Final Scientific/Technical Report

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

DE-SC0006904

PI: L. Bartley, University of Oklahoma
CoPI: Y. Wu, Oklahoma State University
Collaborators: L. Zhu, Oklahoma State University,
E. C. Brummer and M. Saha, Noble Foundation

May 31, 2016

Prepared by:
L. Bartley
Department of Microbiology and Plant Biology
University of Oklahoma

With contributions from:
Kangmei Zhao, University of Oklahoma
Chengcheng Zhang, University of Oklahoma
Matthew Peck, University of Oklahoma
Sandra Thibivilliers, University of Oklahoma
Aruna Nandety, University of Oklahoma
Prasenjit Saha, University of Oklahoma
Malay Saha, Samuel Roberts Noble Foundation
E. Charles Brummer, University of California, Davis
Ananta Achayra, University of Georgia
Yanqi Wu, Oklahoma State University
Samuel Hazen, University of Massachusetts, Amherst
Pubudu Handukumbura, University of Massachusetts, Amherst
John Ralph, Great Lakes Bioenergy Research Center
Steve Karlen, , Great Lakes Bioenergy Research Center
Lan Zhu, Oklahoma State University
Xuesong Li, Oklahoma State University
Alex Tsai, Joint BioEnergy Institute
Henrik Scheller, Joint BioEnergy Institute

Table of Contents.

This report does not contain patentable information.

1. Executive Summary

Inefficient conversion of biomass to biofuels is one of the main barriers for biofuel production from such materials. Approximately half of polysaccharides in biomass remain unused by typical biochemical conversion methods. Conversion efficiency is influenced by the composition and structure of cell walls of biomass. Grasses such as wheat, maize, and rice, as well as dedicated perennial bioenergy crops, like switchgrass, make up ~55% of biomass that can be produced in the United States. Grass cell walls have a different composition and patterning compared with dicotyledonous plants, including the well-studied model plant, Arabidopsis. This project identified genetic determinants of cell wall composition in grasses using both naturally occurring genetic variation of switchgrass and gene network reconstruction and functional assays in rice. In addition, the project linked functional data in rice and other species to switchgrass improvement efforts through curation of the most abundant class of regulators in the switchgrass genome.

Characterizing natural diversity of switchgrass for variation in cell wall composition and properties, also known as quality, provides an unbiased avenue for identifying biologically viable diversity in switchgrass cell walls. To characterizing natural diversity, this project generated cell wall composition and enzymatic deconstruction data for ~450 genotypes of the Switchgrass Southern Association Collection (SSAC), a diverse collection composed of 36 switchgrass accessions from the southern U.S. distribution of switchgrass. Comparing these data with other measures of cell wall quality for the same samples demonstrated the complementary nature of the diverse characterization platforms now being used for biomass characterization. Association of the composition data with ~3.2K single nucleotide variant markers identified six significant single nucleotide variant markers co-associated with digestibility and another compositional trait. These markers might be used to select switchgrass genotypes with improved composition in breeding programs for biofuel and forage production. Because the SSAC continues to be characterized by collaborators in the bioenergy community, the data generated will be used to identify additional markers in higher resolution genotyping data to approach identifying the genes and alleles that cause natural variation in switchgrass cell wall quality. For example, these markers can be surveyed in the 2100-member Oklahoma Southern and Northern Lowland switchgrass collections that this project also characterized.

An orthogonal approach to biodiversity studies, using comparative functional genomics permits systematic querying of how much regulatory information is likely to be transferable from dicots to grasses and use of accumulated functional genomics resources for better-characterized grass species, such as rice, itself a biomass source in global agriculture and in certain regions. The project generated and tested a number of specific hypotheses regarding cell wall transcription factors and enzymes of grasses. To aid identification of cell wall regulators, the project assembled a novel, high-depth and -quality gene association network using a general linearized model scoring system to combine rice gene network data. Using known or putative orthologs of Arabidopsis cell wall biosynthesis genes and regulators, the project pulled from this network a cell wall sub-network that includes 96 transcription factors. Reverse genetics of a co-ortholog of the Arabidopsis MYB61 transcription factor in rice revealed that this regulatory node has evolved the ability to regulate grass-specific cell wall synthesis enzymes. A transcription factor with such activity has not been previously characterized to our knowledge, representing a major conclusion of this work. Changes in gene expression in a protoplast-based assay demonstrated positive or negative roles in cell wall regulation for eleven other transcription factors from the rice gene network. Eight of fifteen (53%) of these have not previously been examined for this function. Some of these may represent novel

grass-diverged cell wall regulators, while others are likely to have this function across angiosperms. A parallel effort of this project to expand knowledge of enzymes that have evolved to function in grass cell wall synthesis, revealed that a grass-diverged enzyme in rice, OsAT5, ferulates monolignols that are naturally incorporated into grass cell walls. This finding opens potential natural selection avenues for improving biomass composition for downstream processing by weak base pretreatment. Thus, this project has significantly expanded knowledge of cell wall synthesis and regulation in rice, information that can be used in reverse genetics and synthetic biology approaches to re-engineer cell walls for improved production of biofuel and high-value products.

To lay the foundation for translating these results directly for switchgrass improvement, the project employed a comparative phylogenetic analysis of the major group of cell wall transcription factors that have been found to function in cell wall regulation, the R2R3 MYBs. This analysis concluded that known cell wall regulators are largely conserved across switchgrass, rice, maize, poplar, and Arabidopsis. This interpretation is also largely consistent with the gene network analysis described above, though both approaches provide evidence that some co-orthologs of Arabidopsis regulators have diminished or increased in importance based on gene expression patterns. Also, several clades containing dicot cell wall regulators have expanded, consistent with the evolution of new cell wall regulators. This latter result is supported by functional analysis of the R2R3 MYB protein SWAM1 in a collaboration between this project and the DOE-funded group of Dr. S. Hazen at the University of Massachusettes. The curation of the switchgrass genome through this project provides specific targets for future engineering of switchgrass cell wall regulation and may also facilitate identification of regulators that underlie the molecular markers that are genetically linked to differences in cell wall quality.

With the goal of spurring further research and technological developments in lignocellulosic biofuel production, this work has been communicated to the bioenergy and cell wall communities though various presentations and publications. To date, three manuscripts have been published, two others are near to publication, three others are in an advanced state, and two to four more are likely to be written based on analyses still in progress. In addition, project participants have presented thirteen posters and talks at regional, national, and international meetings about aspects of this project. In sum, the work supported by this funding has made and communicated significant progress in identifying the genes that grasses use for cell wall synthesis and regulation, information that will be used by project participants and others to improve the efficiency of conversion of lignocellulosic biomass to biofuels.

2. Comparison of the actual accomplishments with the goals and objectives of the project.

References produced by this project are cited, with citations listed in section 4 of this report.

Objective 1 Proposed:
Select candidate phenylpropanoid regulatory genes primarily through gene network analysis and identify single nucleotide polymorphisms in switchgrass in those genes in the Switchgrass Southern Association Collection (**SSAC**).

Objective 1 Status: Fully completed along with additional related results.

- We completed an initial list of rice secondary cell wall synthesis genes and orthologs of known Arabidopsis regulatory genes. That list was used to mine a novel, deep, high-quality rice

functional gene network to identifying additional putative regulators and associated genes (Zhao et al. In Prep).

- In addition, with partial support from this grant, we conducted reverse genetics analysis of a group of grass-diverged acyltransferases to better define the grass genes that act in phenylpropanoid cell wall metabolism. Our work has contributed to an expansion of understanding of phenylpropanoid modifications to grass cell walls (Karlen et al. Under Revision). In some cases, this work has provided insight into novel components of grass cell walls that are important for grass cell wall recalcitrance to biological conversion to biofuels.
- We conducted in depth curation of R2R3 MYB proteins of switchgrass v1.1 genome relative to rice and Arabidopsis (Zhao and Bartley 2014). Most known secondary cell wall regulators belong to this family. Several other important functions, such as abiotic stress tolerance, are also regulated by MYBs.
- We and collaborators at the Noble Foundation and the BioEnergy Research Centers applied reduced representation genotyping by sequencing and exome capture sequencing to the SSAC.
- We measured phenylpropanoid gene expression in important gentoypes of switchgrass and found that both lowland and upland ecotypes use similar predominant lignin biosynthesis genes and that cell wall gene expression precedes cell wall accumulation (Saha et al. In Prep). We have also measured variation in gene expression in divergent genotypes of switchgrass to improve understanding of the function of loci and their homoeologs and alleles to compare with mapping data and to identify potential novel regulators (Zhang et al. In Prep). These data are also being used to improve the annotation of the switchgrass genome.

