areas: first, the development of genetic and genomic resources for switchgrass, and second, the development of a high-throughput phenotyping platform to assess biomass traits in segregating plant populations as a means to expedite gene discovery.

Using transcriptome data for two cultivars of switchgrass that exhibit significantly different biomass phenotypes, 'Alamo' and 'Dacotah', we produced *de novo* transcriptome assemblies, and used these to generate gene models and variant calls. We evaluated SNP discovery based on an increasing scale of included reads to estimate the variation within each cultivar. Even with deep transcriptome sequencing, the full breadth of variation in switchgrass was not fully captured, yet we identified differentially expressed transcripts that may relate to biomass differences. To develop genetic resources for biomass studies, we created a pseudo F2 population of 180 individuals by crossing 'Alamo' and 'Dacotah'. This population was established in replicated field plots, and is being evaluated for various agronomic/biomass traits. To create a coding sequence-based genetic linkage map for switchgrass, we are sequencing cDNA from these F2 plants. These data will provide (1) genome-wide SNP markers for QTL analysis, (2) quantitative data on gene expression differences segregating in this population, (3) data on recombination fractions and linkage among polymorphisms in transcripts. These latter data will allow for improving the genome assembly.

Thanks to technical innovations, measuring genome-wide DNA polymorphism is becoming trivial, leaving high throughput, precision phenotyping as the gauntlet for crop breeding. We have optimized a tractor-based spectral imaging system to measure 5,000 test plots in 2 h. We are using this tractor-based high throughput phenotyping system to discover alleles and high throughput traits that can be used to predict biomass and yield components. Field trials with 1400 rice RILs conducted in spring 2013 were scanned on a weekly basis during the growing season. In this trial, a large effort was focused on optimizing the spectral data collection system, which includes a suite of proximal sensors allows capture of phenotypic data for large populations via a number of spectral vegetation indices, such as NDVI for biomass, NBI for nitrogen content, and several other indices for photosynthetic activity, leaf area index (LAI), etc. To ground-truth the spectral data, we collected flowering time and various biomass-relevant data. Our results show a high level of genetic correlation between spectral measurements and key traits including biomass, height, and yield. The heritability of these proxy spectral traits looks promising for QTL mapping. This high throughput system will facilitate the identification of new

non-destructive measures that predict biomass and genes controlling other physiological traits important for plant growth and yield that are not accessible after typical end-of-experiment destructive harvesting.

Functional Genomics of Sugar Content in Sweet Sorghum Stems
David Braun

Project Director: David Braun, University of Missouri, Columbia, braundm@missouri.edu

Co-PDs: Ismail Dweikat, University of Nebraska, Lincoln, idweikat2@unl.edu
Benjamin Babst, Brookhaven National Laboratory, bbabst@bnl.gov
Rich Ferrieri, Brookhaven National Laboratory, rferrieri@bnl.gov

Abstract:
Enhancing the production and conversion of plant feedstocks to utilizable sources of energy will decrease the U.S. dependency on fossil fuels, while reducing greenhouse gas emissions. Sweet sorghum is a rapidly growing, high biomass, widely adaptable crop with tremendous potential for biofuel production. Sweet sorghum accumulates very high concentrations of sucrose in the stem, which can be efficiently converted to ethanol. We hypothesize that sucrose accumulation in sweet sorghum can be further improved if we understand the mechanisms regulating carbon allocation to stems. Our project is using a combination of approaches, spanning genomics, molecular genetics, biochemical phenotyping, and detailed physiological studies to identify bioenergy-relevant genes and to understand their functions in carbohydrate partitioning in sweet sorghum. Toward determining the mechanisms regulating the transport and partitioning of sucrose to sweet sorghum stem tissues, we are 1) mapping and characterizing quantitative trait loci (QTL) related to stem biomass, total biomass, and sugar accumulation in the stem in a population derived from a grain sorghum crossed by a sweet sorghum; 2) are determining the expression patterns and functions of a family of proteins, called sucrose transporters (SUTs), which are hypothesized to play essential roles in regulating carbon partitioning of sugar to sweet sorghum stems; and 3) are using detailed radiotracer (carbon-11) physiological and biochemical approaches to map and delineate the bottlenecks to sucrose allocation in sweet sorghum. From these data, we will generate a physiological knowledge base and determine which genes function in transport and allocation of assimilated carbon to stem tissues in sweet sorghum. The knowledge and germplasm generated by the project will make important contributions to advancing sweet sorghum for sustainable bioenergy production.

Genomic and Breeding Foundations for Bioenergy Sorghum Hybrids
Stephen Kresovich

Stephen Kresovich, Clemson University, skresov@clemson.edu; Andrew H. Paterson, University of Georgia; and F. Alex Feltus, Clemson University

The USDA (June 23, 2010 release) forecast that ≥ 90% of U.S. cellulosic bioenergy needs will be met through biomass production in the South, with sorghum identified as a key feedstock. In

contrast with other regions, the South has levels of sunlight and rainfall favorable for efficient production of biomass for bioenergy, and also has ample land not used for food production. The new sorghum bioenergy belt is likely to encompass a region from the Mississippi Valley, east across the Gulf states, and north through the Atlantic coast to Virginia. This new production area is vastly different from the current production areas for grain sorghum and is likely to have a unique set of biotic and abiotic production conditions and constraints that need to be addressed to maximize sorghum's potential for bioenergy production. As such, much of the previous breeding and genetics of sorghum will be of limited value for optimizing sorghum to the new bioenergy belt of the United States. Therefore, a new systems-oriented vision and scientific understanding are required for a new crop, in a new region, for a new use in a new century.

Sorghum is a species that possesses significant underexploited genetic and phenotypic variation for agricultural production of bioenergy. Historically genepools have been developed by genetic resources scientists and breeders for the production of grain and forage from sorghum. These classic ideotypes focus on a high harvest index for grain and leaf tissue, respectively. To make transformational progress in breeding sorghum for bioenergy production, new ideotypes, and their genetic bases, need to be targeted and genetically characterized. For both sweet and biomass types, hybrids must be developed. We envision sweet ideotypes to be integrated with current sugar cane production in Florida, Louisiana, and Texas. These ideotypes are constrained in their potential production region by the stability of the sugar solution accumulating in the stalk. We expect biomass types to be utilized for production of bioenergy based on cellulosic conversion and not be constrained by possible sugar solution degradation in the stalk. Therefore, biomass types can be produced throughout the new bioenergy belt.

We now have the genomic tools and datasets [DNA sequence (Paterson et al. 2009; Mace et al. 2013) and genotyping-by-sequencing (Morris et al. 2013) to dissect the genetic bases of useful traits, to establish associations of evolutionary haplotypes with environmental conditions to target new sources of genetic variation, and to use DNA-based markers to expedite breeding cycles and to test and optimize genomic selection strategies for crop improvement. This approach recently has been employed to associate desirable agroclimatic traits with DNA sequence variation (Morris et al. 2013) and to elucidate the genetic bases of grain quality traits in sorghum (Sukumaran et al. 2012). Building on this progress, now is the time to harness biological and genomic strengths of sorghum to address our energy needs in the U.S.

Genomic Selection of Biomass Traits in a Global Collection of 976 Sorghum Accessions
Jianming Yu

Project Director: Jianming Yu, Iowa State University, jmyu@iastate.edu

Co-PDs: Tesso, T., Kansas State University, ttesso@ksu.edu; Roozeboom, K., Kansas State University, sstaggen@ksu.edu; Wang, D., Kansas State University, wang@ksu.edu; Bernardo, R., University of Minnesota, bernardo@umn.edu; Wang, M., Plant Genetic Resources Conservation Unit, USDA-ARS, mingli.wang@ars.usda.gov; Pederson, G., Plant Genetic Resources Conservation Unit, USDA-ARS, gary.pederson@ars.usda.gov

Substantial genetic diversity exists in sorghum (*Sorghum bicolor* L. Moench), a key lignocellulosic biofuel species in the United States. The implementation of genomic, genetic tools to select and enhance current germplasm will greatly accelerate new variety development. Our objectives in this study are to address several key questions: 1) *How to tap into the vast plant germplasm collections for biomass crop improvement?* 2) *How to increase the information contained in genotypic and phenotypic data for the selected germplasm?* 3) *How to leverage a high-throughput phenotyping method such as Near Infrared (NIR) to facilitate plant biomass composition investigation?* 4) *How robust are the various genomic prediction models on biomass traits?* In this study, genotyping by sequencing (GBS) with the Illumina HiSeq platform was conducted for 976 sorghum accessions sampled from germplasm bank and generated 0.72 million SNPs. A set of 300 accessions, selected to be most representative from this panel were extensively phenotyped for biomass yield, plant height, stem diameter, stalk number, stalk lodging, and root lodging at Kansas and Taxes in 2012 and 2013. Cellulose and lignin content were investigated through both wet chemical and high throughput NIR methods. Bringing phenotype and genotype data together, genomewide prediction models were established with all common methods. Biomass traits of the 700 untested accessions were predicted by using the optimal genomic selection model and will be validated through phenotyping the accessions with extreme high and low biomass potential.

Gene Expression vs. Cell Wall Composition Correlations Vary Dramatically Among Switchgrass Genotypes

Laura Bartley

Laura Bartley[1], Prasenjit Saha[1], Fan Lin[1], and Nicholas Santoro[2]

[1] University of Oklahoma, Norman, OK
[2] GL-BRC, Michigan State University, East Lansing, MI

Efficient conversion of lignocellulosic biomass, such as that of switchgrass, into "second-generation" biofuels is influenced by the biomass' composition and structure. In grasses, lignin and other phenylpropanoids, such as *p*-coumaric acid (*p*-CA) and ferulic acid, reduce cell wall sugar accessibility hampering biochemical fuel synthesis. Transgenic silencing of lignin biosynthesis enzymes and regulators improves cell wall digestibility; however, questions remain as to the relative contribution of different molecular players in controlling biomass content. In this study, we characterized cell wall parameters and lignin biosynthesis gene expression in three switchgrass genotypes (AP13, A4, and VS16) for tillers at three developmental stages (V3, E4, and R3) and for the E4 stage, three segments (S1-S3). The genotypes have similar patterns of cell wall content across development but distinct gene expression. Correlation analysis between gene expression and components reveals that comparing expression at stage N to cell wall content at stage N+1, abolishes many unexpected correlations between sugar content and lignin gene expression compared with N vs. N comparisons. For the whole data set, expression in the Nth stage of "COMT3" is most highly correlated with lignin and *p*-CA in the N+1*th* stage. Still, gene-component correlation patterns for each genotype are distinct, and in some cases opposite in different genotypes. This may be due to the complexity of functions of cell walls, expansion of biosynthesis gene families in tetraploid switchgrass, and multiple layers of regulation. Moreover,

approaches for improving biomass content in switchgrass that rely on gene silencing may not be generalizable across genotypes.

Quantifying Phenotypic and Genetic Diversity of *Miscanthus sinensis* as a Resource for Knowledge-Based Improvement of *M.* ×*giganteus* (*M. sinensis* × *M. sacchariflorus*)

Erik J. Sacks

PIs: Erik J. Sacks, University of Illinois, esacks@illinois.edu
Joe Brummer, Colorado State University, Joe.Brummer@ColoState.EDU
Xiaoli Jin, Zhejiang University, jinxlzju@163.com
Stephen Long, University of Illinois, slong@illinois.edu
Toshihiko Yamada, Hokkaido University, yamada@fsc.hokudai.ac.jp
Chang Yeon Yu, Kangwon National University, cyyu@kangwon.ac.kr

Abstract:
Miscanthus is a perennial C_4 grass that is a leading potential feedstock for the emerging bioenergy industry in North America, Europe, and China. However, only a single sterile genotype of *M.* ×*giganteus*, a nothospecies derived from diploid *M. sinensis* and tetraploid *M. sacchariflorus*, is currently available to farmers for biomass production. To facilitate breeding of *Miscanthus*, we characterized genetic diversity and population structure of *M. sinensis*. We studied 767 accessions, including 617 *M. sinensis* from most of its native range in China, Japan, and South Korea, 77 ornamental cultivars from the U.S. and Europe, and 43 naturalized individuals from the U.S. The accessions were screened with 21,207 RAD-Seq SNPs obtained via the UNEAK pipeline in TASSEL, 424 GoldenGate SNPs, and ten chloroplast microsatellite markers. Population structure was analyzed with STRUCTURE 2.3.4, and via Discriminant Analysis of Principle Components (DAPC) using the R package adegenet.

We identified six genetic clusters of *M. sinensis* from geographically distinct regions in Asia. Genetic data indicated that 1) Southeast China was the origin of *M. sinensis* populations found in temperate eastern Asia, which was consistent with this area having likely been a refugium during the last glacial maximum (LGM), 2) *M. sinensis* migrated directly from Southeast China to Japan before migrating to the same latitudes in China and Korea, which was consistent with the known sequence of warming post-LGM, 3) ornamental *M. sinensis* cultivars from the U.S. and Europe were derived from the Southern Japan population, and U.S. naturalized populations were derived from a subset of the ornamental cultivars, 4) many ornamental cultivars previously considered to be entirely *M. sinensis* have, in actuality, hybrid ancestry from *M. sacchariflorus* and *M. sinensis*, whereas U.S. naturalized populations of *M. sinensis* do not, and 5) curiously, the four *M. floridulus* we studied, which were from China, Japan, New Caledonia and Papua New Guinea, grouped with the SE China population of *M. sinensis*, calling into question the validity of two different species names.

Allelic diversity of the candidate gene, *Hd1*, which in grasses typically acts as repressor to flowering under long day conditions, was also studied. We cloned and sequenced *Hd1* for 38 *M. sinensis* accessions that originated from a broad range of latitudes in Japan and China. Non-functional alleles due to insertions and deletions were found. MITEs that produced a premature

stop codon were found in *Hd1* alleles of Japanese accessions but not Chinese accessions. Differing patterns of evolution for *Hd1* among populations suggest opportunities to obtain transgressive segregants for flowering time.

These results have, for the first time, provided a regional-level understanding of population structure in *M. sinensis*, and insights into its most recent phase of evolution. The genetic bottleneck associated with U.S. *M. sinensis* germplasm was a previously unknown limitation to breeding improved bioenergy feedstock cultivars of *Miscanthus*. The present study indicates that there is an opportunity to broaden the genetic diversity of *M. ×giganteus* by using *M. sinensis* parents from populations other than Southern Japan.

Functional Interactomics: Determining the Roles Played by Members of the Poplar Biomass Protein-Protein Interactome

Eric Beers

Eric Beers, Amy Brunner, Allan Dickerman, Richard Helm,Virginia Tech, Blacksburg, VA 24061

The main goal of the poplar biomass interactome project is to use yeast two-hybrid assays to discover novel networks of interacting proteins, which in turn become the subjects of more detailed analyses involving independent interaction assays and transgenic poplar. To accomplish this we have cloned more than 300 ORFs corresponding to genes that are highly upregulated in the xylem of *Populus trichocarpa*. Selected ORFs have been screened in a binary yeast two-hybrid matrix and against a xylem cDNA prey library (http://xylome.vbi.vt.edu/index.html). Full-length clones based on selected interaction sequence tags are prepared as DNA-binding domain and activation domain fusions and retested for ability to activate yeast reporters. To date we have detected over 200 protein-protein interactions (library and binary screens combined) involving 74 bait proteins. These interactors comprise some previously reported as well as novel protein-protein interaction networks involving cytoskeletal and trafficking proteins, signal transduction GTPases and protein kinases, transcription factors, ubiquitin/26S proteasome pathway proteins, disordered proteins, domain of unknown function (DUF) proteins, chromatin modifiers, and cell wall biosynthesis proteins. Edges connecting some of these networks have been identified, facilitating the assembly of a more integrated poplar biomass protein-protein interaction network. The following three interaction networks are currently being functionally characterized. **1)** XND1 (a NAC domain protein and a strong negative regulator of xylem differentiation) and its orthologs from several eudicots and monocots are RETINOBLASTOMA-RELATED (RBR)-binding proteins. XND1 also interacts with VND6 and VNI2, two other NAC proteins that regulate xylem differentiation. We are investigating the potential for RBR to negatively regulate xylem differentiation through interactions with NAC proteins. **2)** PtrDRIF is composed of two distinct domains: an N-terminal Myb–like domain, which interacts with Myb proteins PtrRAD and PtrDIV, and the C-terminal DUF3755-containing domain, which interacts with the xylem-associated homeodomain proteins PtrHB1.3 and PtrKNAT7.1. Results indicate that PtrDRIF serves as a scaffold for Myb-homeodomain protein complexes in poplar wood-forming tissue. **3)** PB504, a novel endomembrane protein of unknown function, was found to interact with cellulose synthase. We have extended the PB504 network, which now includes a

small GTPase, a Ser/Thr protein kinase, and several uncharacterized putative cytoskeletal proteins. We are currently characterizing constitutively active and dominant negative forms of the PB504-interacting small GTPase and protein kinase in transgenic plants. In November 2013, we planted over 500 transgenic poplar trees for field studies. We are also using *PtrNAC154* overexpression lines to test the hypothesis that genetic modification of poplar trees results in chemotypes that can be assessed by global metabolite profiling.