Objective 2 Proposed:

- Measure the association between cell wall quality traits and polymorphisms in plants from the SSAC.

Objective 2 Status: Completed. Aspects are being redone with better data that has recently become available.

- We measured cell wall quality in terms of HCAs, enzymatic digestibility, and for a subset of samples, sugars.
- To determine the relationships between different measures of cell wall content and quality, we conducted correlation with the various directly measured data and the predicted composition based on near infrared spectroscopy of SSAC samples.
- We conducted modeling of the power of a mixed linear model to identify associations in given the switchgrass resequencing data (Nandety et al. In Prep).
- We used association analysis with a mixed linear model to associate significant alleles with cell wall composition traits (Nandety et al. In Prep).
- Higher resolution association analysis is still underway with exome capture data.

Objective 3 Proposed:

- Validate significant sequence variant-phenotype associations through targeted analysis of two additional, independent switchgrass collections (Oklahoma northern and southern lowland collections, ONLC and OSLC)

Objective 3 Status: Partially completed. Full completion pending generation of higher resolution molecular associations.

- We harvested and measured cell wall composition of ONLC and OSLC switchgrass with near infrared spectroscopy.

- We have harvested tissue for genotyping and established a rapid method for DNA preparation.
- Testing of significant sequence variants awaits higher resolution association data, which will be collected through the analysis of the exome capture data for the SSAC.

Objective 4 Proposed:
- Test the hypothesis that altered expression of genes associated with extreme wall qualities are "functional markers" that have direct molecular roles in cell wall biosynthesis regulation through reverse genetics.

Objective 4 Status: Completed using a network approach to identify candidate regulators.
- We screened mutant lines for altered expression of candidate transcription factors and protein kinases associated with phenylpropanoid biosynthesis in the rice network. We have confirmed the function of MYB61A in regulating both conserved and grass-evolved cell wall biosynthesis genes (Zhao et al., In Prep).
- These funds have also supported our contribution to the work of Dr. Sam Hazen (U. Mass) to functionally characterize the novel "Secondary Wall Associated MYB 1" gene (Handankumbura et al. Under Revision).
- We found that MYB-L1 may enhance stem growth.
- Transient assay analysis of other transcription factors reinforces some similarities between dicot and grass regulatory pathways and provides experimental evidence for 8 additional transcriptional regulators of cell wall genes in grasses.

3. Summary of Project Activities.

The research goal of this project was to identify regulatory genes that control cell wall composition and properties in grasses, especially switchgrass (*Panicum virgatum*). Grass biomass composes over half of the feedstock available in the US for biofuel production. Cell wall content is a key determinant of the efficiency of production of biofuels via biochemical conversion. Lignin and other phenylpropanoids within secondary cell walls function as crosslinking molecules in cell walls preventing enzymatic saccharification of cell wall polysaccharides. Grasses possess different cell wall composition compared with dicots and different patterns of cell wall deposition. The strategy for identifying regulators in this project combined comparative genomics (Objective 1), examining switchgrass genetic diversity through association genetics (Objectives 2 and 3), and reverse genetics and other functional assays (Objective 4).

Objective 1 Proposed:
Select candidate phenylpropanoid regulatory genes primarily through gene network analysis and identify single nucleotide polymorphisms in switchgrass in those genes in the Switchgrass Southern Association Collection (**SSAC**).

Objective 1 Status: Fully completed along with additional related results that improve understanding of grass gene synthesis and comparative genomics with switchgrass.

Overview. One of two main efforts of this objective was to create a novel high-quality grass gene network and to use comparative genomics to identify grass genes likely to function in secondary cell wall regulation within that network and the switchgrass genome. Only recently is the genome for the bioenergy crop, switchgrass, achieving relative completeness and switchgrass transcriptomics data have not been deeply sampled for building genomic networks. Thus, the project used rice as a

diploid reference and bridge species from diploid Arabidopsis to polyploid switchgrass (Obj 1A). This analysis included a detailed examination of the R2R3 MYB transcription factor family, the most well-represented family among known secondary cell wall regulators (Obj 1B).

As we proceeded in the effort to gather a list of all known cell wall-related enzymes and regulators in grasses, we recognized that the catalog of grass genes important for phenylpropanoid-mediated crosslinking was incomplete, saliently missing the gene that functions in incorporation of the phenolic acid, ferulic acid into grass cell walls. As ferulic acid is one of the distinguishing characteristics of grass and other recently evolved cell walls and has been correlated with cell wall recalcitrance, with partial support from this grant, we also undertook a reverse genetics study of an enzyme that preliminary analysis had suggested incorporates ferulic acid into cells walls (Obj 1C).

The other major goal of this objective as proposed, was to use gene capture technology to catalog the sequence variation within the Switchgrass Southern Association Collection (**SSAC**) specifically targeting regulatory genes and enzymes identified through the network analysis mentioned above. However, soon after commencing this Feedstocks project, my collaborators M. Saha and E.C. Brummer, who possess the SSAC, secured support to conduct sequencing of the whole exon-containing portion of the current switchgrass genome, i.e., exome capture GBS. In addition, Dr. Brummer also undertook a low depth (~0.35X) double restriction enzyme digest GBS of the SSAC (Obj 1D). Most of the genomic regions that this project would have examined were encapsulated within the exome capture sequences by the GBS. Thus, in consultation with my DOE-BER program manager, Dr. Cathy Ronning, we decided to direct the resources that would have been spent in genotyping to transcriptomics. This allowed us to improve the transcriptomics data availability for switchgrass and address the hypothesis that genes with major effects on phenylpropanoid biosynthesis and enzymatic recalcitrance would be more highly expressed in recalcitrant genotypes than easily digested ones (Obj. 1E).

The key results and the significance of the research conducted under Objective 1 are described below.

Objective 1A. Network analysis to identify cell wall synthesis regulators in the reference grass rice.

Hypothesis: Using a low-depth, but high-quality rice functional gene network to calibrate deeper, but poorly characterized gene networks can improve the number of high quality inference related to grass-sepcific biology including regulation of cell wall synthesis.

This subobjective explored the regulation of grass cell wall biosynthesis pathways using rice as a model and examining the conservation and divergence of the regulatory machinery between dicots and grasses using network inference and comparative genomics. To address limitations of depth or quality of available rice gene networks, we developed a generalized linear model-based scoring system to construct the Rice Combined mutual Ranked (RCR) network with data from three publicly available rice coexpression and functional genome-scale networks, namely, ROAD, PlaNet and RiceNet. The RCR network covers ~90% of the genome and to our knowledge, it is the most complete rice network developed so far (**Fig. 1A**). The RCR network also appears to be of high quality and is able to recall 92% of known interactions between transcription factors and cell wall synthesis genes (**Fig. 1B**). We used an initial list of rice secondary cell wall synthesis genes and orthologs of known Arabidopsis regulatory genes to mine the RCR for likely regulators. From this, we identified 96 additional putative cell wall-associated transcription factors from 19 protein families. Functional

analysis of a subset of these genes is described under Obj 4. Along with those data, this work is being prepared for submission to a high profile journal in summer of 2016 (Zhao et al. In Prep).

Objective 1B. Phylogenetic analysis of gene families important to the regulator of Arabidopsis cell wall synthesis to facilitate comparative genomics among species.

Hypothesis: Secondary cell wall regulatory pathways are conserved among dicots and grasses.

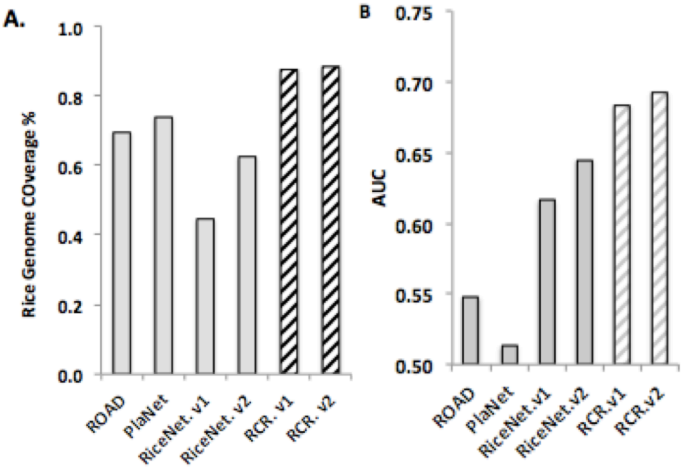

Fig 1. The Rice Combined mutual Ranked (RCR) networks (v1 and v2) have improved features over the original networks in terms of **A.** Genome coverage and **B.** The area under the curve (AUC) of a receiver-operator relationship that compares the network predictions to gene ontology similarity. Zhao et al. In Prep.

Our analysis of R2R3 MYB transcription factors included the first detailed analysis of the Joint Genome Institute (JGI) switchgrass genome annotation (v0). We applied phylogenic, OrthoMCL, and sequence identity analyses to classify the R2R3 MYB family proteins from the annotated proteomes of Arabidposis, poplar, rice, maize and the initial genome (v0.0) and transcriptome of switchgrass (*Panicum virgatum*, Pv). We find that the R2R3 MYB proteins of the five species fall into 48 subgroups, including three dicot-specific, six grass-specific, and two panicoid grass-expanded subgroups. We observed four classes of phylogenetic relationships within the subgroups of known SCW-regulating MYB proteins between Arabidopsis and rice, ranging from likely one-to-one orthology (for AtMYB26, AtMYB103, AtMYB69) to no homologs identifiable (for AtMYB75). Due to genomic co-linearity among grasses as an allotetraploid, we expect switchgrass to possess roughly twice as many gene loci as rice, with each rice locus represented on both switchgrass homoeologous chromosome pairs. As summarized in **Table 1**, out of 11 SCW regulatory clades, we found that 2 are true to this expectation, 2 are missing a homoeolog, and 7 possess more than the expected gene complement (Zhao and Bartley, 2014). Indeed, during the analysis we corresponded with JGI regarding unexpected aspects of the data and helped J. Schmutz and his team improve their annotation pipeline, including collapsing likely splice variants to single loci. We concluded that the v0 switchgrass genome annotation is largely complete, though likely missing or containing expansions of homoeologous loci. We also found evidence that some putative switchgrass SCW regulators possess different patterns of expression, suggesting that suggests subfunctionalization of putative switchgrass secondary cell wall regulators and that some are better candidates for SCW control than others (Zhao and Bartley, 2014). This work provides an initial road map for translating data regarding SCW regulation in other taxa to improving switchgrass. We continue to work with JGI regarding the gene content and ordering of the switchgrass genome and our input will be included in a genome paper expected for submission in 2016. One challenge that we have become aware of through this is that the various methodologies for assigning "orthology," an underlying assumption of comparative genomics, give different results, highlighting the need for further research and/or consensus building on this topic.

Table 1. Summary of R2R3 MYB transcription factors from bioenergy-relevant species the paralogs or orthologs of which in Arabidopsis function in secondary cell wall regulation. Proteins highlighted with the same color have ≥99% sequence similarity and are likely allelic to each other. (From Zhao et al. 2014)

Arabidopsis	Poplar	Rice[a]	Maize	Switchgrass
AtMYB26	POPTR_0001s20370	Os01g51260	GRMZM2G088783	AP13ISTG69224
AtMYB103	POPTR_0003s13190	Os08g05520	GRMZM2G325907	AP13CTG15561
	POPTR_0001s09810			AP13ISTG58495
AtMYB69	POPTR_0007s04140	Os11g10130	GRMZM5G803355	Pavirv00031864m
	POPTR_0005s06410			Pavirv00029353m
				Pavirv00020802m
AtMYB46	PtrMYB3	OsMYB46	ZmMYB46	AP13ISTG55479
AtMYB83	PtrMYB20			AP13ISTG55477
	POPTR_0009s05860			
	POPTR_0001s26590			
AtMYB20	POPTR_0004s08480	Os08g33150	GRMZM2G169356	Pavirv00023587m
AtMYB43	POPTR_0017s02850	Os09g23620	GRMZM2G126566	AP13ISTG57686
			GRMZM2G169356	Pavirv00069978m
				Pavirv00053167m
				KanlCTG16207
				AP13ISTG67468
				Pavirv00023586m
				Pavirv00051815m
				Pavirv00011866m
AtMYB42	POPTR_0003s11360	Os09g36250	GRMZM2G138427	AP13ISTG65795
AtMYB85	POPTR_0001s07830		GRMZM2G037650	AP13CTG22878
	POPTR_0015s14600		GRMZM2G104551	AP13CTG08064
	POPTR_0012s14540			
AtMYB52	POPTR_0012s03650	Os03g51110	GRMZM2G455869	AP13ISTG34280
AtMYB54	POPTR_0015s05130		GRMZM2G077147	AP13ISTG43780
	POPTR_0017s04890			Pavirv00048592m
	POPTR_0007s01430			Pavirv00048591m
				Pavirv00005610m
AtMYB58	POPTR_0007s08190	Os02g46780	GRMZM5G833253	AP13ISTG56055
AtMYB63	POPTR_0005s09930	Os04g50770	GRMZM2G097636	Pavirv00047040m
			GRMZM2G097638	Pavirv00055045m
			GRMZM2G038722	Pavirv00019950m
				AP13ISTG56056
				Pavirv00053415m
AtMYB61	POPTR_0005s00340	Os05g04820	GRMZM2G127490	AP13CTG04029
	POPTR_0013s00290	Os01g18240	GRMZM2G171781	Pavirv00042495m
	POPTR_0002s18700		GRMZM2G017520	Pavirv00021467m

	POPTR_0014s10680			Pavirv00035679m
				Pavirv00041312m
AtMYB4	POPTR_0005s11410	Os09g36730	GRMZM2G000818	AP13ISTG73550
AtMYB32	POPTR_0009s13640	Os08g43550	ZmMYB31	AP13ISTG73836
	POPTR_004s18020		GRMZM2G084583	PvMYB4.a ~ e
			ZmMYB42	AP13ISTG65360
				AP13ISTG63786

[a] MSU v7 locus identifiers, all are preceded by LOC_.

Objective 1C. Expand the catalog of phenylpropanoid-modifying enzymes by identifying the enzymes that incorporate ferulic acid into the cell walls of grasses and other commelinid monocots.

Hypothesis: Rice enzyme OsAT5 and orthologs in other species contribute to the incorporation of ferulic acid esters into grass cell wall polymers.

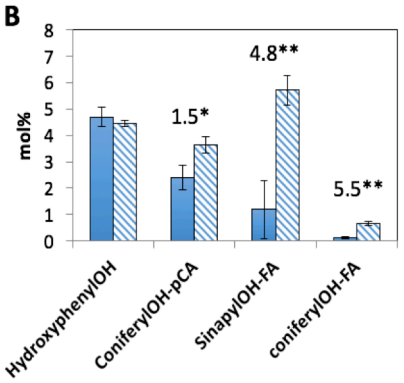

Fig 2. Increased lignin-FA conjugates in *Ubi$_{pro}$-AT5* overexpression mutant-3 (hatched) and negative segregant (solid). Error bars represent 2×SE of 3 biological replicates. * $p<0.05$, ** $p<0.01$ via Student's t-test. Karlen et al. Submitted.

Grasses, and other recently evolved monocot species, esterify the hydroxycinnamic acids, ferulic acid (FA) and *p*-coumaric acid (*p*CA), to lignin and to the matrix polysaccharide, arabinoxylan. FA undergoes radical-oxygen mediated coupling to form dehydro-dimers (diferulates, diFA) and covalently links cell wall polymers. Work partly supported by this project followed up on the observation that over expression of a so-called BAHD acyl-coA acyltransferase, *OsAT5*, increased ferulate content of rice mutant cell walls (Bartley *et al.* 2013. *Plant Physiology* **161**:1615-1633). Surprisingly, detailed characterization that we conducted in collaboration with Dr. J. Ralph (GL-BRC) revealed that *OsAT5*-over expression increases ferulate-lignin conjugates in rice, with no effect on polysaccharide conjugates (**Fig. 2**). Consistent with this, feeding yeast expressing OsAT5 and a 4-coumaryl ligase with phenylpropanoids causes accumulation of FA-monolignol conjugates. Remarkably, feruloylated monolignols have not been previously observed in wild-type grasses, though Ralph and colleagues recently introduced ferulate monolignol transferase (FMT) activity into poplar with a distantly related dicot enzyme from a medicinal herb (18% identity, 30% similarity). The FMT introduces mild base-labile ester bonds into lignin resulting in improved poplar biomass susceptibility to cellulases. Though lignin still functions in grass cell wall integrity, the FMT activity of OsAT5 may explain recent evidence that grass lignin is relatively labile and less important in grass wall recalcitrance to breakdown compared to dicot lignin. Much of these data are part of a manuscript under consideration for publication in Science Advances (Karlen et al. Submitted). Our work expands understanding of phenylpropanoid modifications to grass cell walls. In addition to being a possible avenue for breeding of grasses with less recalcitrant cell walls, this provides an avenue for controlling the degrees of polymerization of lignin for production of high value chemical products again via breeding or genetic engineering.