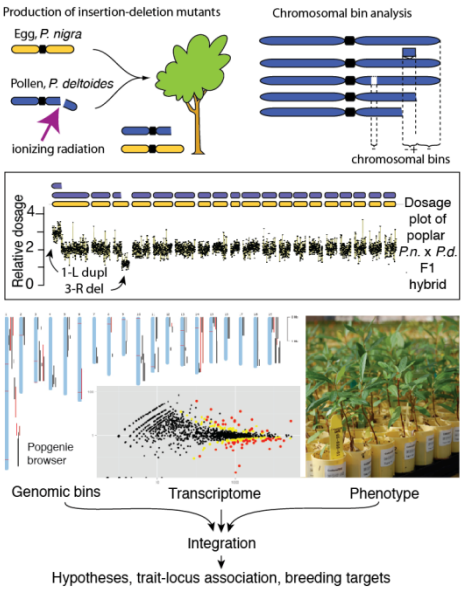
Creation and High-Precision Characterization of Novel Poplar Biomass Germplasm

Luca Comai

Luca Comai, P.D.[1,2], Isabelle M. Henry[1,2], Matthew S. Zinkgraf [3], Andrew Groover [3,1], co-P.D.
[1]Department of Plant Biology, [2]Genome Center, University of California at Davis;
[3]US Forest Service, Davis

The core goal of poplar biomass breeding is the creation of novel genotypes with superior performance. The poplar clones with the highest biomass yields are created through hybridization between species. Altered gene dosage relationships are believed to play a causative role in hybrid performance. The goal of our project is to identify, generate and characterize dosage-dependent variation and to facilitate its manipulation to breed poplar for biomass production. Our objectives are:

1. Survey genomic composition in commercial pedigrees of Populus F1 hybrids.
2. Manipulate gametic contribution for functional genomic and germplasm enhancement
3. Correlate variation in karyotype, gene dosage, and transcriptional modules with superior biomass production



We established a method for high throughput dosage and genotype analysis through Illumina sequencing. Our bioinformatics platform parses dosage information per genomic bin yielding dosage variation ranging from whole chromosome aneuploidy to segmental indels. Variation in commercial germplasm ranged from pure triploids, to aneuploids displaying extra copies of chromosomes, to segmental aneuploids displaying missing or extra segments of one or more chromosomes. Surprisingly, we found aneuploidy in natural, adapted individuals of *P. trichocarpa*. Anecdotal and preliminary evidence indicates that these variants can be superior biomass producers. To create more diverse germplasm, we produced a population of dosage variants via gamma-irradiation of mature pollen. High pollen irradiation doses (125 Gy or higher) resulted in complete catkin abortion. Samples originating from pollen irradiated with 100 Gy were selected for an initial dosage analysis, from which 41/74 (55%) exhibited at least one dosage lesion. Our goal is to produce and characterize enough dosage variants to cover each gene at least once, preferably several times. Replicating these treatments, we created a core collection of ~600 poplar F1 hybrid clones predicted to saturate the poplar genome with 1-10Mb indels. The first cohort is undergoing

propagation and we expect to gather the first replicated phenotypic measurement in the coming months. Preliminary phenotypic observations are encouraging. Many of the aneuploids lacked an obvious phenotype and grow well, sometimes better than their diploid siblings. The aneuploids display the expected wide variation in trait values, for example in phenology, leaf shape or apical dominance. We have produced RNA-Seq libraries from a subset of individuals and are awaiting sequencing data. In summary, this population provides an excellent tool to investigate the effect dosage variation of specific chromosomal pieces or specific genes. The project will result in the creation of a valuable functional genomics resource for poplar, increased understanding of phenotypic variation associated with dosage variation, and insights into the mechanisms influencing superior performance of hybrids and ploidy variants. This will in turn facilitate breeding of well-characterized, superior biomass germplasm for direct use by the biofuel industry.

SPEED TALKS ABSTRACTS (In Presentation Order)

Functional Manipulation of Root Endophyte Populations for Feedstock Improvement

Sarah Lebeis

It is becoming increasingly clear that the plant function is co-dependent on the microbial community that exists within and around them, both above ground (phyllosphere) and underground (rhizosphere). The complex associations of plants and microbes can often benefit plant health and productivity in the face of changing environmental conditions by affecting host physiological traits including nutrient uptake, growth allocation, hormone signaling, catabolism of toxic compounds, and resistance to pathogens. Hence, the *plant microbiome* contributes to the *extended phenotype* of the combined plant and microbial genotypes to determine plant productivity. We will define and exploit plant-rhizosphere and plant-endophytic microbes that lead to changes in biomass and growth under various nutrient regimes. In particular, we are interested in microbes that are (1) enriched in the zone directly adjacent to, or on the surface of, the root (the rhizosphere and rhizoplane, respectively) and (2) that gain access to and colonize the interior of the root (endophytes). To fulfill this goal we have assembled two large, and expanding sets of microbial strains isolated from root systems of three diverse hosts. We will define strains, and mixtures of strains, that confer conserved plant phenotypes (performance under nutrient stress of in the presence of a root pathogen) that maintain taxonomic and functional diversity. The objective is to identify endophytic strains that robustly re-colonize and induce a similar host plant phenotype, such as biomass, across all three plant species, thereby providing potential candidate strains that act across host species. By developing plant-microbe re-colonization screens for subsets of cultured endophytes applied to sterile plants as a community, we can test (1) defined nutrient stresses (Phosphate, Nitrogen) or water stress, (2) spatial and developmental aspects of endophyte re-colonization, and (3) transcriptional changes in host and defined microbiome in combination with the single and multiple strain consortia. Additionally, we address how host genotype contributes to the assembly and function of the host plant extended phenotype. To date, only very small quantitative effects of endophyte or rhizosphere community structure have been attributed to host genotype. This may be due to taxonomic elasticity in the microbiome, at least as assessed by 16S sequence at the OTU level. We will deploy the *Arabidopsis thaliana* MAGIC lines and additional traditional two parent recombinant inbred lines (RIL), *Setaria italica* RILs (from Jeff Bennetzen's Lab), and 20-100 selected individuals from the BESC *Populus trichocarpa* association population to identify plant host genotypes that express different re-colonization profiles and plant productivity phenotypes from those of the reference genome. Emphasis will be placed on testing hypotheses that will ultimately result in design principles for constructing PGPR consortia for enhanced feedstock sustainability traits. We anticipate that our findings will be applicable over a broad host evolutionary space, and therefore translatable and transformative for feedstock production in general.

Unraveling the Genetics of Two Key Biomass Traits that Differentiate Upland and Lowland Tetraploid Switchgrass Ecotypes, Colonization by Mycorrhizal Fungi and Frost Tolerance

Katrien Devos

Katrien Devos, University of Georgia, Athens

Goal: To develop strategies for increased frost tolerance of lowland switchgrass through (1) identification of the genetic pathway(s) that provide frost-tolerance in upland switchgrass, and (2) studying the potential of beneficial fungi to minimize host cold stress. This project seeks to leverage the high biomass yield of southern-adapted lowland types and the frost tolerance of northern-adapted upland types to identify candidate genes that can be exploited to enhance biomass production of switchgrass under cold stress.

Genetic Control of Flowering in Switchgrass

Yiwei Jiang

Yiwei Jiang, Purdue University, West Lafayette IN

Goal: To elucidate the genetic mechanisms and identify candidate genes controlling flowering time in switchgrass. Late-flowering genotypes yield more biomass because the growing season is extended; having a better understanding of the genes that control flowering time will help to develop a rational strategy for creating improved switchgrass lines. The knowledge generated will aid breeding programs in developing late flowering varieties of switchgrass that fully utilize the growing season and achieve high biomass yield.

Accelerated Development of Optimal Pine Feedstocks for Bioenergy and Renewable Chemicals Using Genome-Wide Selection

Matias Kirst

Matias Kirst, University of Florida, Gainesville

Goal: To hyper-accelerate pine breeding using genome-wide selection, generating cultivars of loblolly and slash pine tailored to produce high energy yields that are ready for deployment. Traditional genetic improvement of pines is logistically complex and expensive, and a single breeding cycle takes almost two decades to complete. Thus, new breeding strategies that accelerate the development of cultivars that are suitable for bioenergy production will be developed and applied.

Pyramiding Genes and Alleles for Improving Energy Cane Biomass Yield
Ray Ming

PI: Ray Ming, University of Illinois at Urbana-Champaign, Urbana, IL 61801
Co-PI: Chifumi Nagai, Hawaii Agriculture Research Center, Kunia, HI 96759
Qingyi Yu, Texas A&M AgriLife Research, Weslaco, TX 78596

Summary

Energy cane and sugarcane cultivars are generally derived from interspecific hybridization between high sugar content and biomass yield *Saccharum officinarum* and wild species *S. spontaneum*. Commercial energy cane and sugarcane cultivars are subsequently developed through additional rounds of backcrossing to *S. officinarum* or hybrids to recover the high biomass yield and high sugar content while retaining biotic and abiotic stress resistance provided by *S. spontaneum*. This scheme has been practiced for a century by sugarcane breeders due to the need to recover high sugar content. We propose a new paradigm for energy cane breeding to utilize the transgressive segregation in true F2 populations from interspecific crosses, because sugar content is not a limiting factor for selecting high biomass energy cane cultivars. Our initial field trial of an F2 population yielded clones with 338% increase of biomass yield compared to its high yielding parent *S. officinarum* LA Purple. Such extraordinary yield performance is due to pyramiding genes/alleles for biomass yield in autopolyploid genome, an advantage of 8 potential alleles for each gene in autooctoploid. Our long term goal is to establish a new paradigm to accelerate energy cane breeding programs and maximize the biomass yield for biofuel production. The implementation of this project will uncover the genes and alleles contributing to the high biomass yield in the high end extreme segregants of the F2 population, including the mode of interaction among various dosages of the same haplotype. Understanding the mechanisms of the extraordinary transgressive segregation in autopolyploid sugarcane will accelerate the application of this new paradigm in energy cane breeding programs, and may have implication in crop improvement programs of other autopolyploid crops such as potato.

Global Analysis of Epigenetic Regulation of Gene Expression in Response to Drought Stress in Sorghum
A.S.N. Reddy

A.S.N. Reddy, Colorado State University, Fort Collins

Goal: To investigate the impact of drought stress on epigenetic modifications and alternative splicing in sorghum. Using recently developed high-throughput tools, this project will examine genome-wide changes in the chromatin landscape and patterns of alternative splicing in drought sensitive and tolerant cultivars under normal conditions and in response to drought stress. Understanding how plants respond and adapt to drought stress at the molecular level will help in developing plants that can grow under water- limiting conditions.

Structural Polymorphisms as Causes of Heterosis in Populus
Steven Strauss

Steven Strauss, Oregon State University, Corvallis

Goal: To characterize the extent of structural polymorphisms (SPs) between and within species of Populus that are used to produce wood and bioenergy, and examine their relationship to growth, stress tolerance, and breeding efficiency. This project will study wild black cottonwoods and interspecies hybrids important in plantations in the Pacific Northwest USA and other parts of the world, with a focus on the extent to which assay of SPs could improve hybrid breeding compared to alternative approaches.

PROJECT REPORTS (Alphabetical Order of Lead Project Director)

Association Mapping of Cell Wall Synthesis Regulatory Genes and Cell Wall Quality in Switchgrass

Laura Bartley

PI: L. Bartley, University of Oklahoma
Co-PI: Y. Wu, Oklahoma State University
Collaborators: L. Zhu, Oklahoma State University,
E. C. Brummer and M. Saha, Noble Foundation,
N. Santoro, Michigan State University

The research goals of this project are to test the association between polymorphisms in phenylpropanoid regulatory networks and biosynthetic genes with altered cell wall quality in switchgrass (*Panicum virgatum*) and to conduct reverse genetics of demonstrated or likely candidate genes in model grasses. This work aims to advance both control over the critical trait of biomass recalcitrance and genomic-breeding knowledge in the tetraploid native bioenergy grass, switchgrass. Our progress is described below.

Under typical conditions the efficiency of recovery of sugars from lignocellulose for biochemical conversion to biofuels is about 50%. Hence, improving cell wall deconstruction by genetically manipulating the structure and content of cell walls has potential to enormously improve the overall efficiency of biofuel production, provided such changes can be engineered without compromising plant fitness. Though grasses constitute the major source of biomass in the U.S. available for conversion to lignocellulose, our knowledge of cell wall synthesis and control has been based mostly on that gleaned from Arabidopsis and other dicots. Since dicots and grasses have different cell wall content and patterning, synthesis and possibly regulatory mechanisms may have diverged between the two clades. A further complication for translating data from models to biomass crops is the expansion of known cell wall related gene families in grasses, especially in species with complex genomes, such as switchgrass.

To address the question of conservation and divergence of grass cell wall regulatory mechanisms, we have mined three publicly available rice gene association networks, ROAD, PlaNet and RiceNet, and various combinations of them. From these networks, we extracted secondary cell wall (SCW) subnetworks around homologs of known cell wall-related transcription factors, lignin biosynthesis genes, and putative grass-diverged crosslinking enzymes. Of the various networks, the SCW subnetwork built using a general linear model (GLM) that combined the three networks showed the largest biological pathway Gene Ontology-term match ratio, indicating high network quality. The GLM-network is also enriched with non-bait cell wall metabolic terms. So far, we have tested eight predicted protein-promoter interactions predicted in the GLM-subnetwork between transcription factors and SCW biosynthesis pathway genes with a luciferase transient assay. Of these, six of them (62.5%) show increased luciferase production in the presence of the transcription factor relative to its absence. Genetic analysis of the roles of some of the novel, but validated transcription factors in the MYB-R2R3 and MYB-Like families is now underway in rice. As an orthoganol approach to experimental validation, we are also analyzing the patterns of promoter motifs in the rice gene

network. Furthermore, we have initiated a modest study of the rice stem developmental gradient proteome and phosphoproteome toward addressing the question of posttranslational modifications that may impact grass secondary wall development.

With the goals of translating the Arabidopsis and rice results to switchgrass and in supporting the annotation of the switchgrass genome, we have conducted a detailed phylogentic analysis of the annotated R2R3-MYBs in the switchgrass v0.0 and v1.1 Phytozome genomes, compared with those in the genomes of Arabidopsis, rice, maize, and poplar. From this we have learned that R2R3 MYBs involved in SCW development in Arabidopsis fall into three classes relative to those of the grasses: conserved, conserved and expanded/contracted, and non-conserved. Furthermore, The work reveals that the current switchgrass genome and transcriptome resources when combined are likely mostly complete, though have collapsed homoeologs and alleles into single loci. However, many of these issues have been resolved in the pre-released, v1.1 genome.

To translate knowledge of secondary wall synthesis to differing switchgrass genotypes, we have also conducted a study of gene expression and cell wall content correlations in three well-studies swtichgrass genotypes. Strikingly, this work suggests that different lignin biosynthesis genes and gene orthologs demonstrate predominant expression in different genotypes. We have expanded this work via RNA Seq analysis of four tissues (leaf, young stem, old stem, whole tiller) of four genotypes with divergent deconstructability that we are now in the process of analyzing.

All of the molecular analyses described above lay the ground-work for and are complementary to the switchgrass association genetics project that is underway. This work focuses on the Switchgrass Southern Association Collection which consists of ~16 genotypes from each of 36 diverse switchgrass accessions assembled by collaborators Brummer and Saha from across the Southern U.S., from Florida to New Mexico. By focusing on diverse switchgrass genotypes, we will obtain data about older recombination events that led to successes in diverse environments. Furthermore, the advantage of this collection is that it has been phenotypically characterized in detail in two different sites (Noble Foundation and Georgia) over 3-4 years and continuing.

Complementing the DOE-BESC project of my collaborators, we have obtained quantitative wet chemistry cell wall data for this collection as follows: hydroxycinnamates, ferulic acid and *p*-coumaric acid; acetyl bromide soluble lignin; and enzymatic deconstruction subsequent to mild pretreatment. Total sugar measurements and quantitation of hemicellulose sugar monomers are underway. These data will be used to identify quantitative trait loci for this traits through association analysis using a mixed linear model of Yu et al. 2006 (Nature Genet.) using single nucleotide polymorphism data obtain through reduced representational genotyping by sequencing. Simulation studies by collaborator Zhu with the mixed linear model, especially when a correction factor is included, reveal that this model has high power and specificity, as expected. Significant SNPs identified in the association collection will be tested in the regional switchgrass collection of Y. Wu at Oklahoma State University. We have found that this collection shows even greater cell wall content variation via near infrared spectroscopy than the Southern Association Collection.

Taken together, we expect that this project will significantly advance our understanding of secondary cell wall biosynthesis and regulatory networks in grasses, in general, including the subtleties of specific gene functions among diverse switchgrass. Toward actual improvements in production of switchgrass for biofuel production, already, this project has identified switchgrass genotypes with highly divergent cell wall qualities for use in breeding and we hope that associated markers will soon be forthcoming.