Objective 1D. Cataloging sequence variation in the Switchgrass Southern Association Collection.

Collaborators Saha, Brummer, Buell, and Kaeppler at the Noble Foundation and/or the BioEnergy Research Centers applied reduced-representational genotyping by sequencing (GBS) and exome capture sequencing to the SSAC. The reduced representational GBS used two restriction enzymes, so as to limit library complexity. At 5M reads per genotype (equivalent to 0.35X per haploid genome), that this approach yielded 3.2K SNPs with 5X coverage, speaks to the effectiveness of the double digest method for reducing missing data in low-depth GBS. Phylogenetic reconstruction based on the sequence data revealed that the SSAC consists of the upland and lowland ecotypes and that the lowland ecotype can be further subdivided into two populations that roughly correspond to the southeastern and Great Plains regions (**Fig. 3**). The Brummer group and my group have both used this relatively low coverage marker data for association analysis, as described under Objective 2.

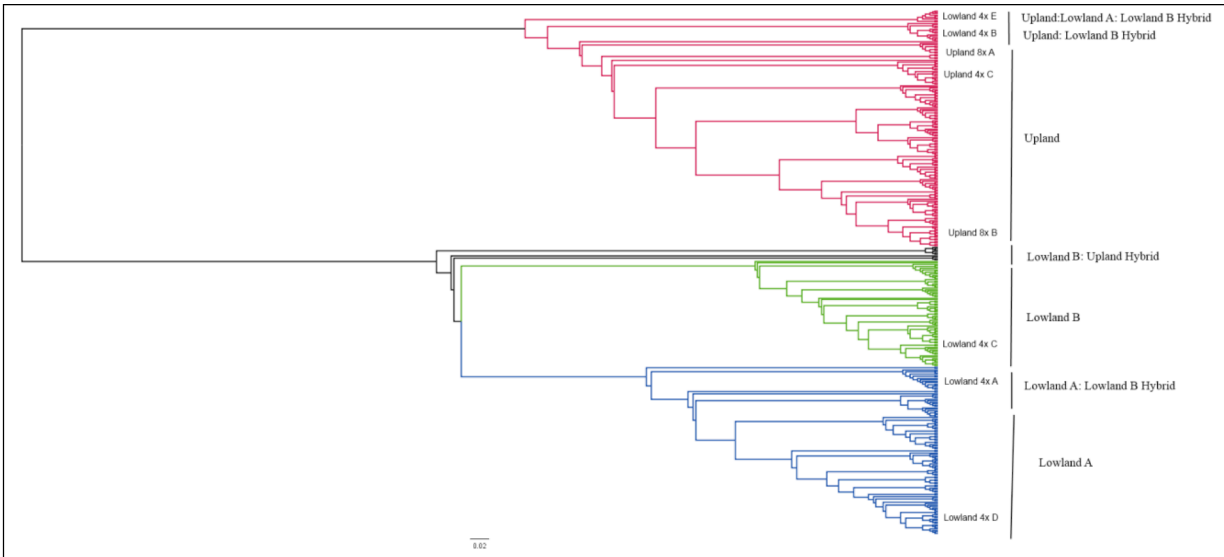

Fig. 3. The neighbor-joining phylogeny of 327 SSAC members from shared allele genetic distance with SNP markers. Blue represents southern Great Plains lowland (Lowland A), green represents Gulf Coast lowland (Lowland B), and red represents Upland. Black are hybrids of Upland and Lowland B. (From Acharya, Saha, Devos, Brummer. 2014. Genetic Diversity, Population Structure, and Association Mapping of Biofuel Traits in Southern Switchgrass Germplasm. Ph.D. Dissertation.)

In addition, the SSAC has also been characterized by exome capture GBS, though these data took approximately four years to generate and were not available until after the no cost extension period of this project. Conceived of in Winter 2011, I did not realize at the time how long it would take to get data for the collection when I moved to abandon the genotyping that we had proposed with this project. The rationale was that the effort would have been redundant and wasteful with the full exome-capture effort that Buell and colleagues worked for two years to establish the switchgrass (Evans *et al.* 2014. *Plant J.* **79**:993-1008.) In 2015, this platform was used to generate ~1M single nucleotide variants for the SSAC, a much greater depth than the GBS conducted by Brummer and better distributed across the genome than the targeted genotyping by exome capture data that I might have generated. Still, in retrospect, though the genotyping data generated through this project would have been largely subsumed by the exome capture data, we might have moved forward more quickly if we had conducted our own analysis. As described below, we have passed our phenotype data to the BESC scientists who are working on association analysis with the exome capture data. Thus, the results will eventually be published and the goals of this aspect of the project met.

Objective 1E. Identify genes the expression of which correlates with enzymatic recalcitrance.

Hypotheses: Expression of genes that encode lignin biosynthesis enzymes correlates with cell wall recalcitrance and precedes cell wall accumulation.

If a gene is differentially expressed in genotypes with a contrasting phenotype, this can be used to help identify genes underlying quantitative trait loci for that trait. In this case, our goal was to collect gene expression data for switchgrass with contrasting enzymatic saccharification yields. However, to run this experiment, we first had to decide at what developmental stage to collect RNA. Toward this end, we ran a pilot experiment to measure expression of the major lignin biosynthesis genes of switchgrass along with cell wall quality in different important genotypes from the two major ecotypes, up- and lowland, at different developmental stages. Our results indicate that though the two ecotypes appear to have similar gene expression patterns, the magnitude of expression varies, consistent with differences in recalcitrance. In addition, we found that expression in the preceding (N-1) developmental stage correlates better with cell wall accumulation in a particular stage (N), (**Fig. 4**). This work will be submitted in Summer 2016 for publication (Saha et al. In Prep). We have also measured variation in gene expression in divergent genotypes of switchgrass to improve understanding of the function of loci and their homoeologs and alleles to compare with mapping data and to identify potential novel regulators. In addition, we have used this data as the basis for securing a DOE Community Science Project in coloration with the JGI to measure expression among developmentally matched, field grown plants of the SSAC, which will allow us to get gene expression data underlying quantitative trait loci for genotypes with different alleles to help narrow down candidate genes controlling various traits, including cell wall composition and recalcitrance.

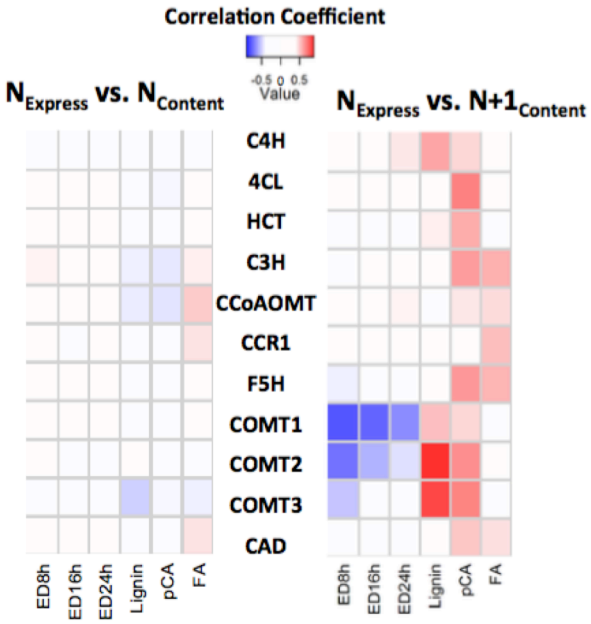

Fig. 4. Gini correlation analysis between expression of lignin biosynthesis genes (rows) and cell wall components or enzymatic digestibility (ED, columns). **Right Panel.** Expression in a developmental stage (N) correlates better with cell wall accumulation in the subsequent stage (N+1) than **Left Panel.** expression and composition in the same stage. (Saha et al. In prep.)

Objective 2 Proposed: Map associations between cell wall quality traits and polymorphisms in plants from the SSAC.

Objective 2 Status: Fully completed along with a comparison among different cell wall quality trait measurements used within the bioenergy field. Association analysis is now being redone with higher resolution exome-capture data.

Hypothesis: Cell wall quality variation among genotypes of the SSAC can be associated with genetic variation of underlying genes.