Functional Interactomics: Determining the Roles Played by Members of the Poplar Biomass Protein-Protein Interactome
Eric Beers

Project Director: Eric Beers, Virginia Tech, ebeers@vt.edu
Co-PDs: Amy Brunner, Virginia Tech, abrunner@vt.edu; Allan Dickerman, Virginia Tech, allan@vbi.vt.edu; Richard Helm, Virginia Tech, helmrf@vt.edu

Project website: http://xylome.vbi.vt.edu/index.html

Objectives and Accomplishments: Our main objective is to use yeast two-hybrid assays to discover novel networks of interacting proteins for more detailed analyses involving independent interaction assays and transgenic poplar. To date we have detected over 200 protein-protein interactions involving over 70 bait proteins. The following three interaction networks are being functionally characterized. **1)** XND1 (a NAC domain protein and a strong negative regulator of xylem differentiation) and its orthologs from several eudicots and monocots are RETINOBLASTOMA-RELATED (RBR)-binding proteins. XND1 also interacts with VND6 and VNI2, two other NAC proteins that regulate xylem differentiation. We are investigating the potential for RBR to negatively regulate xylem differentiation through interactions with NAC proteins. **2)** PtrDRIF is composed of two distinct domains: an N-terminal Myb–like domain, which interacts with Myb proteins PtrRAD and PtrDIV, and the C-terminal DUF3755-containing domain, which interacts with the xylem-associated homeodomain proteins PtrHB1.3 and PtrKNAT7.1. Results indicate that PtrDRIF serves as a scaffold for Myb-Homeodomain protein complexes in poplar wood-forming tissue. **3)** PB504, a novel endomembrane protein of unknown function, was found to interact with cellulose synthase. We have extended the PB504 network, which now includes a small GTPase, a Ser/Thr protein kinase and several uncharacterized putative cytoskeletal proteins. We are currently characterizing constitutively active and dominant negative forms of the small GTPase and Ser/Thr protein kinase in transgenic plants. In November 2013, we planted over 500 transgenic poplar trees for field studies. We are also using *PtrNAC154* overexpression lines to test the hypothesis that genetic modification of poplar trees results in chemotypes that can be assessed by global metabolite profiling.

Broad Impacts: We are focusing on previously uncharacterized pathways for transcriptional regulation, signaling, and cytoskeletal activities to advance our knowledge of wood formation. Our discoveries related to the XND1-RBR interaction have led us to a novel testable hypothesis with broad implications for the evolution of vascular plants.

Deliverables

Publications (2011-2013):

- Zhao, C., E.P. Beers, 2013. Alternative splicing of Myb-related genes *MYR1* and *MYR2* may modulate activities through changes in dimerization, localization, or protein folding. Plant Signal. & Behav. 8: http://dx.doi.org/10.4161/psb.27325
- Petzold, H.E., M. Zhao, E.P. Beers, 2012. Expression and functions of proteases in vascular tissues. Physiol. Plant. 145:121-129
- Rodgers-Melnick E., S.P. Mane, P. Dharmawardhana, G.T. Slavov, O.R. Crasta, S.H. Strauss, A.M. Brunner, S.P. DiFazio, 2012. Contrasting patterns of evolution following

whole genome versus tandem duplication events in *Populus*. Genome Res. 22:95-105

- van Doorn, W.G., E.P. Beers, J.L. Dangl, et al. (19 authors), 2011. Morphological classification of plant cell deaths. Cell Death Differ. 18:1241-1246
- Zhao, C., A. Hanada, S. Yamaguchi, Y. Kamiya, E.P. Beers, 2011. The Arabidopsis Myb genes *MYR1* and *MYR2* are redundant negative regulators of flowering time under decreased light intensity. Plant J. 66:502-515

Oral/ Poster Presentations (for 2013):

Posters presented at IUFRO, Tree Biotechnology, May 26-June 1, 2013, Asheville, N.C.

- Rigoulot, S., E. Beers, A. Brunner, A. Dickerman, X. Sheng, C. Zhao, B. Chanda, E. Petzold, X. Jia, 2013. Poplar biomass interactome project reveals novel protein-protein interaction modules.
- Petzold, E., C. Zhao, A. Brunner, E. Beers, 2013. The poplar RWI protein binds to MYB domain and homeodomain proteins through its distinct MYB-like N-terminal and DUF3755 C-terminal domains.
- Zhao, C., L. Bakó, A. Cruz-Ramírez, B. Scheres, A. Brunner, E. Beers, 2013. The ability of XYLEM NAC DOMAIN 1 to block the terminal differentiation of xylem vessel elements depends on its RETINOBLASTOMA RELATED interaction domain.

Poster presented at the annual meeting of the American Society of Plant Biologists, July 20-24, 2013, Providence, RI.

- Chanda, B., E. Petzold, X. Jia, A. Dickerman, A. Brunner, E. Beers, 2013. The interactome network of a novel endomembrane protein in *Populus trichocarpa.*

Community Resources Generated: The project website (http://xylome.vbi.vt.edu/index.html) provides information on available clones and interactions discovered.

Other products/ outcomes:

Training: *Postdoctoral*; Bidisha Chanda (Dr. Chanda was awarded a travel grant from the American Society of Plant Biologists to attend the annual meeting in Providence, RI, July 20-24, 2013, and present her interactome poster.)

PhD students; Xiaoyan Jia, H. Earl Petzold, Steven Rigoulot

Undergraduates trained in poplar tissue culture and transgenesis; Grace Lee, Rosa Angeles, Chihyun Hwang, Nina Wilson

Collaborations: Collaborating with Pankaj Jaiswal to expand the poplar protein-protein interactome.

Functional Genomics of Sugar Content in Sweet Sorghum Stems
David Braun

Project Director: David Braun, University of Missouri, Columbia, braundm@missouri.edu
Co-PDs: Ismail Dweikat, University of Nebraska, Lincoln, idweikat2@unl.edu
Benjamin Babst, Brookhaven National Laboratory, bbabst@bnl.gov
Rich Ferrieri, Brookhaven National Laboratory, rferrieri@bnl.gov

Project website: http://agronomy.unl.edu/sweetsorghum

Objectives and Accomplishments:
Understanding the mechanisms regulating the transport and partitioning of sucrose to sweet sorghum stem tissues will lead to new approaches for modifying and enhancing this crop for bioenergy production. Our project seeks to determine the molecular basis of sucrose hyperaccumulation in sweet sorghum stem tissues through a combination of genetic, genomic, biochemical, and physiological approaches. Objective 1 will map and characterize quantitative trait loci (QTL) related to stem biomass, total biomass, and sugar accumulation in the stem in a population derived from a grain sorghum crossed by a sweet sorghum. Objective 2 will determine the expression patterns and functions of a family of proteins, called sucrose transporters (SUTs), which are hypothesized to play essential roles in regulating carbon partitioning of sugar to sweet sorghum stems. Objective 3 will use detailed radiotracer physiological and biochemical approaches to map and delineate the bottlenecks to sucrose allocation in sweet sorghum. From these data, we will generate a physiological knowledge base and determine which genes function in transport and allocation of assimilated carbon to stem tissues in sweet sorghum. The knowledge and germplasm generated by the project will make important contributions to advancing sweet sorghum for sustainable bioenergy production.

Objective 1: The sweet x grain sorghum F_6 recombinant inbred line (RIL) mapping population and parental lines were planted at two locations in Nebraska and phenotypically characterized during the last two growing seasons. Out of 1184 SSR markers screened with parental lines (Wray and Macia) of the RIL population, 177 markers were polymorphic (codominant markers) and are being placed on the mapping population. Additionally, we are in the process of genotyping by sequencing the RILs.
Objective 2: To characterize SUT gene expression, we are performing quantitative RT-PCR comparing the two lines at different developmental stages and across multiple tissues. Tissues sampled include immature and mature leaves, stem internodes at different developmental stages (adult vegetative stage, pre-flowering, post-flowering), and panicles. Detailed SUT expression studies are in progress to assess which genes display differences in expression between sweet and grain sorghum in the different tissues and during plant development. These experiments will inform our understanding of which SUTs may contribute to the high sucrose accumulation in sweet sorghum stems.
Objective 3: The carbon-11 (^{11}C) radiotracer experiments are being led by Co-PDs Dr. Ben Babst and Dr. Rich Ferrieri at Brookhaven National Laboratory. They have hired a postdoctoral researcher to work on the project who is responsible for the radiotracer transport assays, imaging, and metabolomics flux experiments. Using $^{11}CO_2$ to pulse label leaves of sweet and grain sorghum plants, preliminary studies suggest that sorghum plants are not source limited in terms

of carbohydrate partitioning as there was no significant differences between the genotypes. However, there were strong differences in allocation of recently fixed carbon to sink tissues, reflecting the difference in carbon storage as sucrose in stems of sweet sorghum vs. starch in seeds of grain sorghum.

Deliverables:
Publications:
David M. Braun, Lu Wang, and Yong-Ling Ruan, (2013) "Understanding and manipulating sucrose phloem loading, unloading, metabolism, and signaling to enhance crop yield and food security" *Journal of Experimental Botany*, published online, doi:10.1093/jxb/ert416.
Benjamin A. Babst, Abhijit A. Karve, and Tatjana Judt, (2013) "Radio-metabolite analysis of carbon-11 biochemical partitioning to non-structural carbohydrates for integrated metabolism and transport studies" *Plant and Cell Physiology*, 54: 1016-1025.
Saadia Bihmidine, Charles T. Hunter III, Christine E. Johns, Karen E. Koch, and David M. Braun, (2013) "Regulation of assimilate import into sink organs: Update on molecular drivers of sink strength" *Frontiers in Plant Science, 4:177. doi: 10.3389/fpls.2013.00177*
David M. Braun (2012) "SWEET! The pathway is complete" *Science*, 335: 173-174
R. Frank Baker, Kristen A. Leach, and David M. Braun, (2012) "SWEET as sugar: New sucrose effluxers in plants" *Molecular Plant*, 5: 766-768

Training:
Cassandra Hoffner, undergraduate student in Braun lab
Saadia Bihmidine, postdoctoral fellow in Braun lab
Kanokwan Teingtham, graduate student in Dweikat lab
Abhijit Karve, postdoctoral fellow in Babst lab

Collaborations:
Collaboration with Cliff Weil at Purdue University
Collaboration with Seth DeBolt at University of Kentucky

Functional Discovery and Characterization of Genes and Alleles Affecting Wood Biomass Yield and Quality in Poplar Using Activation Tagging and Association Analysis

Victor Busov

PI - Victor Busov
Co-PIs: **Yordanov, Y** - MTU, Postdoctoral scientist, **Sykes, R –** National Renewable Energy Laboratory, Research Scientist, **Muchero, W -** Oak Ridge National Laboratory, Staff Scientist, **Tuskan, G -** Oak Ridge National Laboratory, Distinguished Scientist

PI's contact information:
Michigan Technological University
1400 Townsend Dr.
Houghton, MI 49931
Phone: 906-487-1728
Fax: 906-487-2915

Email: vbusov@mtu.edu

The goal of this project is to discover and characterize novel genes and alleles that affect wood biomass yield and quality in *Populus*. These discoveries can enable knowledge-based approaches for development of specialized bioenergy poplar cultivars. We employ a two stage approach, using activation tagging as a functional gene discovery tool followed by intensive search and validation of functional alleles that can be deployed through traditional breeding or genetic engineering. Our approach is combining the strengths of mutagenesis for functional identification of genes with the power of next generation sequencing technologies for identification of alleles with breeding values.

To date we have screened 1,100 activation tagged poplars. A total of 133 biomass yield and 63 biomass quality mutants were discovered. Biomass yield mutations affect stem diameter, height, wood dry weight and density. In addition to effect on one of these traits, some of the mutations impact simultaneously as many as 5-6 traits. A total of 63 mutations affecting cell wall composition were discovered including 38 with altered lignin concentration and S/G ratio and 25 with changes in C5 and C6 sugar content. We have positioned and validated the activation in more than 40 mutant lines which were prioritized by the severity and importance of the phenotype. For 6 tagged activated genes we have initiated assembly of constructs and transformation experiments to recapitulate the phenotype. All 40 genes that were molecularly characterized were sent to the scientists in the DOE OakRidge National Laboratory for association studies. These studies will identify native alleles affecting the studied trait(s).

Identifying Differences in Abiotic Stress Gene Networks between Lowland and Upland Ecotypes of Switchgrass (DE-SC0008338)

Kevin Childs

PI - Kevin Childs, Michigan State University
Co-PIs: C. Robin Buell, Michigan State University; Bingyu Zhao, Virginia Tech; Xunzhong Zhang, Virginia Tech

Switchgrass (*Panicum virgatum*) is a North American prairie grassland species that has been identified as a desirable lignocellulosic biofuel species for use in the United States due to its local adaption and high biomass accumulation. It is envisioned that switchgrass will be grown, in part, on marginal lands to supplement and diversify farmers' traditional crop incomes. In the last twenty years, interest in using switchgrass as a biofuel feedstock has lead to more intensive efforts to improve switchgrass yields. Despite impressive productivity gains, additional yield increases should be possible. One area for improvement will be the tolerance of switchgrass to abiotic stresses such as drought and salinity. Not only is little known about the genes in switchgrass that affect tolerance and sensitivity to these two abiotic stresses, but only a relative handful of switchgrass cultivars have been phenotyped under drought and salt stress.
For this project, we are screening 49 switchgrass cultivars/varieties from both the lowland and upland ecotypes for drought and salt tolerance. The initial experiments have been performed in a greenhouse at Virginia Tech. Clones of established cultivars/varieties were transplanted to pots

and grown for seven weeks before treatments began, and plants were subjected to drought/salt treatments for 30 days. For drought treatments, soil moisture was decreased from ~35% to ~5% over 30 days. Control treatments for the drought experiment consisted of plants that were watered to full soil moisture. For the salt stress experiment, plants were provided either half-strength nutrient solution or nutrient solution with NaCl. For the salt treatment, NaCl was initially provided at 100 mM, but it was increased to 250 mM after the first week. Ca^{+2} was provided in the nutrient solution.

A wide variety of morphological and physiological data have been collected every 5 or 6 days during the experiments, and tissues were harvested at the end of the experiments for metabolite analyses. Morphological measurements included plant height, tiller number, leaf blade length, leaf blade width and sheath length. Physiological analyses included leaf relative water content, electrolyte leakage, net photosynthetic rate, stomatal conductance, water use efficiency and transpiration rate. Metabolites related to abiotic stress responses in plants were also analyzed by GC-MS and LC-MS/MS. Metabolites that were quantified included phytohormones (ABA, JA, JA-Ile), proline, polyamines (putrescine, betaine, spermidine, spermine) and simple sugars (glucose, fructose, inositol, sucrose, trehalose and raffinose). Additionally, for the salt stress treatments, leaf Na^{+} and K^{+} content was measured at the end of the experiment.

The morphological and physiological responses of the 49 cultivars/varieties to drought and salt stress have been quite varied. In reaction to both drought and salt treatments, leaf electrolyte leakage increased, photosynthetic rate decreased, stomatal conductance was reduced and transpiration rate declined in all varieties. Nonetheless, variation in responses was seen among the cultivars/lines that we examined. We calculated salt and drought tolerance indices relative to photosynthetic rates (Pn) to determine relative stress tolerance of each of the 49 cultivars/lines. Stress tolerance indices were calculated as $Pn_{stress}/Pn_{control}$. The results showed that lowland ecotypes were generally more tolerant to both salt and drought stresses. However, two notably tolerant cultivars included Caddo, an upland, and T-2086, a lowland. The upland cultivars, Dacotah, Cave-in-Rock, and Genville-2, were more sensitive to salt and drought stresses than all other cultivars/lines that were tested. Metabolite analyses are ongoing with these samples, but with ABA for which we have the most complete data, we have observed large variations in concentrations suggesting that differences in ABA production in response to drought may be playing a role in the differences in physiological responses that we have recorded. We will make a final decision in the near future about the tolerant/sensitive varieties that will be transplanted and cultured for the next phase of this work.

From these initial results, we will choose four cultivars that are notably tolerant/susceptible to salt and drought stresses. The candidate cultivars/varieties will be subjected to appropriate drought or salt stress treatments again and morphological and physiological measurements will be taken at regular intervals following the onset of stress treatments. Additionally, tissue samples will be collected for metabolomic and RNA-seq analyses at the same time points. We will use Weighted Gene Coexpression Analysis (WGCNA) to identify modules of highly correlated genes. WGCNA will also be used to attempt to correlate morphological, physiological and metabolic measurements with gene expression. Our goal will be to identify groups of genes belonging to common genetic pathways but to also find groups of genes that are involved with specific physiological and metabolic responses of switchgrass to drought and salt stress.

In addition to our phenotypic analyses of morphological, physiological, metabolic and gene expression within specific cultivars of switchgrass, we plan to target several transcription factors

known to be involved in abiotic stress responses. We will overexpress these genes in switchgrass and subject the resulting transgenics to drought or salt stress as described above. One line that we have already tested overexpresses the *Arabidopisis thaliana LOV1* gene. Switchgrass that overexpress *AtLOV1* are more tolerant to drought than non-transgenic plants as measured by relative water content, electrolyte leakage and more resilient CO_2 assimilation. Of the metabolite analyses performed on these plants, proline was found to increase in control plants in response to drought, but proline increased to a significantly lesser degree in the *AtLOV1* overexpressing plants under drought conditions. Conversely, raffinose levels were significantly higher in drought stressed *AtLOV1* switchgrass than in non-transgenic control plants. There were changes in many of the other metabolites that we analyzed in response to drought (ABA, JA, JA-Ile, putrescine, inositol, fructose, glucose, trehalose), but the observed levels of these compounds were not notably different between the transgenic and non-transgenic plants.

Creation and High-Precision Characterization of Novel Poplar Biomass Germplasm

Luca Comai

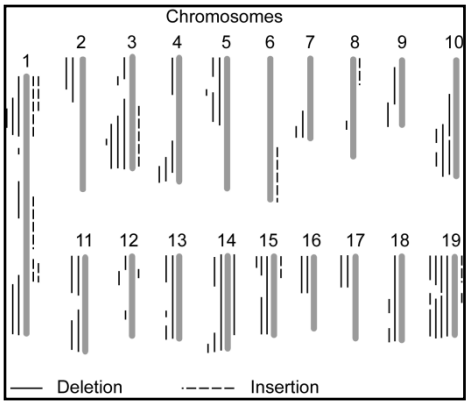
Project Director: Luca Comai, UC Davis Genome Center and Dept. Plant Biology, lcomai@ucdavis.edu
Co-PDs (Name, institution, email): Andrew T. Groover, Institute of Forest Genetics, US. Forest Service, agroover@fs.fed.us

Project website (URL): http://comailab.genomecenter.ucdavis.edu/index.php/Poplar

Objectives and Accomplishments:
Previous research has indicated that hybrid poplar germplasm exhibited frequent variations in genome content. Triploids, aneuploids, and related dosage variants are particularly suited for poplar breeding because they can be maintained in vegetatively propagated system and they constitute a rapid alternative resource for functional genomics and gene discovery.

Objectives

1. Survey genomic composition in commercial pedigrees of Populus F1 hybrids.
Of 326 samples, including five different hybrid types, 8 triploids (2.5%) and 13 aneuploids were identified (4%). Aneuploidy originated from the male gamete in 12/13 samples, suggesting that aneuploid gametes are at least as likely to successfully compete against balanced haploid pollen grains. We also characterized the progeny of pure species crosses. In a population of *P. trichocarpa*, 3/53 trees exhibited aneuploidy (5.7%). This confirmed the high rate of aneuploidy, even in a situation where no selection was performed on the germinated seeds. Similarly, we found 2/428 aneuploid in collection of pure *P. trichocarpa* from natural population. *Taken together, our results suggest that aneuploid and triploid poplar trees are successful both in nature and in breeding programs, consistent with the idea that they might exhibit advantageous traits.*



2. Manipulate gametic contribution for functional genomic and germplasm enhancement
To create more diverse germplasm, we produced a population of dosage variants via gamma-irradiation of mature pollen. Samples originating from pollen irradiated with 100 Gy were selected for an initial dosage analysis, from which 41/74 (55%) exhibited at least one dosage lesion. Lesions vary in length (up to a whole chromosome) and there were up to six lesions in a single individual, included deletions (82%), and insertions (18%). With just 74 individuals, 68% of the genome is represented in at least one dosage lesion (see Figure, depicting the position and length of each indel on the 19 chromosomes of poplar). A second round of pollen irradiation produced > 500 new irradiated plants now under analysis. *As anticipated, this population provides an excellent tool to investigate the effect dosage variation of specific chromosomal pieces or specific genes. Our goal is to produce and characterize enough dosage variants to cover each gene at least once, preferable several times.*

3. Correlate variation in karyotype, gene dosage, and transcriptional modules with superior biomass production
In this objective, we aim to use the material presented above to assess the effect of altered gene dosage on phenotypes, focusing on those phenotypes associated with biomass production. The trees generated in year 2012 are currently in their second year and are ready for propagation this winter. We have performed an initial phenotypic assessment of the dosage variants with the following conclusions: Many of the aneuploids lacked an obvious phenotype and grow well, sometimes better than their diploid siblings. The aneuploids display the expected wide variation in trait values, for example in phenology, leaf shape or apical dominance. We have produced RNA-Seq libraries from a subset of individuals and are awaiting sequencing data.

Broad Impacts:
We have demonstrated that both triploidy and aneuploidy were common in interspecific crosses, confirming that aneuploidy *per se* is not detrimental in poplar, and can lead to novel phenotypes. Between the "natural" dosage variants that we have identified in breeding programs and those we have produced through gamma mutagenesis, we have assembled a unique collection of variants that will be instrumental in investigating the relationship between phenotype and gene dosage, and represents a new pool of phenotypic variants for poplar.

Deliverables

Publications:

Henry IM, Groover AT, Comai L. Creation and characterization of dosage variants for functional genomics in Populus, in preparation

Groover A, and Cronk Q (2013). From Nehemiah Grew to Genomics: the emerging field of evo-devo research for woody plants. International Journal of Plant Science 174(7), 959-963.

Oral/ Poster Presentations:

Eight presentations in the past year by Andrew Groover (5), L. Comai (2), I. M. Henry (1) including an invited talk at PAG2013

Community Resources Generated

- All sequencing reads will be deposited in the NCBI SRA database
- Lists of single nucleotide polymorphisms (SNPs) between the various populus species characterized will also be made available in the context of this manuscript as well as on our

website. The following lists are currently available: *P. deltoides* x *P. nigra* (155,689 SNPs), *P. deltoides* x *P. trichocarpa* (105,525 SNPs), *P. deltoides* x *P. maximowiczii* (53,718 SNPs) and *P. maximowiczii* x *P. nigra* (39,310 SNPs).
- Data about our tree collection will be available in the Populus Genome Integrative Browser (PopGeniE.org). Specifically, the list of dosage variants, along with information about their origin, the nature of the dosage lesions and which genes are located within the indels is being developed as a new resource within the PopGeniE web portal, called "PopIndels". The resource is currently being developed, in collaboration with the PopGenIE team. A demo version is available at the following url: *http://popindels.popgenie.org/* (under construction).
- The trees exhibiting dosage variation are currently being propagated for public access**.**

Training

Matthew Zinkgraf: Post-doctoral researcher; Jian Gao: Visiting graduate student from China, Elisha Garcia, Dale Hill, Danae Sugaoka and Peter Deng: Undergraduate researchers at UCD. Shani Sarena, Courtney Castle, Txai Gomez: Undergraduate student interns at USFS. Jenna Arciero: High School intern at USFS.:

Collaborations

Greenwood Resources: Poplar breeding company.University of Minnesota Natural Resources Institute. Cees Van Oosten, poplar breeding community. Nathaniel Street and the PopGenie team from the Umeå Plant Science Centre, Umeå, Sweden Charles Hefer, Shawn Manfields and Carl Douglas, UBC – Dosage characterization of *P. trichocarpa* pure species pedigree.

Poplar Interactome for Bioenergy Research
Pankaj Jaiswal

Project Director: Pankaj Jaiswal, Oregon State University, jaiswalp@science.oregonstate.edu
Co-PDs: Palitha Dharmawardhana, Oregon State University, dharmawd@science.oregonstate.edu, Amy Brunner, Virginia Tech, abrunner@vt.edu and Eric Beers, Virginia Tech, ebeers@vt.edu.
Project website: N/A
Objectives and Accomplishments:
Aim #1: Poplar functional interaction network: A network for poplar was constructed based on AraNet (http://www.functionalnet.org/aranet/) and Inparanoid orthology projections. We have enriched this network with experimentally validated data coming from the Arabidopsis interactome project (http://interactome.dfci.harvard.edu/A_thaliana/) and poplar xylem interactions generated through the poplar biomass interactome project (http://xylome.vbi.vt.edu/). The enriched poplar functional interactome included 17,467 unique poplar genes and 1.9 million interactions.

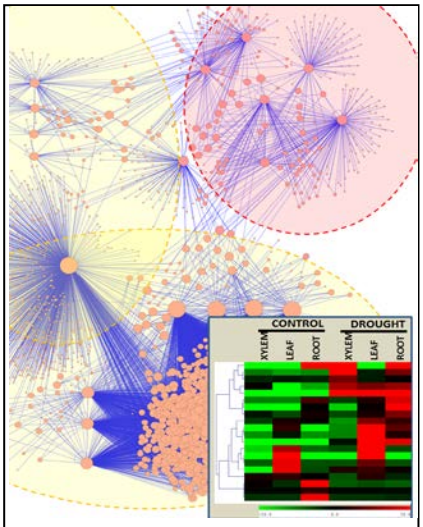


Figure-1: Coexpression network showing nodes (genes) discovered as interaction hubs. Inset shows gene expression profile of the genes represented as interaction hubs.

Aim#2 and #3: Develop subnetworks for abiotic stress:
In order to add confidence on interacting edges and nodes in the largely computationally projected network (Aim#1) we carried out an extensive tissue specific genome-scale transcriptome profiling study under 4 abiotic stress conditions (drought, salt, heat and cold) for the reference species *Populus trichocarpa* Nisqually-1. The transcriptome data is being analyzed for developing coexpression networks. This experiment also helped us to establish baseline drought and salt treatment conditions to be used for the planned transcriptome and expression network studies on non-reference drought and salt resistant genotypes of Poplar. RNA sampling from all treatments on the reference Nisquali-1 have been processed and RNA-Seq analysis has been completed for tissues from the drought experiment. A subset of data from this experiment was used to analyze the drought response subnetworks extracted from the poplar interactome (aim#1). Preliminary results of the ABA biosynthesis and signaling interaction network and the expression pattern of the major interacting hub genes (more than 10 interactions) were analyzed. The core 31 genes from the ABA biosynthesis and signaling processes show 7,767 interactions involving additional 2666 unique genes but connected to 20 major interaction hubs (genes). These hubs included the key ABA receptor & signaling components (yellow highlight) and biosynthetic enzymes (red highlight) (Figure-1). After a baseline analysis, we are in the process of conducting a similar transcriptome and genotyping analysis on the five selected genotypes, which show distinct shoot biomass gain differences under drought (Figure-2).

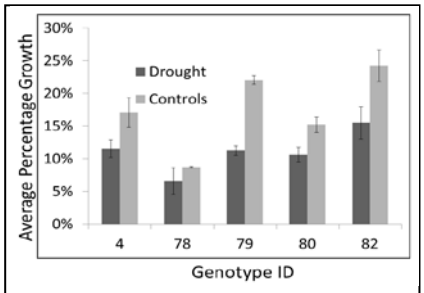


Figure-2: Percentage above ground biomass increase in the 5 poplar genotypes under drought (2 weeks) and control conditions.

Aim#4 Y2H screening of a set of candidate genes associated with abiotic stress stress response:
A graduate student was recruited and began to work on library construction for this project in fall 2013. Two libraries were constructed using the CloneMiner II cDNA Library construction kit. The first library was constructed using mRNA isolated from *Populus trichocarpa* axillary buds collected on October 17. Starting with approximately 3 µg of mRNA, an entry library with a titer of 5.4×10^6 cfu/ml or 6.48×10^7 primary clones was constructed. Inserts from 15 randomly selected clones ranged from 300 – 1100 base pairs. The second library using pooled RNA from salt- and drought-treated leaves was constructed using 1.5 µg of starting mRNA. This generated an entry library with a titer of 7.56×10^6 cfu/ml or 9.08×10^7 primary clones. Inserts from 15 randomly-selected clones ranged from 500 – 1200 base pairs. These two cDNA libraries are currently being used for preparation of AD fusion prey libraries for yeast two-hybrid assays The preparation of the cDNA library from the remaining RNA isolated from salt- and drought-treated roots will be completed after the AD fusion prey libraries are prepared and evaluated. Once the libraries are ready we will start screening by including the candidate genes such as the ones discovered as interaction hubs in Aim#2.

Aim#5 A web-based resource for Poplar Interactome: We are working with the DOE KBase Networks group by sharing annotation and expression profiles of publically available Poplar gene expression datasets for building the prototypes for the community resource.

Broad Impacts : The Poplar interactome with the verified subnetworks and of the drought and salt responsive co-regulated candidate genes will provide both a genomics resource for investigating effect of abiotic stress on bioenergy traits in poplar.

Deliverables

Publications: Coming soon.

Oral/ Poster Presentations:

1. Plant Genomes and Biotechnology meeting. CSHL, Dec. 2013 -Abiotic stress gene networks in poplar
2. IUFRO Tree Biotechnology Meeting. Asheville, May 2013 – Drought response gene networks in poplar
3. PAG 2013, San Diego, Jan 2013 - Construction of a functional gene network for poplar

Community Resources Generated: Coming soon.

Other products/ outcomes:

Training:
At Oregon State University, Postdoctoral Associates Palitha Dharmawardhana, Vindhya Amarasinghe, Sergei Filichkin and Justin Elser were involved in experimental setup, sample collection, sequencing and data analysis using bioinformatics. Undergraduate students Dylan Beorchia, Teague Green and Alexander Roos were trained on plant propagation and maintenance in green house conditions and plant physiology data collection. At Virginia Tech, Stephen Rigoulot a graduate student is trained in the yeast two-hybrid library construction and screening.

Collaborations :
We collaborate with the DOE KBase Plants and Network groups lead by Doreen Ware and Sergei Maslov.

The Genetics of Biofuel Traits in *Panicum* grasses: Developing a Model System with Diploid *Panicum hallii*

Thomas E. Juenger

Project Director: Dr. Thomas E. Juenger, University of Texas at Austin, tjuenger@austin.utexas.edu

Co-PDs: Dr. Ed Wolfrum, National Renewable Energy Lab, Ed.Wolfrum@nrel.gov

Project website: http://w3.biosci.utexas.edu/juenger_lab/

Objectives and Accomplishments:
The overall goal of the project research is to utilize genetic and genomic analyses to better understand the growth and development of *Panicum* grasses. Our objectives are centered on: (1)

characterizing the phenotypic diversity in *Panicum virgatum* and *Panicum hallii* accessions to better understand growth architecture in natural conditions; (2) mapping QTL for tillering, biomass, and tissue composition (NIR) in new mapping populations; (3) exploring the patterns of gene expression and transcriptional control of tillering in *Panicum* grasses under harvest and nutrient conditions; (4) developing bioinformatic resources furthering the use of diploid *Panicum hallii* as a models system for biofuels research.

We have established extensive *P. virgatum* and *P. hallii* common garden sites in Austin Texas. The existing garden includes a number of well studies agronomic cultivars (Alamo, Kanlow, VS16) as well as many new collection made by the Juenger lab. Phenotypic information from these lines has been collected over the course of our grant. We have developed a new 4-way genetic mapping population for *P. virgatum* that will segregate both upland and lowland traits. This population consists of 400 F_2 progeny from reciprocal crosses of F_1 parents. F_1 parents were derived from 4 grandparental lines including Alamo AP13 (lowland, JGI sequencing reference), Summer/VS16 (upland), WBC (lowland), and Dakotah (upland). These genotypes are currently being resequenced by the DOE Joint Genome Institute. We have collected basic growth and morphological measurements from these 400 progeny and have screened the population using a double-digest RAD protocol for genotyping. These data are currently being analyzed to develop a new linkage map and to complete preliminary QTL analyses.

We have completed QTL mapping with a *P. hallii* F_2 population and have made progress developing of a new recombinant inbred (RIL) mapping population. Our mapping studies have identified a number of QTL controlling flowering time, tillering, tiller characteristics, seed size, flower number, and sterility. We have progressed our RIL lines to F_5, plan to genotype the population at F_6, and bulk lines at F_7. An important thrust of the proposed work is to develop near-infrared spectroscopy methods for characterizing tissue in *Panicum* grasses. We have completed our initial NIRs run and have prepared samples from our F_2 QTL mapping population.

A key component of the project is to study the plasticity of growth architecture of *P. virgatum* and *P. hallii* grown under differing nitrogen and harvest environments. This project will focus on clonal replicates of the parents of the 4-way cross (AP13, VS16, WBC, Dakotah) as well as single seed decent lines of the two core *P. hallii* genotypes (FIL2 and HAL2). The experiment has been established, treatments applied, and phenotypic measurements and RNA-sequencing studies will be initiated during the 2014 field season.

Finally, we have worked closely with the DOE Joint Genome Institute and partners at HudsonAlpha Institute for Biotechnology to develop genomic tools for *P. hallii*. To this end, we continue to develop protocols, share materials, and facilitate ongoing sequencing, assembly, and annotation efforts.

Broad Impacts:
Our project will provide basic information on the growth architecture, development, and tissue characteristics of *Panicum* grasses. Moreover, genetic mapping and transcriptome analyses will help to identify genomic regions and genetic networks driving this natural variation. Ultimately, the development of an assembled *P. hallii* genome will facilitate functional genomic studies of C4 perennial grasses as biofuel feedstocks.

Deliverables
Publications: two manuscripts are in review, several manuscripts are in preparation

Oral/ Poster Presentations: Seminars presented at: Texas Tech University, Zurich Plant Sciences Institute Switzerland, University of Bern Switzerland, Switchgrass II, UC Davis, Oklahoma State University, Noble Foundation, University of Missouri Columbia, HudsonAlpha Institute of Biotechnology. Posters presented at: PAG 2012/2013, Switchgrass II meeting

Community Resources Generated: RNA-sequencing, genetic mapping, and genome assembly resources will be made available to the community. A preliminary version of the *P. hallii* genome assembly (PH version 0.5) is planned as an early release genome for Phytozome in January 2014. The *P. hallii* RIL mapping population will be made available through the GRIN stock center when completed.