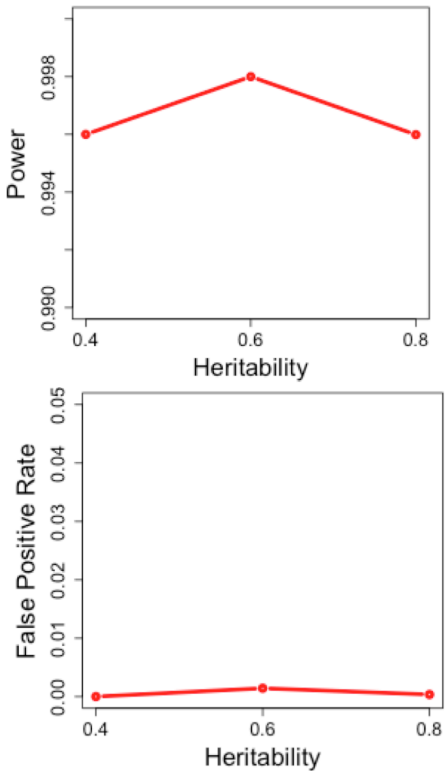

Fig 5. Simulation of the ability of a mixed linear model to detect an in vitro dry matter digestibility quantitative trait locus (QTL) with different levels of heritability (i.e., fraction of trait determined by genetics) indicates that the model functions with very high sensitivity (power) and specificity (low false positive rate). Nandety et al. In prep.

This work represents an initial use of genome wide association analysis to identify molecular markers associated with improved compositional and saccharification properties and examines the relationships among various measures of biomass quality. The switchgrass southern association collection (SSAC) examined here consists of 372 genotypes belonging to 36 accessions, originating from the southern United States. As mentioned above, double-digest reduced representational genotyping-by-sequencing of the collection identified 3,194 single nucleotide variants with 20% imputation of missing data (i.e., ~1 SNP per 39 kb in the partially assembled v1.1 genome). We used a mixed linear model to associate this dataset and biomass traits with sequence variants. Simulations with the data and traits with different heritability, indicated a good power and low false positive rate for the study design (**Fig. 5**).

For the SSAC in a grown in 2010 in Atherton, GA with triplicate clones per genotype, we measured acetyl bromide soluble lignin, ferulic and *p*-coumaric acid content and enzymatically digested pentose and glucose after mild alkaline pretreatment (Glc.ED.alk, **Fig. 6**). We also mapped the forage trait, in vitro dry matter digestibility (IVDMD) as estimated by near infrared spectroscopy, for biomass collected in three years and two locations. As co-association boosts confidence in this situation with low linkage disequilibrium and marker density, **Table 2** highlights seven markers associated with altered IVDMD and another trait. Though we believe these markers are indeed significant, their low density restricts their usefulness. Thus, higher resolution association analysis is still underway with the exome capture data and the traits measured in this study.

In addition, we examined correlations among the mapped traits and with biomass traits estimated using a switchgrass-specific NIRS bioenergy model; with lignin analysis via pyrolysis molecular beam mass spectrometry; and with enzymatic digestibility after hot acid pretreatment. Pearson's correlation coefficients revealed few strong relationships among putative "bioenergy" traits measured with various assays (**Fig 7**). The measures of lignin content, in particular, show variable agreement ($0.0 < R < 0.7$). However, we do observe numerous weaker, though still highly significant, correlations. For example, Glc.ED.alk correlates with IVDMD as predicted via either the NIRS forage or switchgrass models ($R = 0.5$; $R = 0.3$, respectively). Furthermore, Glc.ED.alk is negatively correlated with *p*-coumaric acid content, measured through chemical analysis ($R= -0.3$), with lignin content predicted by the forage NIRS equation ($R = -0.4$), and with Klassen lignin content predicted with the switchgrass NIRS equation ($R = -0.5$). Though we have used the acetylbromide solubilization (LIG.con

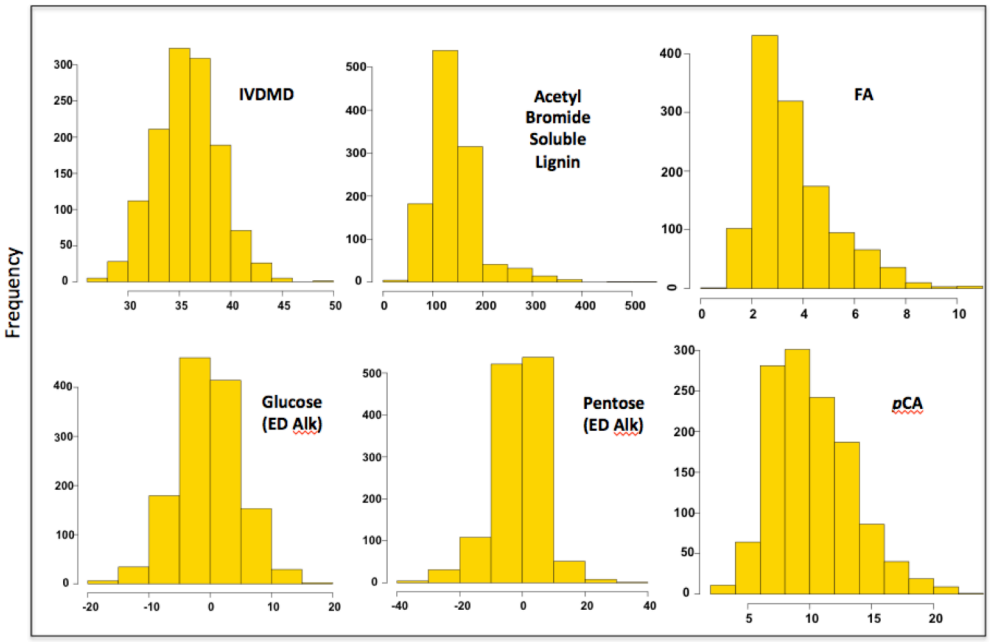

Fig. 6. Distribution of biomass traits of the Switchgrass Southern Association Collection measured through this project and in vitro dry matter digestibility (IVDMD) determined by collaborators with a near infrared spectroscopy model. ED Alk is enzymatic digestibility after mild alkaline pretreatment.

Table 2. Common significant single nucleotide polymorphism (SNPs) between Cell Wall Traits

Chromo-some	Common SNP	Position	Traits	Homozygous/heterozygous effect
2a	108816_29	7062980	IVDMD (P <0.0001) *p*CA (P <0.0001)	IVDMD: 14/21 *p*CA : -1.3/ -1.4
2a	108816_38	7062989	IVDMD (P <0.0001) *p*CA (P <0.0001)	IVDMD: 14/21 *p*CA : -1.3/ -1.4
3a	174855_57	20130494	IVDMD (P = 0.0059) Pentose (P <0.0001)	IVDMD: -1.4/ -1.2 Pentose: 5.5 / 4.4
4b	171276_43	49489964	IVDMD (P <0.0001) Glucose (P = 0.241)	IVDMD: -1.7/ -0.5 Glucose: -3 / -0.8
6b	166754_58	48176274	IVDMD Pentose (minor allele could not be counted)	IVDMD: -1 /0 Pentose: 4/0
Contig	167773_60	15306060	IVDMD (P <0.0001) FA (P <0.07)	IVDMD: 2.13/0 FA: 0.34 / 0
Contig	24667_61	10763061	IVDMD (P = 0.108) Pentose (P <0.0001)	IVDMD: 0.42/0.79 Pentose: 1.09 / -1.67

in **Fig 7**) for gathering reproducible data on lignin content of different stages of development, this measure of lignin in particular seemed to correlate poorly with others. Thus, this analysis provides an initial set of markers for mapping biomass quality traits in switchgrass. Furthermore the data reinforce that biomass quality traits possess complex relationships and that no single compositional feature seems likely to determine biofuel yield with current technologies. Moreover, different genotypes yield differently under different biofuel production conditions, such as dilute base versus hot acid pretreatment.