Training:
Postdoctoral Researchers: David Lowry, Eugene Shakirov, Kyle Hernandez, Dhivya Arasappan, Benjamin Getz, Chris Blazier
Lab technicians: Tierney Logan, Jacob Heiling
Field technicians: Jason Bonnette, Briana Whitaker, Jacob Heiling
Graduate Students: Liz Milano, Juan Diego Palacio
Undergraduate Researcher: Wes Skillern,

Collaborations:
Zeng-Yu Wang (Noble Foundation)
Jeremy Schmutz (HudsonAlpha)
Jerry Jenkins (HudsonAlpha)
Dan Rohksar (DOE Joint Genome Institute

Genomic and Breeding Foundations for Bioenergy Sorghum Hybrids
Stephen Kresovich

Project Director: Stephen Kresovich, Clemson University, skresov@clemson.edu
Co-PDs: Andrew H. Paterson, University of Georgia
F. Alex Feltus, Clemson University

Objectives and Accomplishments:
The specific objectives of the proposed research are: (1) to develop ten nested association mapping populations (NAMs) and a diversity panel necessary to dissect the genetic bases of carbon accumulation and partitioning of cellulosic sorghum; (2) to phenotype these NAMs and diversity panel for patterns of carbon accumulation and partitioning and to correlate these traits with DNA sequence variation that will underlie future breeding/genetic studies; (3) to lay the foundation for integrating genomic selection and other genomics-based strategies into cellulosic sorghum breeding programs; and (4) to identify and create and characterize cellulosic male-sterile (A lines), maintainer (B lines), and restorer (R lines) germplasm necessary to exploit heterosis specifically targeted at energy production.

Following winter nursery activities in Mexico for 2013-2014, we will have made significant progress in population development. For each of the five races of sorghum, we have at least two NAM populations advancing to the F_3 or F_4 generation. In all, 15 NAM populations have been generated. We are a generation behind schedule because of the complexities of handling and inducing flowering in photoperiod sensitive lines. Particularly in the summer increase in South Carolina, plants are exceedingly tall (and difficult to pollinate and collect viable seeds) and frequently the seeds do not mature. When this project ends, we intend to continue advancing populations through support to be provided by the national sorghum commodity group.

Three F_3 populations have been phenotyped for a significant number of agronomic and compositional traits in replicated trials in South Carolina and Georgia. These populations then will be genotyped and association studies will be performed for particular bioenergy traits. It is expected that when all populations are completed, all individuals will be genotyped-by-sequencing.

A-lines of six selected bioenergy and sweet sorghum accessions have been developed. Seed of these lines (at the F_4 stage) will be made available on a limited basis.

In addition, we have conducted a complementary series of association analyses based on various diversity panels representing bioenergy, grain, and mineral phenotypes.

Broad Impacts:
Interest in both genetic information and germplasm are high for these bioenergy materials (both cellulosic and sweet types). Collaborations with private and public organizations have been established over the past two years. In addition, in areas of rural economic distress along the I-95 corridor, there is public interest in sorghum to reinvigorate the agricultural sector (see http://www.thestate.com/2013/10/19/3048058/sc-nurtures-crops-to-help-state.html) and provide raw materials for the production of energy, food or feed.

Deliverables:
Publications:
G.P. Morris et al. 2013. “Dissecting genome-wide association signals for loss-of-function phenotypes in sorghum flavonoid pigmentation traits”. G3: Genes, Genomes, Genetics: 11 2085-2094.
G.P. Morris et al. 2013. “Population genomic and genome-wide association studies of agroclimatic traits in sorghum”. Proceedings from the National Academy of Sciences: 110 453-458.
N. Shakoor et al. Accepted. “A Sorghum bicolor expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums”. BMC Plant Biology.

Oral/Poster Presentations:
Geoff Morris. "Dissecting genome-wide association signals: lessons from flavonoid pigmentation traits in sorghum". *Crop Science Society of America Annual Meeting* -Tampa, FL. November 2013.

Geoff Morris. "New resources and strategies for genome-wide mapping in sorghum". *Sorghum Improvement Conference of North America*- Lubbock, TX. August 2013.
Zach Brenton. “Pericarp pigmentation and its effect on grain quality in sorghum”. *University of South Carolina Magellan Scholars Discovery Day*-Columbia, SC. April 2013.
Geoff Morris. "Population genomic and genome-wide association studies of agronomic traits of sorghum". *Plant and Animal Genome Conference* -San Diego, CA. January 2013.
Geoff Morris. "Genomic resources for complex trait dissection and molecular breeding in sorghum". *Plant and Animal Genome Conference* -San Diego, CA. January 2013.
Davina Rhodes. “Diversity, genetics, and health benefits of sorghum polyphenols". *Crop Science Society of America Annual Meeting* -Tampa, FL. November 2013.
Rick Boyles. Genome-wide association studies of grain traits in Sorghum [*Sorghum bicolor* (L.) Moench]”. *Crop Science Society of America Annual Meeting* -Tampa, FL. November 2013.
Zach Brenton. “Towards more climate resilient cereal crops: insights from genomic diversity studies in sorghum”. *University of South Carolina Magellan Scholars Discovery Day* -Columbia, SC. April 2013.
Davina Rhodes. “Diversity genetics and health benefits of sorghum polyphenols". *Healthy Eating in Context: Bridging Gaps, Linking Communities* -Columbia, SC. March 2013.
Zach Brenton. “Towards more climate resilient cereal crops: insights from genomic diversity studies in sorghum”. *Healthy Eating in Context: Bridging Gaps, Linking Communities* - Columbia, SC. March 2013.
N. Shakoor. “Whole genome gene expression profiling of multiple tissue types and development stages in Sorghum Bicolor”. *Plant and Animal Genome Conference*-San Diego, CA. January 2013.
Geoff Morris. “Genomic basis of agro-climatic adaptation during crop diffusion in sorghum”. *Plant and Animal Genome Conference* -San Diego, CA. January 2013.

Community Resources Generated:

Other Products/Outcomes: We are the initial stages of negotiating with seed companies for access to germplasm.

Training:
Post-doctoral researchers: Geoff Morris (now Kansas State University) and Thomas Stefaniak (now North Dakota State University)
Graduate students: Rick Boyles, Davina Rhodes, and Nadia Shakoor, University of South Carolina; Mark Taylor (now at UC-Davis)
Technical personnel: Zach Brenton, Adrienne Lewandowski, and Matt Myers, University of South Carolina

Collaborations:
For access to genetic resources and complementary genomic studies of bioenergy traits, we have had collaborations with Chromatin, Inc., NexSteppe, Inc., and 3e-bioenergy (all bioenergy/biotechnology-focused companies). For ionomic studies, we have worked with the Ivan Baxter of the Danforth Center. For ongoing genotyping-by-sequencing efforts, we collaborate with Ed Buckler, USDA, ARS and Sharon Mitchell, Cornell University.

An Integrated Approach to Improving Plant Biomass Production
Jan E. Leach

Jan E. Leach (PD); Colorado State University; jan.leach@colostate.edu
Daniel Bush; Colorado State University; D.Bush@colostate.edu
Andrew Kern; Rutgers University; kern@biology.rutgers.edu
Hei Leung; International Rice Research Institute; h.leung@irri.org
John McKay; Colorado State University; jkmckay@colostate.edu
Bingyu Zhao; Virginia Tech; bzhao07@exchange.vt.edu

Objectives and Accomplishments

We are translating functional genetic understanding of traits important for biomass production from rice to switchgrass. In addition, we are developing a high-throughput phenotyping platform to assess biomass traits in segregating plant populations as a means to expedite gene discovery. Below, we provide highlights of progress for these two research objectives.

1) Discover gene models, polymorphism and biomass QTL in switchgrass. Transcriptome data for two cultivars of *P. virgatum*, 'Alamo' and 'Dacotah', was generated to identify sequences responsible for their large biomass differences. The current effort to establish a high-quality reference genome and annotation set for *P. virgatum* has yielded a first-draft assembly and set of gene models. This effort -- lead by the US DOE Joint Genome Institute -- created a preliminary assembly consisting of hundreds of thousands of genome contigs for the AP13 variety of Alamo. Thus far, RNA-seq data has been used to improve annotation beyond computational gene prediction. Here we do not analyze their data, but we use it as a framework to align our transcriptome reads to a common reference. We have also done a *de novo* assembly of the transcriptome of each cultivar, and generated gene models and variant calls. To estimate the variation captured within each cultivar, we evaluated SNP discovery based on an increasing scale of included reads. Even with deep transcriptome sequencing, the full breadth of variation in switchgrass has not been fully described. We will report on differential expression of transcripts shared between the two cultivars.

We developed a pseudo F2 population by crossing 'Alamo' and 'Dacotah'. The 180 F2 plants were planted in replicated field plots, and are being evaluated for various agronomic/biomass traits. To add gene-based markers and increase marker density for QTL mapping in switchgrass, cDNA from these F2 plants are being sequenced to create a coding sequence-based genetic linkage map for switchgrass. Barcoded cDNA libraries for the F2 were submitted for RNA-Seq. These data will provide (1) genome-wide SNP markers for QTL analysis, (2) quantitative data on gene expression differences segregating in this population, (3) data on recombination fractions and linkage among transcripts. Also, these data will allow for improving the genome assembly, but using linkage to constrain the possible states of order and orientation of the existing 200,000 scaffolds.

2) **Combining Field-based High Throughput Phenotyping and Genomics to Increase Crop Biomass Yield.** Improving crop yields through breeding requires genetic variation and precise

prediction of a genotype's fit to a target environment. Private sector maize breeding demonstrates the power of large numbers, combining data on yield, underlying traits, and genomic variation to maximize selection. Measuring DNA polymorphism is now trivial, leaving precision phenotyping as the gauntlet for crop breeding. Collaborating with the International Rice Research Institute (IRRI), we have optimized field-based spectral imaging to measure 5,000 test plots in 2 h, data that predict variation in rice yield. We are using this field-based high throughput phenotyping to discover alleles and high throughput traits that can be used to predict biomass and yield components.

We will report on field trials with 1400 RILs that were conducted in spring 2013 at IRRI. Fields were scanned with spectral imaging on a weekly basis during the growing season. The sensors used include Sonar, Crop Circle ACS-430, and IR. Together, this suite of proximal sensors allows capture of phenotypic data for large populations via a number of spectral vegetation indices, such as NDVI (normalized difference vegetation index) for biomass, NBI (nitrogen balance index) for nitrogen content, and several other indices for photosynthetic activity, leaf area index (LAI), etc. A large effort was focused on optimizing the spectral data collection system. For ground-truthing of the spectral data, flowering time and various biomass-relevant data were collected. Initial results from 2013 indicate a high level of genetic correlation between spectral measurements and key traits including biomass, height, and yield. The heritability of these proxy spectral traits also looks promising for QTL mapping. These data will identify: 1) new non-destructive measures that predict biomass and 2) genes controlling other physiological traits that control plant growth and yield that are not accessible after typical end-of-experiment destructive harvesting.

Broad Impacts: The genetic and genomic resources developed for switchgrass, the discovery of gene candidates in the model rice and their translation to switchgrass, and the development of high throughput phenotyping systems for evaluating biomass traits at the field level are key tools needed for the improvement of switchgrass (and other grass species) as bioenergy crops.

Deliverables

Publications in 2013:

Yang Z, Z Shen, H Tetreault, L Johnson, B Friebe, T Frazier, LK Huang, B Xu, C Burklew, X-Q Zhang, **B Zhao**, 2013. Production of autopolyploid lowland switchgrass lines through in vitro chromosome doubling. Bioenergy Research (10.1007/s12155-013-9364-x).

Sathitsuksanoh N, B Xu, **B Zhao**, YH Zhang, 2013. Overcoming biomass recalcitrance by combining genetically modified switchgrass and cellulose solvent-based lignocellulose pretreatment. PLoS ONE 8(9): e73523.

Tanger, P, JL Field, CE Jahn, M DeFoort, **JE Leach**. 2013. Biomass for thermochemical conversion: targets and challenges. Frontiers in Plant Science 4:218 doi: 10.3389/fpls.2013.00218.

Bandillo N, Raghavan C, Muyco PA, Sevilla MA, Lobina IT, Dilla-Ermita CJ, Tung CW, McCouch S, Thomson M, Mauleon R, Singh RK, Gregorio G, Redoña E, **Leung H.** 2013. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding Rice doi: 10.1186/1939-8433-6-11

Xu B, N Sathitsuksanoh, Y Tang, MK Udvardi, JY Zhang, Z Shen, M Balota, K Harich, PY-H Zhang, **B Zhao**, 2012. Overexpression of *AtLOV1* in switchgrass alters plant architecture, lignin content, and flowering time. PLoS One, 7(12): e47399.

Hupalo D, **AD Kern**. Conservation and functional element discovery in 20 angiosperm plant genomes. 2013. Mol Biol Evol Mar 11 doi:pii:g3.113.005637v1.

Oral/ Poster Presentations:

20 Oral presentations and 4 posters were presented.

Training: Post-doc Julius Mojica, PhD student Paul Tanger, and undergraduate Paul Langlois are engaged in the high-throughput phenotyping experiment and in mapping rice biomass QTL. Post doc Bettina Broeckling, PhD student Brad Tonnessen, MS student Rashad Reed, and undergraduate Emily DeLorean analyzed rice mutants for discovery of biomass related genes. Taylor Frazier, PhD student, developed the switchgrass F2 population, generated and characterized transgenic switchgrass plants. Bin Xu, PhD student, developed transgenic switchgrass lines with altered plant architecture and cell wall lignin content. Caitlin Burklew, technician, helped with tissue culture and switchgrass field work. Zhiyong Yang, postdoc, generated lowland polyploid switchgrass lines through tissue culture. Daniel Hupalo (PhD student) analyzed the switchgrass transcriptome and established a browser.

Collaborations: Collaborating with Jesse Poland and Mike Gore on developing standards and data flow for reviewing and sharing high throughtput phenotyping data.

Modulation of Phytochrome Signaling Networks for Improved Biomass Accumulation Using a Bioenergy Crop Model

Todd Mockler

Project Director: Todd Mockler, Ph.D., Donald Danforth Plant Science Center, tmockler@danforthcenter.org
Co-PD: Sam Hazen, Ph.D., University of Massachusetts-Amherst, hazen@bio.umass.edu

Project website (URL): http://www.brachypodium.org/ and http://www.mocklerlab.org/

Objectives and Accomplishments

We have cloned and validated several hundred Brachypodium transcription factor (TF) ORFs. These include a number of TFs specifically of interest in this project based on their predicted functions in phytochrome signaling networks. We have cloned and integrated into yeast strains a collection of 768 synthetic promoters for use in yeast one-hybrid assays with Brachypodium TFs. We have developed and deployed a Thermo Automation robotics system for automating high-throughput yeast one-hybrid screening. We are currently screening the collection of synthetic promoters using selected TFs.

In collaboration with a team of researchers we have designed and deployed a LemnaTec-based high-throughput plant phenotyping facility at the Danforth Center. The system has the capacity to image 1140 Brachypodium plants daily with visible, NIR and fluorescent imaging

stations. The phenotyping platform's controlled-environment chamber enables precise control of environmental variables, including light, temperature, and humidity, and weighing and watering stations allow for precise control and quantification of water availability and usage.

We have developed and deployed a Brachypodium tissue culture and transformation pipeline at the Danforth Plant Science Center. To facilitate our study of Brachypodium transcription factors implicated in phytochrome signaling networks, in collaboration with Dr. James Carrington's group (Danforth Center) we have developed new constructs and methods for amiRNA knockdown of target genes in Brachypodium. These constructs are being used in the Brachypodium transformation pipeline to knock down target transcription factors.

We used Brachypodium whole-genome microarray based gene expression profiling datasets to generate a gene network model through co-expression relationships among Brachypodium genes. We identified a network module strongly over-represented with genes implicated in or associated with photosynthesis. Closer examination of this module and identification of TFs having connections to the photosynthesis genes in this module revealed a subnetwork of putative photosynthesis regulating Brachypodium TFs and their targets that includes several of the TFs we had originally prioritized for in planta perturbation in this project. We are focusing current yeast one-hybrid studies on interrogating the promoters of the identified subnetwork of Brachypodium photosynthesis genes with the identified putative phytochrome-signaling associated TFs. We are also prioritizing these TFs for in planta overexpression and knockdowns.

Broad Impacts

The genesis of this project was our observation that biomass accumulation was greater in Brachypodium, Switchgrass, and rice plants grown in monochromatic red light (i.e. with a very high red:far-red ratio) than in blue or white light, despite the levels of photosynthetically active radiation (PAR) being the same in all conditions. Our studies suggest that the enhanced biomass productivity observed in monochromatic red light may be a consequence of phytochrome signaling that modulates shade-avoidance and photosynthesis. Modulation of photosynthesis by manipulating phytochrome signaling may open the door to engineering crops for enhanced *intrinsic yield*, which unlike *yield protection,* has not necessarily been a target for crop improvement efforts to date.

Deliverables

Publications:

Trabucco GM, Matos DA, Lee SJ, Saathoff AJ, Priest HD, Mockler TC, Sarath G, Hazen SP. Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in Brachypodium distachyon. BMC Biotechnol. 2013 Jul 31;13:61. doi: 10.1186/1472-6750-13-61. PMID: 23902793

Priest HD, Fox SE, Rowley ER, Murray JR, Michael TP, and Mockler TC. Analysis of global gene expression in Brachypodium distachyon reveals extensive network plasticity in response to abiotic stress. Manuscript in Review.

Oral/ Poster Presentations:

Interrogating transcriptional networks in grasses. University of Missouri - Columbia. April 3, 2012

Interrogating transcriptional networks in grasses. University of California, Los Angeles. May 21, 2012

Community Resources Generated:
T-DNA lines and transcription factor cDNA clones being generated in this project will be made available in the future following appropriate validation.

Other products/ outcomes:

This project has led to two invention disclosures that were submitted to the Danforth Plant Science Center and subsequently licensed by an early stage biotechnology company, Benson Hill Biosystems, Inc.