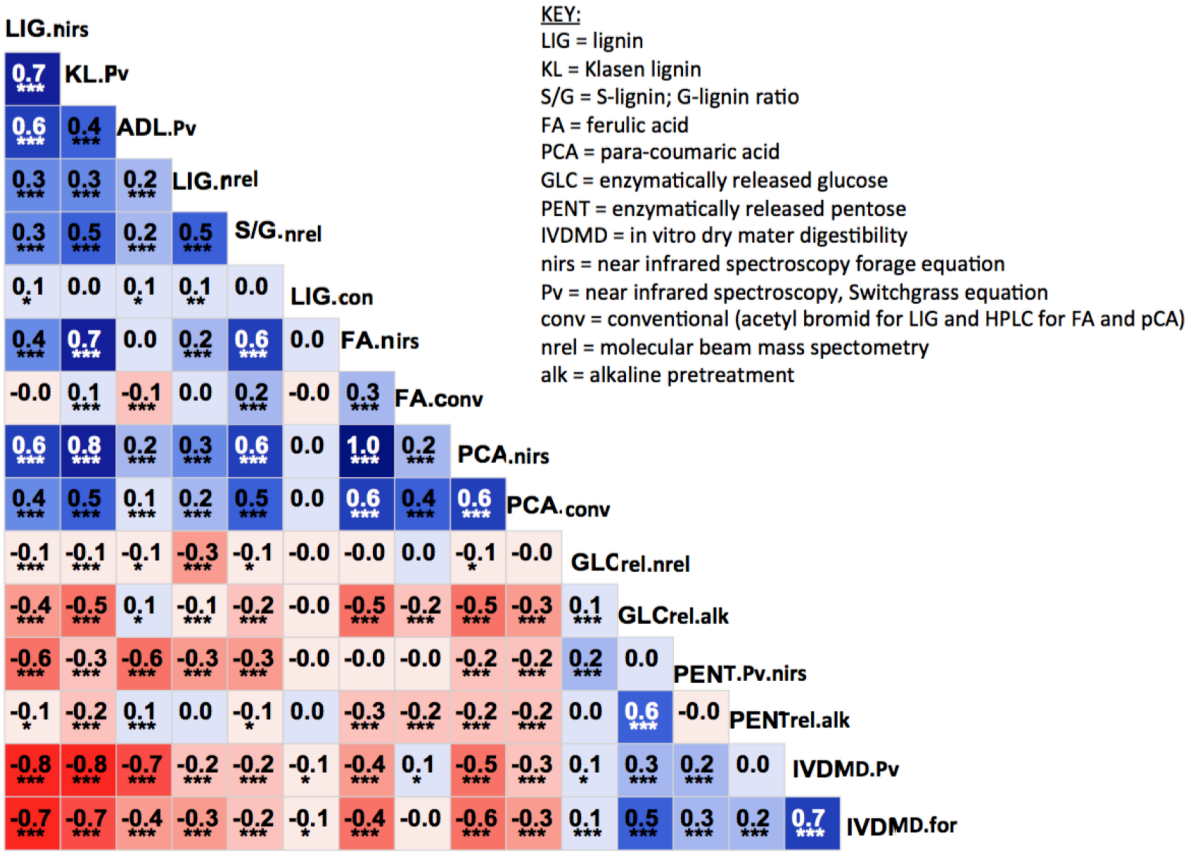

Fig. 7. Heat map of correlations among association mapping traits and other measures of similar traits with different assays. Red colors indicated negative correlations. Blue colors indicate positive correlations. * 0.01<P<0.05; ** 0.001<P<0.01; ***P<0.001

Objective 3 Proposed:
Validate significant sequence variant-phenotype associations through targeted analysis of two additional, independent switchgrass collections.

Objective 3 Status: Partially completed. Completion awaits higher confidence markers.

Hypothesis: Markers associated with alterations in cell wall quality in the SSAC are also associated with differences in cell walls in other lowland switchgrass collections.

We harvested tillers from the Oklahoma northern and southern lowland switchgrass collections and a diverse common garden of wild accessions, representing 2051 samples. We used near infrared

spectroscopy to determine the biomass composition of these genotypes (**Fig. 8**). We have also harvested tissue from these collections for genotyping and established a plate-based rapid DNA preparation method. However, due to the fact that the association analysis with the high-resolution exome-capture GBS is still underway, we have opted to hold-off on testing whether significant alleles identified in the SSAC are also associated with variation in traits in these Oklahoma collections. Once these associations have been established, our plan is to write for a small internal grant to support genotyping of these Oklahoma collections for markers with significant associations. Because these genotypes have already been selected for breeding for large stature, this avenue of research has great potential to generate superior locally adapted genotypes for biofuel production.

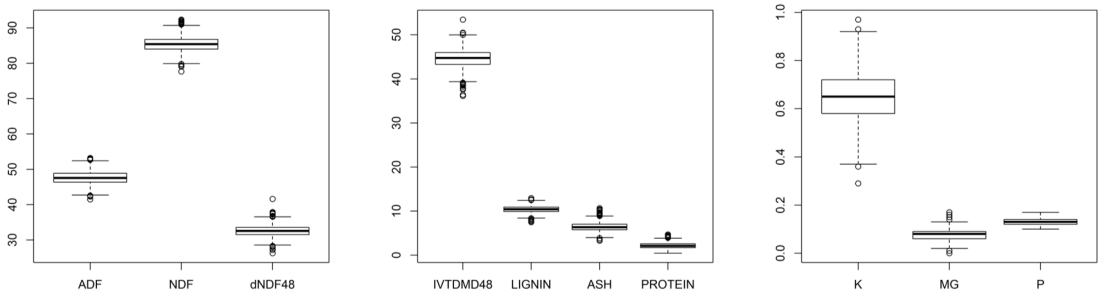

Figure 8. Compositional diversity determined by near infrared spectroscopy (NIRS) for the biomass harvested from Oklahoma State University lowland switchgrass collections at the end of the 2011 growing season. Data are shown for the standard forage equation, though we also have the switchgrass bioenergy equation predictions.

Objective 4 Proposed:
Test the hypothesis that altered expression of genes associated with extreme wall qualities are "functional markers" that have direct molecular roles in cell wall biosynthesis regulation through reverse genetics.

Objective 4 Status: Completed using a network approach to identify candidate regulators.

Hypothesis: Transcription factors with edges to several cell wall enzymes regulate the expression of cell wall genes.

Overview. The major goal of this objective was to gain functional evidence for grass gene products that regulate cell wall biosynthesis. Originally, we had hoped that this objective would serve as a follow-up to the association genetics, allowing for testing of candidate genes underlying cell wall quality and composition markers. However, due to the slow progress of the association genetics portion of the project, we decided to work with transcription factors identified through the rice gene network described under Objective 1. The network built from secondary cell wall enzymes includes 96 annotated transcription factors from 19 protein families. As technology available at the time did not allow us to screen all of them, our functional analysis focused on two classes of regulators with a large number of cell wall edges, 1) orthologs of Arabidopsis regulators connected in the network to "grass-specific" cell wall synthesis enzymes, and 2) unstudied transcription factors from families not well-associated with cell wall regulation. Our aim with this was to support comparative analyses between grasses and dicots, with the goal of applying information learned in Arabidopsis to

improving grasses and then to delineate the caveats and limitations of this approach by identifying grass-specific (or Arabidopsis absent) regulators. Functional analyses took two approaches, reverse genetics in rice and a transient gene expression assay in rice protoplasts.

Objective 4A. Genetic analysis of candidate grass cell wall transcription factors.

We characterized rice insertional and activation-tagged mutants for three transcription factors, MYB61A, MYB61B, and MYBL1, for cell wall or developmental phenotypes. These mutant analyses took advantage of the existing (and growing) collection of indexed rice mutant resources and were obtained through collaboration with researchers in South Korea. Of these, we were unable to isolate mutants for MYB61B, possibly due to low fertility of the mutant line or disruption of the primer binding sites near the T-DNA insertion.

We were able to confirm a cell wall phenotype for homozygous for knockout of a rice co-ortholog of a known dicot secondary wall transcriptional regulator, MYB61A. The knockouts exhibit a mild dwarf phenotype and have reduced expression of several secondary cell wall transcripts (**Fig 9**). Remarkably, among the genes with decreased expression in this mutant is a cellulose synthase-like F (CslF) gene family member and some of the grass-diverged acyltransferases. CSLF functions in mixed linkage glucan synthesis, a grass-specific hemicellulose. As described under objective 1, growing evidence suggests that the acyltransferases incorporate phenylpropanoid acids onto various polymers in commelinid monocot cell walls. Thus, both of these target genes are effectively absent from dicots and to our knowledge, this is the first evidence of a mechanism of regulation of grass-specific cell wall synthesis genes. In support of the gene expression result, we also find that the myb61a mutant has less mixed linkage glucan (**Fig. 9**). This analysis suggests orthologs of Arabidopsis known transcription factors tend to maintain similar functions in grasses; however, neofunctionalization appears to have occurred to add control grass cell wall specific

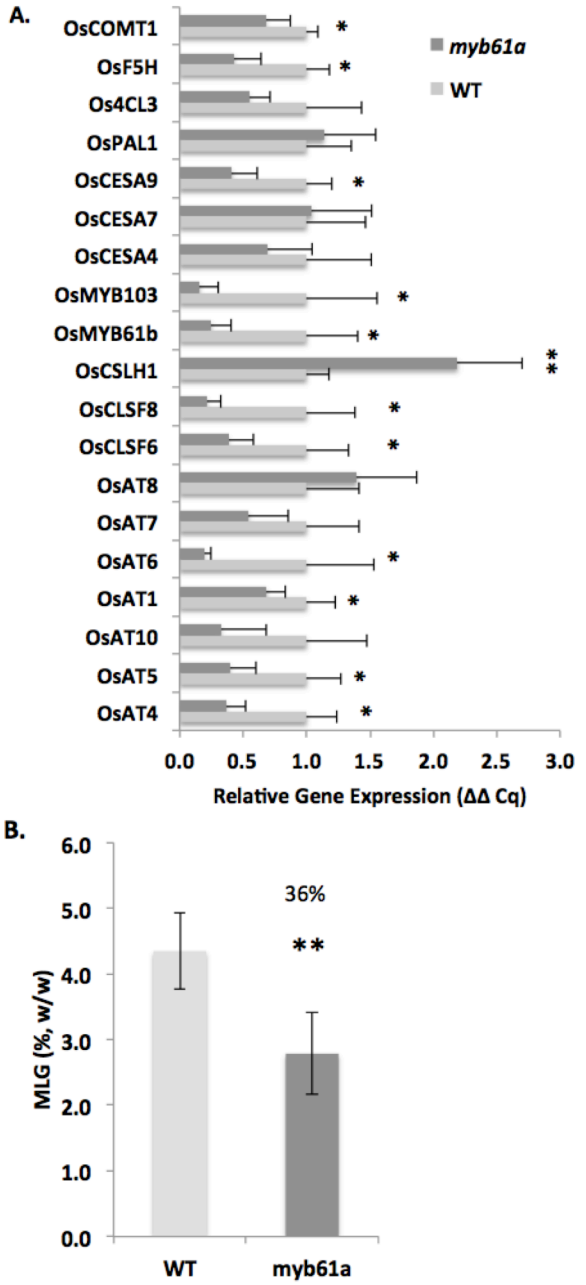

Fig. 9. Gene expression and cell wall content are altered in the rice myb61a knockout mutant compared with the wild type. **A.** qRT-PCR data show reduced expression of several cell wall enzymes and regulators, except for enhancement of CslH1, which maybe compensatory. Data are averages and standard error of four replicates. **B.** myb61a mutants have significantly less mixed linkage glucan. Data are averages and standard deviation of five biological replicates. * t-test p value < 0.05 and ** < 0.01.

enzymes, likely through recruitment of regulatory binding sites to gene promoters.

Studies of an activation tagged mutant of another regulatory gene in the network, MYB-L1, have revealed no cell wall compositional changes, but may have uncovered an enhanced stem growth phenotype (20% increase, $p<0.001$). However, over expression may not be a productive way to study this gene family, which is typically found to participate in complexes. Furthermore, the basis of the increased stature maybe enhanced cell wall thickness that does not change cell wall composition. Additional mutants are being created to understand this gene's roll in growth and/or cell wall development.

In addition, project funds supported our contribution to the work of Dr. Sam Hazen (U. Mass) to functionally characterize the novel Secondary Wall Associated MYB 1 (SWAM1) gene (Handankumbura et al. Under Revision). SWAM1 interacts with cellulose and lignin gene promoters *in vitro* and *in vivo* with preferential binding to AC-rich sequence motifs commonly found in the promoters of cell wall-related genes. *SWAM1* over-expression (*SWAM-OE*) lines had greater above-ground biomass with only a slight change in flowering time while *SWAM1* dominant repressor (*SWAM1-DR*) plants were severely dwarfed with a striking reduction in lignin of schlerenchyma fibers. Cellulose, hemicellulose, and lignin genes were significantly down-regulated in *SWAM1-DR* plants and up-regulated in *SWAM1-OE* plants. Considering lignin is inversely correlated with bioconversion efficiency phenotypes, ethanol yield was measured after culturing stems with *Clostridium phytofermentans*. There was no reduction in ethanol yield in *SWAM1-OE* lines; however, yield was significantly increased for *SWAM1-DR* samples. Phylogenetic and syntenic analyses strongly suggest that the SWAM1 clade was present in the last common ancestor between eudicots and grasses, but was lost from the Brassicaceae. Collectively, these data suggest that *SWAM1* is a transcriptional activator of secondary cell wall thickening and biomass accumulation in *B. distachyon*.

Objective 4B. Transient analysis of induction of changes in gene expression

To accelerate functional analysis of cell wall regulators, we implemented a protoplast-based transient assay to measure gene expression changes resulting from over expression of putative regulators. With this approach, we gathered evidence to support a roll in cell wall gene expression regulation of all (4 out of 4) orthologs of Arabidopsis secondary cell wall regulators and 53% (8 out of 15) of previously unstudied transcription factors (**Table 3**). Though most of these appear to be enhancers, the rice ortholog of SND2 functions as a repressor in this assay. Consistent with the whole-plant studies, MYB61A is an activator in this assay, though we obtained evidence for regulation of only a subset of genes, suggesting that some signals in this system may be false negatives. In addition, a variant of the protoplast assay with dexamethasone-induced release of the transcription factor from the plasma membrane along with translation inhibition, provides evidence that the regulation by MYB61A of grass-specific cell wall enzymes is direct. While the protoplast assay was developed in rice, it seems likely that a similar approach could be taken with switchgrass protoplasts, facilitating transfer of the results to this bioenergy crop.

Table 3. Transient gene expression results summary of tested transcription factors.
Values represent the average normalized relative gene expression ($\Delta\Delta Cq$) of each cell wall enzyme in protoplasts transformed with the target transcription factor under the control of the 35S promoter normalized to expression of with controls transformed with a similar amount of empty vector. Average and standard deviation of three replicates are shown. Shaded boxes represent results that were similar when duplicated independently.

Table 3. Part 1 of 3. Orthologs of Arabidopsis Secondary Cell Wall Regulators.

	Gene Name	OsMYB58a	OsMYB61a	OsMYB61b	OsSND2
Lignin	Os4CL3	4.69±0.30*	1.02±0.16	1.41±0.19*	NA
	OsCOMT1	5.06±0.71*	1.32±0.28	1.64±0.05*	0.81±0.12
	OsF5H	1.84±0.07*	1.61±0.23*	1.55±0.22*	0.86±0.12
	OsCCR1	1.17±0.16	1.15±0.20	1.33±0.24	0.99±0.13
	OsCAD2	2.45±0.15*	NA	NA	1.31±0.17
Secondary CESA	OsCESA4	0.85±0.06	NA	NA	0.44±0.27
	OsCESA9	0.99±0.30	1.91±0.15*	2.10±0.36*	0.32±0.02*
MLG	OsCSLF6	NA	1.57±0.13*	1.25±0.09	NA
	OsCSLH1	NA	1.06±0.10	1.11±0.24	NA
HCA	OsAT4	1.77±0.21*	1.75±0.08**	1.74±0.12*	0.31±0.17
	OsAT5	1.05±0.03	1.28±0.19	2.00±0.34*	0.33±0.08*

Table 3. Part 2 of 3. Novel Transcription Factors.

Gene Name	Os02g4151 MYB13a	Os04g43680 MYB13b	Os04g08060 C2H2	Os03g08470 AP2	Os12g43950 BLH	Os10g39030 BLH
Os4CL3	1.38±0.04*	1.69±0.12*	0.43±0.12*	1.08±0.13	0.98±0.13	1.16±0.07
OsCOMT1	1.63±0.05*	1.09±0.06	0.65±0.04*	1.00±0.24	0.85±0.08	0.91±0.08
OsF5H	2.39±0.49*	1.33±0.13	2.25±0.26*	2.04±0.28*	NA	0.80±0.12
OsCCR1	1.58±0.19	1.36±0.23	0.98±0.24	0.96±0.48	1.03±0.04	1.16±0.07
OsCAD2	0.84±0.11	1.37±0.09*	0.96±0.21	0.94±0.11	1.76±0.25*	0.76±0.13
OsCESA4	NA	NA	0.15±0.03*	1.19±0.16	1.46±0.37	0.94±0.12
OsCESA9	NA	NA	1.61±0.43	1.17±0.25	0.94±0.13	1.16±0.03
OsCSLF6	NA	NA	0.87±0.43	1.05±0.22	0.92±0.14	NA
OsCSLH1	NA	NA	1.10±0.33	1.42±0.19	1.14±0.11	NA
OsAT4	0.98±0.05	1.04±0.03	1.78±0.06*	0.95±0.09	1.20±0.36	0.77±0.20
OsAT5	1.03±0.16	0.67±0.25	1.20±0.17	NA	NA	0.76±0.12

Table 3. Part 3 of 3. Novel Transcription Factors.