Training
Undergraduate Students: Livingstone Nganga (transcription factor cloning). Graduate Students: Henry Priest (gene network modeling, T-DNA line screening). Technical Personnel: John Gierer, Madeline Wiechert, Darren O'Brien (transcription factor cloning, yeast one-hybrid platform development)

Collaborations
In collaboration with Dr. James Carrington's group (Danforth Plant Science Center) we have developed new constructs and methods for amiRNA knockdown of target genes in Brachypodium (manuscript in preparation). On collaboration with the Carrington group we are also burning-in and benchmarking the LemnaTec high-throughput phenotyping platform for analyzing the T-DNA lines generated in this project for potential growth-rate and biomass associated phenotypes.

Genomics of Bioenergy Grass Architecture
Andrew H. Patterson

Project Director: Andrew H. Paterson, University of Georgia, paterson@plantbio.uga.edu

Objectives and Accomplishments:
The overall objective of this proposed research is to increase knowledge of the genetic determinants of components of plant architecture that are important to productivity and harvestability of bioenergy grasses, also characterizing allelic and haplotype variation in salient sorghum genes toward their deterministic utilization in crop improvement.
Specific objectives of the proposed research, and progress to date, are:

a. *Genetically-dissect components of plant architecture that are important to productivity of bioenergy grasses under annual or perennial production systems.* Forward genetics in three populations that span the range of variation in eusorghums will provide baseline QTL data for components of plant architecture and related traits, also assessing inter-relationships with one another and perenniality. Phenotyping is complete in all three populations, with genetic analysis complete in one and ongoing in the others.

b. *Investigate levels and patterns of DNA sequence variation in positional and functional candidate genes for association with phenotypic variation in plant architecture and other traits.* Phenotyping and resequencing of positional and/or functional candidate genes in a validated panel that broadly and deeply samples *S. bicolor* genetic diversity, will narrow the locations of QTLs, reveal haplotype diversity in trait-controlling regions, and perhaps even identify functional variants in some instances. Candidate genes are being investigated based on results from the population for which genetic analysis is complete, and remain to be identified for the other two populations.

c. *Enrich online resources for meta-analysis and deterministic utilization of variants in plant architecture.* We will facilitate searches for positional candidate genes affecting plant architecture by supplementing a CMAP-based QTL resource being developed under current program funding for Miscanthus, with new QTLs as well as orthologs, paralogs or other homologs of major genes cited below (plus new additions and any inadvertent omissions) that qualitatively impact plant architecture. Further, we will facilitate integrative use of positional, diversity, and mutant information in discovery and utilization of genetic variation by using Gramene trait ontologies to 'interleave' three well-characterized but to date isolated genetic resources as described below. The QTL resource was recently published, and progress is ongoing.

Broad Impacts:

Variations in plant architecture are fundamental to human utilization of the Poaceae grasses for a wide range of purposes, and are likely to become more important in the more volatile and stressful (to plant growth) climates anticipated in the future. The 'design' of crop genotypes optimized for production of biomass (or food!) from a range of environments will benefit from better understanding of the genetics of components of plant architecture, as well as the inter-relationships among these with one another and with input levels and production strategies.

Deliverables

Publications:

Dong Zhang, Hui Guo, Changsoo Kim, Tae-Ho Lee, Jingping Li, Jon Robertson, Xiyin Wang, Zining Wang and Andrew H. Paterson. CSGRqtl, a comparative QTL database for Saccharinae grasses. 2013. Plant Physiology, 161(2):594-599**.**

Wenqian Kong, Huizhe Jin, Changsoo Kim, Valorie H. Goff, Tae-Ho Lee and Andrew H. Paterson. 201#. Quantitative Trait Analysis of Branching of Recombinant Inbred Lines for *Sorghum bicolor* × *S. propinquum* ###:###-###. (near submission)

Oral/ Poster Presentations:

Accelerating the domestication of dedicated biomass crops. AgriGenomics World Congress, Frankfurt, Germany, 4 Sept 2012 (Andrew H. Paterson).

USDA-DOE Plant Feedstock Genomics for Bioenergy Awardee Meeting 2013. Genomics of Bioenergy Grass Architecture. Andrew H. Paterson

International Plant and Animal Genome Conference, January 2014 (pending). Exact title pending, talk invited by Perennial Grasses workshop, Wenqian Kong.

Community Resources Generated:
RIL sets are available from the National Plant Germplasm System (BTx623 x IS3620C), the other is distributed by the USDA-ARS unit in Lubbock, TX (BTx623 x S. propinquum).

Other products/ outcomes:
Not applicable.

Training:
Wenqian Kong – Ph.D student, co-leads phenotyping with the PI and leads data analysis.
Dong Zhang – Ph.D. student, leads online resource development.
Jon Robertson – B.S. level research coordinator, assisted in online resource development and in automation of phenotyping.
Valorie Goff - B.S. level research coordinator, oversaw field plot care and coordinates student assistants in phenotyping.
Yelena Lugin – undergraduate assistant in phenotyping and field care.
Jay Cromwell– undergraduate assistant in phenotyping and field care.

Collaborations:
Not applicable.

Quantifying Phenotypic and Genetic Diversity of *Miscanthus sinensis* as a Resource for Knowledge-Based Improvement of *M.* ×*giganteus* (*M. sinensis* × *M. sacchariflorus*)

Erik J. Sacks

Project Director: Erik J. Sacks, University of Illinois, esacks@illinois.edu

Co-PDs:
Joe Brummer, Colorado State University, Joe.Brummer@ColoState.EDU
Xiaoli Jin, Zhejiang University, jinxlzju@163.com
Stephen Long, University of Illinois, slong@illinois.edu
Toshihiko Yamada, Hokkaido University, yamada@fsc.hokudai.ac.jp
Chang Yeon Yu, Kangwon National University, cyyu@kangwon.ac.kr

Objectives and Accomplishments:
Objectives:

1. Determine population structure for *M. sinensis* genotypes from throughout the species natural distribution in China, Japan and Korea.
2. Quantify phenotypic variation in *M. sinensis* for key traits at field trial sites in the U.S., Canada and Asia.
3. Identify genes governing key traits, through the testing of candidate genes from other grasses, and genome-wide association mapping.

Miscanthus is a perennial C_4 grass that is a leading potential feedstock for the emerging bioenergy industry in North America, Europe, and China. However, only a single sterile

genotype of *M. ×giganteus*, a nothospecies derived from diploid *M. sinensis* and tetraploid *M. sacchariflorus*, is currently available to farmers for biomass production. To facilitate breeding of *Miscanthus*, we characterized genetic diversity and population structure of *M. sinensis*. We studied 767 accessions, including 617 *M. sinensis* from most of its native range in China, Japan, and South Korea, 77 ornamental cultivars from the U.S. and Europe, and 43 naturalized individuals from the U.S. The accessions were screened with 21,207 RAD-Seq SNPs obtained via the UNEAK pipeline in TASSEL, 424 GoldenGate SNPs, and ten chloroplast microsatellite markers. Population structure was analyzed with STRUCTURE 2.3.4, and via Discriminant Analysis of Principle Components (DAPC) using the R package adegenet.

We identified six genetic clusters of *M. sinensis* from geographically distinct regions in Asia. Genetic data indicated that 1) Southeast China was the origin of *M. sinensis* populations found in temperate eastern Asia, which was consistent with this area having likely been a refugium during the last glacial maximum (LGM), 2) *M. sinensis* migrated directly from Southeast China to Japan before migrating to the same latitudes in China and Korea, which was consistent with the known sequence of warming post-LGM, 3) ornamental *M. sinensis* cultivars from the U.S. and Europe were derived from the Southern Japan population, and U.S. naturalized populations were derived from a subset of the ornamental cultivars, and 4) many ornamental cultivars previously considered to be entirely *M. sinensis* have, in actuality, hybrid ancestry from *M. sacchariflorus* and *M. sinensis*, whereas U.S. naturalized populations of *M. sinensis* do not. Replicated field trials in IL, CO, Canada, Japan, and South Korea were established in 2012 and data on overwintering ability, and flowering time was first obtained from established plants in 2013; data on yield and yield components is being taken in autumn-winter (post dormancy).

Two candidate genes, *Hd1* and *MsGPI*, were also studied. *Hd1* in grasses typically acts as repressor to flowering under long day conditions. We cloned and sequenced *Hd1* for 38 accessions from a broad range of latitudes in Japan and China. We identified non-functional alleles due to insertions and deletions. MITEs that produced a premature stop codon were found in *Hd1* alleles of Japanese accessions but not Chinese accessions. For *MsGPI*, transient overexpression in *Nicotiana benthamiana* leaves demonstrated that is involved in antioxidant metabolism and that it is a transcriptional regulator of *NbAPX* and *NbPAL*. *MsGPI* gene expression patterns differed significantly when plants were exposed to NaCl, drought, ABA, or methyl viologen (oxidative stress), suggesting that *MsGPI* is related to the abiotic stress response.

Broad Impacts:
The results from our study have, for the first time, provided a regional-level understanding of population structure in *M. sinensis*, and insights into its most recent phase of evolution. The genetic bottleneck associated with U.S. *M. sinensis* germplasm was a previously unknown limitation to breeding improved bioenergy feedstock cultivars of *Miscanthus*. The present study indicates that there is an opportunity to broaden the genetic diversity of *M. ×giganteus* by using *M. sinensis* parents from populations other than Southern Japan.

Deliverables

Publications:

- Clark, L.V., J.E. Brummer, K. Głowacka, M. Hall, K. Heo, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, and E.J. Sacks (Submitted to Annals of Botany) A footprint of past climate change on the population structure of *Miscanthus sinensis*.
- Seong, E.S, J.H. Yoo, J.G. Lee, H.Y. Kim, I.S. Hwang, K. Heo, J.D. Lim, D.K. Lee, E.J. Sacks, and C.Y. Yu. 2013. Transient overexpression of the *Miscanthus sinensis* glucose-6-phosphate isomerase gene (*MsGPI*) in *Nicotiana benthamiana* enhances expression of genes related to antioxidant metabolism. POJ 6:408-414.

Oral/ Poster Presentations:

- Clark, L.V., J.E. Brummer, M. Hall, S. Long, J. Peng, T. Yamada. J.H. Yoo, C.Y. Yu., H. Zhao, and E.J. Sacks. Genetic structure of *Miscanthus sinensis* from Asia and the United States. International Plant & Animal Genome XXI. San Diego, CA. 12-16 January 2013.
- Nagano H., N. Uchino, J. Peng, E.J. Sacks, and T. Yamada. Sequence diversity of *Co*/*Hd1* homologs in *Miscanthus sinensis*. International Plant & Animal Genome XXII Conference. San Diego, CA. 11-15 January 2014.
- Sugisawa S, H. Nagano, M.. Dwiyanti, .E.J. Sacks, and T. Yamada. Relation of Pseudo-Response Regulator (PRR)-like gene to flowering time in *Miscanthus* genome. 124th Meeting of the Japanese Society of Breeding. October 2013.
- Nagano H., N. Uchino, E.J. Sacks, and T. Yamada. Paralogue in *Hd1* gene of *Miscanthus sinensis*. 123th Meeting of the Japanese Society of Breeding. March 2013.

Training

Postdoctoral: Lindsay Clark, Maria Stefanie Dwiyanti, Kossonou Guillaume Anzoua
Graduate students: Ji Hye Yoo
Undergraduate students: Melina Salgado, Shun Sugisawa
Technical: Benjamin Baechle, Logan Smith, Hironori Nagano, Norihiko Uchino, Kumi Green, Maiko Ohta

Collaborations:

New Energy Farms—planted trial in collaboration with industry leading company
C.S. Prakash, Tuskegee University—provided plants and helped establish a field trial in support of their USDA-NIFA project, Linking DNA Markers to Key Bioenergy Traits in *Miscanthus*

Deciphering Natural Allelic Variation in Switchgrass for Biomass Yield and Quality using a Nested Association Mapping Population

Malay C. Saha

Project ID: 0018448, Award Register#: ER65456
Program Manager: Catherine M. Ronning; PI: Malay C. Saha

Progress report: December 20, 2013

Understanding the genetic basis of quantitative traits is essential to facilitate genome-enabled breeding for the improvement of biomass yield and feedstock quality in switchgrass. Nested association mapping (NAM) technique was especially designed for identifying and

dissecting the genetic architecture of complex traits. The long-term goal of the switchgrass NAM project is to understand the genetic basis of the key traits of biomass yield and composition. The specific objectives are to: i) develop a NAM population of switchgrass and construct genetic maps; ii) identify quantitative trait loci (QTL) and molecular markers associated with biomass yield, feedstock quality and other traits and iii) validate markers in breeding populations.

A total of 15 diverse genotypes were crossed to AP13, the recurrent parent. Chain crossing scheme was performed using 10 hybrid individuals selected from each of the 15 crosses (Fig. 1). The hybrids were confirmed by SSR marker profiling. Seedlings from each chain cross (pseudo F_2 progenies) were finally accumulated as switchgrass NAM population consisting of 2,000 individuals.

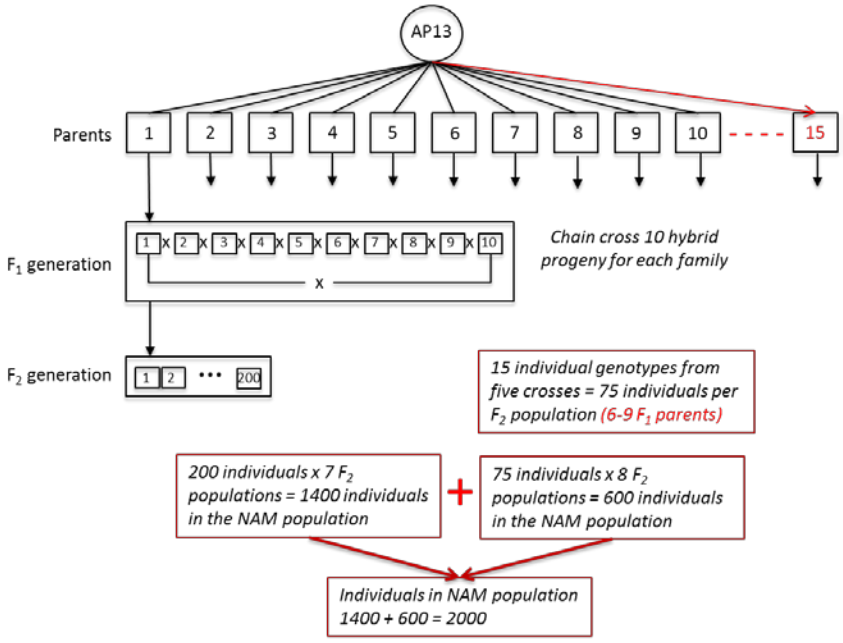


Figure 1: Crossing scheme for the development of a switchgrass NAM population.

Population development and field planting: All the 2,000 individuals, parents and grandparents were grown in gallon pots (Fig. 2A) at 32°C day and 21°C night temperature under 16 h photo period in greenhouse of the Noble Foundation. Multiple copies were then made from every mother plants (Fig. 2B). Field planting was conducted following the Alpha Lattice design at Knoxville, TN and Ardmore, OK (Fig. 2 C and D). Each replication consists of 2,350 genotypes including parents and checks. Planting was done in Knoxville, TN on July 19, 2013 and at Ardmore, OK on August 1, 2013. Significant variations have been observed in the population and data were collected on tiller density, stem rust reactions and panicle color. Harvesting of the population was completed on December 17, 2013 at Knoxville. Due to excessive soil moisture, harvesting has been delayed in Ardmore, OK.

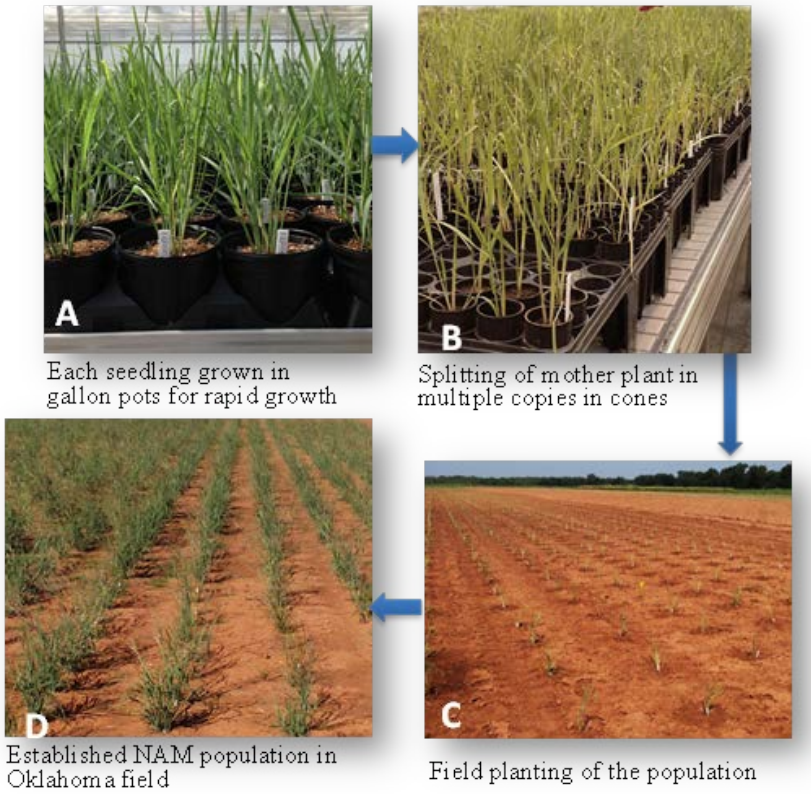


Figure 2: All grandparents, parents and progenies of the NAM population were grown in gallon pots for rapid growth (A); multiple copies of each plant were grown in cones (B); plant in cones were then planted in field (C) and well established plants at Ardmore, Oklahoma (D).