Gene Name	Os01g39330 HLH	Os01g11910 HLH	Os06g43860 KNOX	Os06g46270 NAC	Os07g48550 NAC
Os4CL3	1.31±0.16	1.32±0.09	0.95±0.10	1.21±0.08	0.95±0.12
OsCOMT1	0.86±0.15	1.06±0.05	0.96±0.09	1.02±0.08	1.08±0.08
OsF5H	1.21±0.09*	1.35±0.05**	0.94±0.12	1.24±0.21	1.05±0.13
OsCCR1	1.11±0.09	1.18±0.07*	0.95±0.11	1.19±0.11	0.95±0.07
OsCAD2	NA	NA	NA	NA	NA
OsCESA4	1.34±0.12	1.30±0.14	0.87±0.08	1.02±0.29	0.95±0.12
OsCESA9	1.05±0.05	NA	0.62±0.27	1.61±0.19	1.07±0.08
OsCSLF6	NA	NA	NA	NA	NA
OsCSLH1	NA	NA	NA	NA	NA
OsAT4	1.04±0.17	1.23±0.14	NA	1.24±0.25	0.73±0.31
OsAT5	0.90±0.20	0.86±0.22	1.08±0.06	0.89±0.20	0.92±0.26

***** represents two-tail t-test p value < 0.05.

****** represents two-tail t-test p value > 0.01.

NA indicates the interactions that were not tested in this assay.

4. Products developed under the award and technology transfer activities, to date.

4.A.1. Manuscripts Published with Support from this Grant

- Zhao, K. and **Bartley, L.E.** (2014) Comparative Genomic Analysis of the R2R3 MYB Secondary Cell Wall Regulators in Arabidopsis, poplar, rice, maize, and switchgrass. *BMC Plant Biology* **14**:135. PMID: 24885077
  http://www.biomedcentral.com/1471-2229/14/135

- **Bartley, L.E.**, Xu, T., Zhang, C., Nguyen, H., Zhou, J. (2014) "Switchgrass Biomass Content, Synthesis, and Biochemical Conversion to Biofuels." Invited chapter in *Compendium of Bioenergy Plants: Switchgrass*, Ed. Y. Wu and H. Luo, Science Publishers/ Taylor and Francis/CRC Press, Boca Raton, FL, USA. pp. 109-169. **ISBN-13:** 978-1466596368

- **Bartley, L.E.**, Wu, Y., Saathoff, A., and Sarath, G. (2013) "Switchgrass Genetics and Breeding Challenges." Invited chapter in: *Biomass Crops: Breeding and Genetics*, Ed. Saha, M., Wiley-Blackwell.

4.A.2. Manuscripts Under Revision for Resubmission with Support from this Grant

- Karlen, S. D., C. Zhang, M. L. Peck, D. Padmakshan, Y. L. Tsai, R. A. Smith, K. E. Helmich, A. Eudes, H. C. A. Free, B. G. Smith, C. L. Cass, S. Lee, G. Wang, E. Baidoo, J. Keasling, P. C. Ronald, D. Loque, F. Lu, J. C. Sedbrook, R. Sibout, J. H. Grabber, T. M. Runge, K. S. Mysore, H. V. Scheller, P. J. Harris, **L. E. Bartley*** and J. Ralph*. (Under Revision, Resubmission in June 2016) Monolignol Ferulate Conjugates are Naturally Incorporated into Plant Lignins. *Science Advances.*

* co-corresponding authors

- Handakumbura, P.P., Brow, K., Whitney, I.P., Zhao, K., Sanguinet, K., Lee, S.J., Harrington, M.J., Veling, M.T., **Bartley, L.E.**, Hazen, S.P. (Under Revision, Resubmission date not set) *Brachypodium distachyon SWAM1* is a positive regulator of secondary cell wall synthesis and biofuel feedstock attributes and is not found in the Brassicaceae. *Plant Physiology*

4.A.3. In Preparation Manuscripts with Support from this Grant

- Saha, P., Lin, F. Thibivilliers, S. Santoro, N. and Bartley, L.E. (In Prep, Planned Submission: June 2016) Correlations between cell wall properties and expression of lignin biosynthesis genes in lowland and upland genotypes of switchgrass. For *Frontiers in Plant Sciences.*

- Zhao, K., Lin, F., Goh, H.J., Saha, P., An, G., Jung, K.H., and Bartley, L.E. (In Prep, Planned Submission: June 2016). A Rice Genome-Wide Network Reveals Grass Secondary Cell Wall Regulators. *for Plant Cell.*

- Nandety, A., S. Thibivilliers, X. Li, A. Achayra, F. Lin, N. Santoro, L. Zhu, M. Saha, E. C. Brummer and Bartley, L.E. (In Prep, Planned Submission: August 2016) Genome Wide Association Analysis of Switchgrass Biomass Quality Traits. For *Bioenergy Research*

4. B. Conference Presentations Related to this Project

1. American Society of Plant Biologists Annual Meeting. Austin, TX. Identification of the Regulatory Genes of the Phenylpropanoid Biosynthesis Pathway by Network Analysis of the Model Grass, Rice. July 2012. (Poster)
2. DOE Contractors Meeting. Bethesda, MD. "Switchgrass Genomics toward Improving Cell Wall Quality." January 2013. (Poster)
3. Mid-South Computational Biology and Bioinformatics Society. Stillwater, OK. Identification of Regulators in Grass Secondary Cell Wall Biosynthesis via Gene Network Analysis in Rice. March 2013. (Poster, 2nd place award)
4. American Society of Plant Biologists Annual Meeting. Providence, RI. Identification of putative secondary cell wall regulators through a comparative analysis of the R2R3 MYB family across Arabidopsis, poplar, rice, maize, and switchgrass . July 2013. (Poster)
5. DOE Contractors/USDA PI Meeting. San Diego, CA. "Switchgrass Genomics toward Improving Cell Wall Quality: Genotype Matters." January 2014. (Talk)
6. Plant and Animal Genome Meeting. Identification of Regulators in Grass Secondary Cell Wall Biosynthesis via Gene Network Analysis in Rice. January 2014. (Poster)
7. University of Illinois, Chicago. "Molecular Analysis of Grass-Diverged Cell Wall Synthesis for Biofuels and Food." February 2014. (Talk)
8. Plants and BioEnergy, University of Guelph, Canada. "Switchgrass Genomics toward Improving Cell Wall Quality: Genotype Matters." June 2014. (Talk)
9. American Society of Plant Biologists Annual Meeting. Minneapolis, MN. Identification of Transcriptional Regulators and Corresponding DNA Motifs for Secondary Cell Wall Biosynthesis in Grasses. July 2014. (Poster)
10. Oklahoma State University, Stillwater. Department of Agronomy and Soil Science. "Genetic Determinants of Cell Wall Content in Switchgrass for Improved Biofuel Production." October 2014. (Talk)
11. International Symposium for Rice Functional Genomics, Tucson, AZ. Functional Genomics Reveals Enzymes that Incorporate Hydroxycinnamates into Rice Cell Walls. November 2014. (Talk)
12. International Symposium for Rice Functional Genomics, Tucson, AZ. Exploring the Regulation of Grass Secondary Cell Wall Biosynthesis via Gene Network Analysis in Rice. November 2014. (Talk)
13. Plant & Animal Genome Meeting. Bioenergy Grasses Concurrent Session. San Diego, CA. "Genomics and Genetics of Switchgrass Cell Wall Content." January 2015. (Talk)

4.C. Networks of collaborations fostered.

This funding has been extremely important in establishing connections between the PI's group and others who study switchgrass and cell wall regulations and synthesis, to mutual benefit. These connections are and will continue to produce new knowledge and improvements to bioenergy grasses in future years. For example, these activities have given the PI a basis for leading a group of scientists from multiple groups to use the approaches applied with this funding to curate bioenergy-relevant genes in the first major version (v3) of the switchgrass genome. The aim of that group is to encourage informed utilization of the DOE-funded genome.

Collaborations that have been established with these funds are as follows:

Malay Saha and Yuhong Tang at the Noble Foundation and BESC
Yanqi Wu and Lan Zhu, Oklahoma State University
Breeanna Urbanowicz and Michael Hahn, CCRC with BESC
Nicholas Santoro and David Lowry, Michigan State University and GL-BRC
John Ralph, University of Wisconsin and GL-BRC
Tom Juenger, University of Texas
Jeremy Schmutz and others with JGI
Mark Davis and others with NREL
Sam Hazen, University of Massachusettes, Amherst

4.D. Inventions/Patent Applications, licensing agreements.

Bartley, L.E., Ronald, P.C, and Scheller, H.V. 2015. Modulation of expression of acyltransferases to modify hydroxycinnamate content. US Application No. 14/746,779.