Genotyping: The AP13 genotype, recurrent parent of the NAM population, was used for whole genome sequencing project. The v1.1 version of the AP13 genome has been recently released (http://www.phytozome.net/panicumvirgatum_er.php#A). Genomic DNA from the 15 diverse parents was sequenced at JGI. Analysis of the sequence data is in progress. Preliminary analysis of the sequences obtained from the parent, CDV09_05, indicated a total of 323 M reads, of which 249 M reads were mapped to the switchgrass genome. Number of reads assembled to each switchgrass chromosome varied from 10.3 to 20.6 M. The highest number of reads was assembled in chromosome 2 and the lowest number in chromosome 11. Very high level of pairwise identities (between 88-98%, average of 93%) was observed in the analyzed data set. DNA extraction for the whole population is in progress.

Project management: Three project coordination meetings were completed during recent project period. PI, Malay Saha, visited the NAM planting at Knoxville, TN field. Project update meetings with Co-PIs were organized during professional society meetings and personal visits. Next project coordination meeting has scheduled during the PAG conference at San Diego, CA.

Genetic Architecture of Sorghum Biomass Yield Component Traits Identified Using High-Throughput, Field-Based Phenotyping Technologies
Patrick Schnable

Project Director: Patrick Schnable, Iowa State University, Schnable@iastate.edu

Co-PDs: Lie Tang, Iowa State University, lietang@mail.iastate.edu; Maria Salas-Fernandez, Iowa State University, mgsalas@mail.iastate.edu

Project website (URL): Not applicable

Objectives and Accomplishments:
A systems approach (Genome-wide Association Studies; GWAS) is being used to identify the genetic control of rates of photosynthesis, photo-protection, and biomass growth, as well as a series of biomass yield-related plant architecture traits (e.g., plant height, stalk diameter, leaf number, leaf width, leaf length, leaf angle, leaf area index) in the C4 grass sorghum [Sorghum bicolor (L.) Moench], a promising and productive biomass crop. These experiments will identify SNP markers within or closely linked to the genes that control these traits. Using these identified SNPs it will be possible to predict the phenotypes of sorghum lines based on their underlying genotypes and conduct genomic selection experiments designed to improve biomass yields of sorghum hybrids. Higher photosynthetic rates contribute to increase biomass yields. Although essential for photosynthesis, excess light can lead to the generation of harmful reactive oxygen species (ROS). Plants have developed several mechanisms to protect themselves from the consequences of excess light, which are collectively termed photo-protection. The genetic regulation of photosynthetic rates and photo-protection will be identified. In addition, the hypotheses that significant amounts of variation in growth rate can be explained by variation in photosynthetic rates and/or the amount of photo-protection will be tested. Although total biomass yield is a function of growth rate and growth duration, growth rates are typically not constant throughout the growing season. Hence, the potential exists to identify distinct genetic loci that control growth rates and plant architecture at different times in the growing season. Once favorable alleles of loci that control rates of photosynthesis, photo-protection and biomass growth, as well as plant architecture traits have been identified, breeders can use genomic selection methods to produce sorghum hybrids having higher biomass yields. To identify the genetic control of dynamic changes in biomass growth rates and plant architecture, it will be necessary to collect trait data at multiple times during two growing seasons. It would be extremely challenging to do so using conventional approaches. Instead, these data are being collected using a sophisticated, high-throughput, field-based, plant phenotyping system that is being developed during the project. Over the last several years, substantial progress has been achieved in the development of automated phentoyping systems. But to date most automated phenotyping systems have been laboratory- or greenhouse-based. These systems suffer from the limitation that plant performance in laboratories or greenhouses is often only poorly correlated with field performance. Hence, the field-based phenotyping system that is being developed during the project has the potential to revolutionize the collection of phenotypic data from field-based biomass yield trials. As such, this robotic system is expected contribute widely to the genetic improvement of biomass crops.

Seed stocks were increased for the 300 lines in the Yu Panel of photoperiod sensitive lines during the 2012 winter nursery. Two biological replicates of these lines, plus 315 sorghum lines of the photoperiod insensitive panel were grown in IA during the summer of 2013. Leaf samples were collected from the Yu Panel lines grown in the 2013 summer nursery. Freeze-dried leaf samples were powdered, and DNA was extracted. DNA samples are awaiting QC and subsequent SNP analyses.
Photosynthesis and light-adapted fluorescence parameters were collected from the sorghum lines in the photoperiod insensitive panel. There was significant variation in all parameters. Preliminary analysis of the data demonstrates that, as expected, there was a positive and significant correlation between photosynthesis and stomatal conductance. Data were collected over 11 days from 10:00am to 5:00pm. Weather data were collected to determine the potential effect of variation on temperature, radiation, wind and humidity on photosynthesis and fluorescence parameters. Preliminary analysis suggests that there was no effect of radiation variation within a day on the photosynthesis. Variation between days will be included in the statistical model to correct for these effects.
A phenotypic data acquisition system that automatically collects stereo images of sample plants was constructed. Sorghum images were collected from 4 plant types (single short/young plants, short/young plants with tillers, single tall/mature plants and tall/mature plants with tillers) on which the image processing algorithms will be tested. Biomass yield was determined for a subset of lines (10%) to be used in the algorithm development process and to validate the algorithm. Plant architecture parameters including plant height, leaf number, leaf angle, stem diameter, number of tillers, number of internodes, green weight and dry weight were also manually collected. Images were also collected from these for validation and algorithm development.
Broad Impacts: Not Applicable

Deliverables
Publications:
Oral/ Poster Presentations:
Salas Fernandez, M.G. Improving the photosynthetic capacity of sorghum: current and future perspectives. Great Plains Sorghum Conference and 29th Sorghum Research and Utilization Conference. Kansas State University, August 2012.
Schnable, P., Next Generation Genotyping and Crop Improvement. BASF, Research Triangle Park (RTP), NC, February 2013.
Schnable, P. Next Generation Phenotyping and Breeding. 6th Annual University of Minnesota Plant Breeding Symposium, Minneapolis, MN, March 2013.
Community Resources Generated: Not Applicable

Other products/ outcomes: Not Applicable

Training: Graduate student Ying Bao was trained on the construction and optimization of the phenotyping robot. Graduate student Jing Zhao was trained in the collection and analysis of photosynthesis and light-adapted fluorescent data. Undergraduate student Jieyun Hu received training on the use of Lic-COR to be able to collect high quality data and interpret the parameters collected by the equipment.

Collaborations: Not Applicable

The Genomic Basis of Heterosis in High-Yielding Triploid Hybrids of Willow (*Salix* spp.)
Bioenergy Crops
Lawrence Smart

Project Director: Lawrence Smart, Cornell University, lbs33@cornell.edu
Co-PDs: Christopher Town, J. Craig Venter Institute, CDTown@jcvi.org

Project website: http://willow.cals.cornell.edu

Objectives and Accomplishments: Yield improvement in many crops has been based on capturing heterosis, but even in well-studied species the complex genetic basis for hybrid vigor is poorly understood. Breeding for yield improvement in willow bioenergy crops has relied primarily on capturing hybrid vigor through interspecific hybridization, the best of which are triploids resulting from crosses of tetraploids with diploids. In this project, we are asking how the gene expression patterns in willow hybrids are related to their yield potential and other traits important for biofuels production. In particular, we will test if there is a bias in the expression of key genes from one parent versus the other in species hybrids, and whether there is a gene dosage effect skewing gene expression patterns in triploid progeny compared with their diploid and tetraploid parents.

Objective 1. Quantify heterosis for yield and biomass traits across eight families representing intraspecific diploids and interspecific diploid and triploid progeny. The eight families needed to complete this task have been produced with the desired number of progeny (>100), except for one with only 75 progeny, all of which were established in nursery beds. Cuttings will be made from these progeny to establish a replicated field trial in spring 2014. Two TA instruments Q500 Thermogravimetric Analyzers (TGA) were purchased to quantify cellulose, hemicellulose, lignin, and ash. A set of 100 willow samples that were characterized by wet chemical analysis have been analyzed using the TGA and multivariate analyses were used to develop a new, improved model for compositional analysis.

Objective 2. Determine allele-specific gene expression in intraspecific diploid and interspecific diploid and triploid hybrids. Ten randomly selected progeny from the six families that had been produced by 2012 (SpuF1-082, SpuF2-317, SvixSpu-421, SpuxSvi-407, SpuxSmi-415, and SvixSmi-423, see Table 1) plus all of the parents of those families (94006, 94001, 9882-34, 9882-41, MBG5027, 'Jorr', and 01-200-003) were grown from cuttings in the greenhouse in replication to produce tissue for RNA-Seq analysis. From each of the four replicates from each genotype, shoot tip tissue was pooled and young stem internode tissue was pooled. The tissue was flash frozen in liquid nitrogen, RNA was isolated using the Sigma Spectrum™ kit, and quality assessed using an Experion. Libraries have been constructed from 32 samples using the dUTP method.

Analysis of the first 6 libraries sequenced (2 lanes of HiSeq) showed high levels of read duplication in some samples. We are currently comparing two methods of library construction – dUTP (which has a final PCR step) vs Ovation (which has a final linear amplification step) to better understand and limit the extent of read duplication. Ovation shows slightly lower read duplication, but this does not account for the high duplication levels seen in the first set of samples. Further small-scale analyses (using MiSeq) are in progress to determine the optimal way to proceed with the remainder of the samples.

Young leaf tissue was collected from a single plant of each genotype and sent to Dr. K. Arumuganathan for flow cytometric analysis of nuclear DNA content. The results indicate that all of the DNA content values for the progeny are in the expected range for diploids (SpuF1-082, SpuF2-317, SvixSpu-421, SpuxSvi-407) and triploids (SpuxSmi-415 and SvixSmi-423).

We conducted whole-genome shotgun sequencing of a wild diploid, female accession *Salix purpurea*, clone ID 94006. ALLPATHS-LG was used to assemble sequences representing ~140X coverage of paired-end sequences and all of the mate-pair libraries, producing contigs with L50=46 kb and scaffolds with L50=191 kb. The ALLPATHS-LG assembly has a total size of 348 Mb and a total span of 392 Mb (11% Ns). This assembly is still relatively fragmented due to the high level of heterozygosity (1 SNP per 120 bp, or 0.8%) and low yield of long insert mate pair library. Assessment of the assembly quality against willow BACs and transcripts suggested that ~ 78% to 85% of the willow genome is captured in the current assembly. We further integrated the ALLPATHS assembly with high-resolution genetic maps developed using genotyping-by-sequencing (GBS) to create a chromosomal assembly (version 1). The scaffolds were anchored to 19 pseudo-chromosomes and ordered using a novel algorithm to maximize co-linearity with all genetic maps while simultaneously minimizing the total genetic distance traversed. A total of 276 Mb (70%) sequences can be anchored onto the chromosomes. All pseudomolecules are globally oriented with respect to *Populus trichocarpa* chromosomes. Functional annotation has been completed by JGI and is currently being interrogated prior to public release. We have set up JBrowse using the willow assembly as the reference sequence.

Assessment of parent of origin of transcripts requires accurate determination of SNPs and their proportion in each transcript assembly from the hybrid progeny. We therefore elected to conduct genome re-sequencing on all parents in this project and use the SNPs identified from genomic sequence to characterize parent- and allele-specific expression in the hybrid progeny. This sequencing is currently in progress. We will complete SNP calling in all parents using the genomic resequencing data before beginning the analyses of allele-specific expression analysis using the transcriptome data.

Broad Impacts: Sequence assembly and functional annotation of the *Salix* genome is providing a vital resource for efficient breeding of shrub willow as a biomass feedstock for conversion to biofuels and bioenergy. This project will provide a better understanding of the molecular basis for heterosis in outcrossing and polyploid species that can be applied broadly to the breeding of perennial bioenergy crops.

Deliverables

Publications:

Serapiglia, M.J., Gouker, F.E., and Smart, L.B. Early selection of novel triploid hybrids of shrub willow with improved biomass yield relative to diploids. *Submitted to BMC Plant Biology*.

Oral/ Poster Presentations:

Gouker, F., Serapiglia, M., Tang, H., Town, C., Buckler, E., Mitchell, S., Elshire, R., Hyma, K., Rodgers-Melnick, E., DiFazio, S., Barry, K., Lindquist, E., Schmutz, J., Tuskan, G., Smart, L. "Sequencing and Assembly of the *Salix purpurea* Genome and Transcriptome to Improve Shrub

Willow for Biomass Production", International Plant and Animal Genome Conference (PAG-XXI), Jan. 13, 2013, San Diego, CA. *Oral presentation*.
Smart, L.B. "Genomic approaches to improve yield and biofuels conversion efficiency of shrub willow", Cornell University, Dept. of Plant Biology, Ithaca, NY, Feb. 1, 2013. *Invited departmental seminar*.
Smart, L.B., Gouker, F.E., Serapiglia, M.J., Town, C.D., Tang, H., Buckler, E.S., Elshire, R.J., Mitchell, S.E., DiFazio, S., Rodgers-Melnick, E., Tuskan, G.A., Carlson, J.E., Miller, R.O., Volk, T.A., and Fabio, E.S. "Development of genomic resources and novel species hybrids for the genetic improvement of shrub willow feedstock crops", 2013 Genomic Science Annual Contractor-Grantee Meeting/USDA-DOE Plant Feedstock Genomics for Bioenergy Program Meeting, Feb. 24-27, 2013, Bethesda, MD. *Poster and short oral presentation*.
Carlson, C.H., Gouker, F.E., Serapiglia, M.J., Tang, H., Krishnakumar, V., Town, C.D., Tuskan, G.A., Rokhsar, D., Goodstein, D.M., Shu, S., Barry, K.W., Lindquist, E.A., Zhou, R., DiFazio, S., and Smart, L.B. "Annotation of the *Salix purpurea* L. Genome and Gene Families Important for Biomass Production" International Plant and Animal Genome Conference (PAG-XXII), Jan. 12, 2014, San Diego, CA. *Poster presentation*.
Community Resources Generated: Together with JGI, we have completed the first draft sequence and annotation of the willow genome and will soon release this on Phytozome.

Other products/ outcomes:

Training: Michelle Serapiglia - postdoctoral scientist (Cornell); Craig Carlson, Ph.D. student (Cornell); Haibao Tang and Vivek Krishnakumar- postdoctoral scientists (JCVI)

Collaborations: This project in combination with other funding from USDA has strengthened our collaboration with Steve DiFazio (West Virginia Univ.) and Jerry Tuskan (Oak Ridge National Lab).

The Dual Effect of Tubulin Manipulation on Populus Wood Formation and Drought Tolerance

CJ Tsai

Project Director: CJ Tsai (cjtsai@uga.edu), University of Georgia
Co-PDs: Scott Harding (sharding@uga.edu), University of Georgia; Michael Hahn (hahn@ccrc.uga.edu), University of Georgia; Gary Peter (gfpeter@ufl.edu), University of Florida; Shawn Mansfield (shawn.mansfield@ubc.ca), University of British Columbia

Objectives and Accomplishments: The main objectives are to generate transgenic *Populus* with altered alpha- (TUA) and beta-tubulin (TUB) expression and/or post-translational modifications (PTMs) in a plant-wide (with the 35S promoter), guard cell (GC)-specific or xylem-specific manner. Transgenic plants are subjected to two inductive treatments, tension wood induction and drought stress, in order to understand the effects of tubulin manipulation on cell wall biogenesis and guard cell dynamics, respectively. RNA-Seq, metabolite profiling, stress physiology, cell wall glycan profiling and wood physicochemical analyses will be conducted for a comprehensive description of the transgenic effects.

Our first-year efforts have focused on new transgenic plant production and characterization of existing transgenic lines. Well over 60 transformation trials have been conducted, targeting the native or modified forms of *TUA* and *TUB* genes either individually (single transformation) or in combination (double transformation), under three different promoters. Because of the sheer volume of work and higher-than-expected transformation response, we prioritize transformants expressing the PTM mimics of TUA1 (dY and dEY) for subsequent characterization. The first cohort of transgenics from single transformation trials is being propagated for greenhouse experimentation. Transgenic production from the double transformation trials is ongoing and greenhouse experiments will follow after single transformant characterization.
Characterization of the existing double transformation lines expressing TUA1dY+TUB (A1dYB9) and TUA1dEY+TUB15 (A1dEYB15) under control of the 35S promoter has been ongoing. Proteomics analysis confirmed the presence of TUA1dY and TUA1dEY PTM isoforms in the transgenics. Wood physicochemical analyses revealed no major changes, except significantly reduced lignin S/G ratio in the transgenic wood. Glycome profiling revealed that certain polysaccharide networks and possibly their interaction with the lignin polymers were altered in the transgenics. Tension wood (TW) experiments were conducted. Both wood physicochemical and RNA-Seq analyses identified large TW treatment effects but small genotypic differences. The most strongly affected genes were enriched with cell wall modifying enzymes, thus supporting the glycome profiling results.
We determined that short-term, acute drought stress is effective for evaluating microtubule-dependent stomatal response, whereas long-term chronic water stress leads to reduced growth and other forms of adaptation. Transgenic plants under acute drought stress exhibited a delayed stomatal closure relative to the wild type. A delayed stomatal opening in response to light was also observed in dark-acclimated transgenic leaves. These results showed that transgenic tubulin perturbation affected MT-dependent guard cell dynamics. Replicate experiments are underway and leaf samples will be collected for RNA-Seq, metabolite profiling and selected cell wall analyses. Differential leaf expansion was observed for the A1dYB9 plants, resulting in increased leaf aspect ratios. Epidermal impression analysis suggested that the altered leaf expansion was likely due to prolonged cell expansion during leaf development.

Broad Impacts: Microtubules are evolutionarily conserved with essential functions in mitosis, vesicle trafficking and cell expansion. Microtubules are also pharmacological targets for drug or herbicide development, besides their specialized roles in regulating cell wall synthesis and stomatal behavior. Altering microtubule dynamics via transgenic manipulation of tubulin PTMs, as implemented in this study, is a novel strategy and may have broad application for both basic and applied research.

Deliverables
Publications: one manuscript in preparation

Oral/ Poster Presentations:
Hu H., Xiao H., Swamy P., Tsai C.-J. (2013) Post-translational modifications of tubulins in *Populus*. IUFRO Tree Biotechnology 2013. Ashville, NC, May 26 - Jun 1, 2013 (poster).
Swamy P., Pattathil S., Maloney V., Mansfield S. Chung J.-D., Nyamdari B., Xue L.-J., Harding S., Hahn M., Tsai C.-J. (2013) Alternation in tubulin homeostasis affects cell wall properties

and drought response. IUFRO Tree Biotechnology 2013. Ashville, NC, May 26 - Jun 1, 2013 (poster).

Tsai C.-J. (2013) Beyond wood formation: Pleiotropic effects of tubulin manipulation in *Populus*. IUFRO Tree Biotechnology 2013. Ashville, NC, May 26 - June 1, 2013 (oral).

Tsai C.-J., Swamy P., Pattathil S., Xue L.-J., Johnson V., Hu H., Zhu Y.Y., Hahn M., Harding S., Maloney V., Mansfield S., Chung J.-D. (2013) Pleiotropic effects of tubulin manipulation on *Populus* wood formation and stress response. Annual Meeting of International Academy of Wood Science 2013. Nanjing, China, Oct 17-21, 2013 (oral).

Community Resources Generated: Newly generated RNA-Seq data will be deposited to SRA. Other biological data will be disseminated via publications and conference presentations.

Training:
Tsai lab: Prashant Swamy (PhD student, graduated summer 2013): characterization of existing double transformants; Hao Hu (postdoc): drought experiments, photosynthesis, proteomic data analysis; Yingying Zhu (postdoc): new transformation and RNA-Seq; Liangjiao Xue (bioinformatician): RNA-Seq and glycome profiling data analysis; Virgil Johnson (programmer): RNA-Seq analysis; Daniel Khalek (undergraduate): tissue culture; Anisha Patel (undergraduate): tissue culture.
Hahn Lab: Sivakumar Pattathil (research scientist): glycome profiling
Mansfield lab: Victoria Maloney (postdoc): wood property and chemistry analysis

Collaborations:
Hui Xiao (Department of Pathology, Albert Einstein College of Medicine) for proteomics analysis of tubulin PTMs.
Will York (Complex Carbohydrate Research Center, University of Georgia) for hemicellulose analysis.

Discovering the Desirable Alleles Contributing to the Lignocellulosic Biomass Traits in *Saccharum* germplasm Collections for Energy Cane Improvement

Jianping Wang

Project Director: Jianping Wang, Agronomy Department, University of Florida, wangjp@ufl.edu
Co-PDs: Barry Glaz, Sugarcane Field Station, USDA, ARS; Barry.Glaz@ARS.USDA.GOV; Robert Gilbert, Everglades Research and Education Center, University of Florida, ragilber@ufl.edu

Project website (URL):
http://bioinformatics.ufl.edu/Wang/projects

Objectives and Accomplishments:
Objectives 1: To identify a core collection from the World Collection of Sugarcane and Related Grasses.
In total, 1186 accessions in the world collection of sugarcane and related grasses (WCSRG) comprising 16 species collected from 45 countries were phenotyped and genotyped respectively.

Diversity analysis showed that the WCSRG has a gene diversity of 0.276. The highest gene diversity was found in *S. sponatneum* (0.306) followed by *S. robustum*(0.263). Phylogenetic and structure analysis of 1186 samples revealed 6 major groups. *S. sponataneum* accessions tend to cluster together and S. *hybrids* mostly clustered with *S. officinarum*. Based on the phenotypic and genotypic data on the WCSRG, a diversity panel of 300 clones was selected representing the majority of diversity in WCSRG and were planted in three replications at Canal Point, FL. The detailed results have been compiled into two manuscripts. One is submitted and the other one is under preparation for submission.

Objectives 2: To discover the desirable alleles contributing to biomass components in *Saccharum* spp.
Candidate genes that contribute to lignocellulosic biomass traits in plant have been selected mainly based on the literature keywords search for architecture genes and sugar metabolism genes and a blast search for cell wall related genes. In total, 380 representative unique genes were selected as candidate gene targets involved in cell wall and architecture. Allelic variance of these target genes in addition to other 56,000 genomic regions across 300 diversity panel will be detected through the Agilent sure select hybridization. About 60K probes covered these 380 candidate genes and 56,000 target regions have been designed and 12 DNA samples from 9 different genotypes in the diversity panel were used for hybridization as a pilot experiment. The dosage of allele in the diversity panel will be inferred based on the sequencing depth. We have been estimating the chromosome numbers of *Saccharum* accessions in diversity penal using flow cytometry, which will help us to estimate ploidy levels of each accession to confirm the dosage of alleles.
The diversity panel has been phenotypically evaluated for once.

Objective 3: To characterize and select advanced energy cane breeding lines for energy cane improvement.
Five new energy cane cultivars (UFCP74-1010, UFCP78-1013, UFCP82-1655, UFCP84-1047, and UFCP87-0053) were released for commercial cultivation on September 30, 2013. The released cultivars have significantly lower smut infection and comparatively greater biomass yields than commercial check (L79-1002) based on data collected from 12 location-years (5 years at Citra, 3 years at Lykes Bros., and 2 years each at Tecan and 960 farms). Five manuscripts (one for each cultivar) are under preparation.

Broad Impacts: This project provides a great opportunity to bring scientists in muldisciplinary research areas together including agronomists, breeder, molecular biologist and pathologist. Energy cane is a relatively new crop with very low genetic diversity and very few released cultivars. This project is engaged to discover new germplasm for breeding energy cane and sugarcane with higher yield and disease resistance and provide new genomic tools and resources to the community. The newly released five energy cane cultivars not only improved genetic diversity but also reduced the risk for the nascent cellulosic ethanol industry. Our industry collaborators (BP Biofuels) and two other private companies, Florida Biological Fuels Corp. and Stramit USA have a marked interest in our released energy cane cultivars. Young scientists including psotdocs, technician, graduate and undergraduate students were trained to translate polyploid genomics information into crop breeding. This will contribute to sustaining the critical

plant breeding workforce. All sequences, markers, and the cultivars resulted from this project will be available to the public upon request.

Deliverables

Publications:
Todd J., Glaz B., Irey M.S., Zhao D.,. Hu C., and. El-Hout N (2014) Sugarcane Genotype Selection on a Sand Soil with and without Added Mill Mud, Agronomy Journal. 106 (1).
Todd J., Wang J., Glaz B., Sood S., Ayala-Silva T., Nayak S. N., Glynn N. C., Gutierrez O. A., Kuhn D., Tahir M., and J. C. Comstock. (2013) Phenotypic Characterization of the Miami World Collection of Sugarcane and Related Grasses for Selecting a Representative Core. Crop Science. Submitted.

Oral/ Poster Presentations:
Phenotypic Evaluation of the World Collection of Sugarcane and Related Grasses
PAG, San Diego CA, 2013

Core Selection from the Miami World Collection of Sugarcane and Related Grasses using Phenotypic Markers
ASSCT, Panama City FL, 2013

Phenotyping of Association Mapping Panel of Sugarcane and Related Grasses
PAG, San Diego CA, 2014

Core Selection from the Miami World Collection of Sugarcane and Related Grasses using Genotypic Markers-Seeking Alleles for Biomass
ASSCT, Panama City FL, 2013

Developing a Reference Genomic Map for Sugarcane Resistance Gene
ASSCT, Panama City FL, 2013

Gilbert, R.A. and H.S. Sandhu (2013). Proposed Energycane Releases: UFCP74-1010, UFCP78-1013, UFCP82-1655, UFCP84-1047, and UFCP87-0053. Cultivar Release Advisory Committee meeting, Sept. 6, 2013, Belle Glade, FL.
Gilbert, R.A. and H.S. Sandhu (2013). Proposed Energycane Releases: UFCP74-1010, UFCP78-1013, UFCP82-1655, UFCP84-1047, and UFCP87-0053. Cultivar Release Committee meeting, Sept. 30, 2013, Gainesville, FL.

Community Resources Generated:
A diversity panel of 310 clones representing the majority diversity of WCSRG was planted at Canal Point. The scientific community has access to this panel for research.
The five energy cane cultivars are released publically and community has access to grow these cultivars. We are keeping seed for the releases at EREC and Highlands county and public can request for this seed.

Training:
Postdocs trained: Drs. James Todd and Vanessa Gordon in USDA-ARS, Canal Point; Drs.Spurthi Nayak and Jian Song in University of Florida; Dr. Hardev Sandhu at EREC, Belle Glade, FL,
Ph.D student trained: Xiping Yang in University of Florida;
Biological Scientist: Calvin Howard at EREC, Belle Glade, FL
Undergraduate assistants:; Andrea Villa, Aleksey Kurashev, G, Johnny Molestina in University of Florida.
Agronomic assistants: Jairo Sanchez, Richard Bryant at EREC, FL.

Collaborations:
Dr. Deng Zuhu and Dr. Jisen Zhang at Fujian Agriculture & Forestry University are interested in collaborating with us to extend this project by using the probes we developed in the current project to genotype the modern USA and Chinese sugarcane cultivars to determine their genomic composition and validate their pedigree.

Sorghum Biomass Genomics and Phenomics
Jianming Yu

Project Director: Jianming Yu, Kansas State University, jyu@ksu.edu
Co-PDs: Tesso, T., Kansas State University, ttesso@ksu.edu; Roozeboom, K., Kansas State University, sstaggen@ksu.edu; Wang, D., Kansas State University, wang@ksu.edu; Bernardo, R., University of Minnesota, bernardo@umn.edu; Wang, M., Plant Genetic Resources Conservation Unit, USDA-ARS, mingli.wang@ars.usda.gov; Pederson, G., Plant Genetic Resources Conservation Unit, USDA-ARS, gary.pederson@ars.usda.gov

Project website (URL):
https://sites.google.com/site/quantitativegeneticsmaize/research/genetics-breeding-for-bioenergy

Objectives and Accomplishments:
We have three project objectives. 1. Genotypically characterize a diverse set of sorghum germplasm for selective phenotyping and phenotypically assess the biomass potential of a selected representative set for yield and composition for genomewide selection (GS). 2. Develop a standardized method for high-throughput and cost-effective phenotyping for sorghum biomass composition through near-infrared reflectance (NIR) spectroscopy. 3. Discover additional useful germplasm by genomewide prediction and useful genes by association mapping for biomass yield and composition.

Objective 1 and 2 have been achieved. We selected 1009 biomass sorghum accessions from GRIN and the genotyping-by-sequencing (GBS) was conducted for 976 accessions. 720,000 SNPs were obtained. 300 sorghum accessions were selected based on GBS SNP data and were field tested at Lubbock, TX, and Manhattan, KS in 2012 and 2013.

Broad Impacts:
This strategy employed in this project will generate critical information on how to tap into the vast plant germplasm collections for biomass crop improvement, and how to increase the information contained in genotypic and phenotypic data for the selected germplasm so that this information can generate maximum knowledge to enrich our understanding of the germplasm and genotype-phenotype relationship.

The PI initiated a workshop series: *Genomic Selection and Genome-Wide Association Studies* (GS+GWAS) at the Plant and Animal Genome Conference in 2013.

Deliverables

Publications:

Sukumaran, S., and J. Yu. 2014. Association mapping of genetic resources: Achievements and future perspectives. In R. Tuberosa et al. (ed.) Genomics of Plant Genetic Resources, 207-235.

Xu, F., *et al*. 2013. Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: A mini-review. Applied Energy 104:801–809.

Sukumaran, S., *et al*. 2012. Association mapping for grain quality in a diverse sorghum collection. The Plant Genome 5:126-135.

Li, X., *et al*. 2012. Genic and non-genic contributions to natural variation of quantitative traits in maize. Genome Research 22:2436-2444.

Wu, Y., *et al*. 2012. Presence of tannins in sorghum grains is conditioned by different natural alleles of *Tannin1*. Proceedings of National Academy of Sciences USA 109:10281-10286.

Lin, Z., *et al*. 2012. Parallel domestication of the *Shattering1* genes in cereals. Nature Genetics 44:720-724.

Li, X., *et al*. Computer simulation in plant breeding. *Advances in Agronomy* 116:217-262. .

Oral/Poster Presentations:

Parallel evolution of alleles, genes, chromosomes, and genomes. Genetic, Genomics, and Bioinformatics Program, University of California-Riverside, Oct 16, 2013, Riverside, CA.

Genic and nongenic contributions to natural variation of maize quantitative traits in maize, Plant Genomics Congress USA, Sept. 24, 2013, St. Louis, MO.

The frequency issue in current genetics and genomics analysis. Northwestern Agriculture and Forestry University, Aug. 9, 2013, Yangling, Shannxi, China.

Understanding the role of frequency in “complex” trait dissection in plants and humans. Huazhong Agricultural University, Jul. 31, 2013, Wuhan, Hubei, China.

The critical role of frequency in genetics, genomics,and breeding. T-CAP spring 2013 webinar series, Ahead of the curve: Technologies for next-gen plant breeding, May 1, 2013.

Interdisciplinary genetics analysis across plants, animals, and human. January 29, 2013. Department of Animal Sciences, Iowa State University, Ames, IA.

Genomic selection and application to grain, forage andbio-energy sorghum. December 5, 2012. Session of Genomic Selection in Corn, Sorghum and Wheat, 2012 American Seed Trait Association Meeting, Chicago, IL.

Association mapping of genetic resources: achievement and future perspectives. September 20, 2012, Symposium "Genomics of Quantitative Traits: from QTL to genes", the 56th Annual Congress of Italian Society of Agricultural Genetics, Perugia, Italy.

Opportunities and challenges of statistical genetics in genome-wide association studies. June 23, 2012, Session: Interactions Between Omics and Statistics: Analyzing High Dimensional Data, the 8th International Purdue Symposium on Statistics, West Lafayette, IN.

Genome-wide association studies in crops and comparative genomics for gene cloning. April 23, 2012, Oklahoma State University, Stillwater, OK.

Opportunities and challenges of genome-wide association studies for plant breeding. PlantBreeding and Genetics Symposium, University of Wisconsin, Oct. 28, 2011, Madison, WI.

Enhancing gene discovery and plant breeding by combining genomic technology and genetic design. Chromatin, Oct. 18, 2011, Lubbock, TX.

Li, X., et al. 2013. Genotype by environment interaction of sorghum flowering time. *Maize Genetics Conference Abstracts* 54:P276.

Lin, Z., *et al*. 2013. Parallel domestication of the Shattering1 genes in cereals. *Plant and Animal Genome XXI Conference*.

Li, X., *et al*. 2013. Genotype by environment interaction of sorghum flowering time. *Plant and Animal Genome XXI Conference*.

Wu, Y., *et al*. 2012. Natural genetic variation at Tan1 defines tannin in sorghum grain and offers seedling cold tolerance. *Plant and Animal Genome XX Conference*.

Li, X., *et al*. 2012. The pattern and dynamics of genome and chromosome across species. *Plant and Animal Genome XX Conference*.

Li, X., *et al*. 2011. Computer simulation in plant breeding. *2011 Annual ASA/CSSA/SSSA Meeting*.

Community Resources:

The selection of the 1009 biomass sorghum accessions was shared with the community. Accession name, origin, and the selection process are available upon request. This set of 300 biomass sorghum accessions is being studied in an independently funded project by the same DOE/USDA Plant Feedstocks Genomics for Bioenergy Program, led by Dr. Patrick Schnable. ""Yu Panel" is the term used in that project.

Other products/ outcomes:

Training:

Leonardo de Azevedo Peixoto, visiting student from Universidade Federal de Viçosa, genome-wide selection

Dindo Tabanao, visiting scholar from Philippine Rice Institute, diversity analysis & association mapping

Xiaoqing Yu, postdoctoral research associate, Iowa State University, genome-wide prediction

Xin Li, graduate student, Iowa State University, genome-wide selection

Chengsong Zhu, postdoctoral research associate, Iowa State University, genome-wide selection

Xianran Li, research assistant professor, Iowa State University, genomics and bioinformatics

Feng Xu, graduate students, Kansas State University biomass composition & NIR prediction

Collaborations:
Chromatin, Inc. Chromatin is developing specialized sorghum feedstocks for the renewables industry. These next generation, high-quality feedstocks are being designed by the Chromatin's team to meet the precise yield and performance requirements of the bioprocessing industry.
