USDA and DOE Joint Research

USDA and DOE Joint Research: Plant Feedstock Genomics for Bioenergy

Manipulation of Lignin Biosynthesis to Maximize Ethanol Production from *Populus* Feedstocks

Clint Chapple (chapple@purdue.edu), Richard Meilan, and Michael Ladisch

Purdue University, West Lafayette, Indiana

NON-TECHNICAL SUMMARY: High gasoline prices, global warming, national security, and the limitations of global petroleum resources have reinvigorated worldwide interest in renewable resources as a feedstock for liquid transportation fuels, particularly those derived from cellulose. As a perennial woody plant, hybrid poplar (genus *Populus*) offers several advantages with regard to cellulosic biofuel production including rapid growth rates, the ability to cycle nutrients, a wide geographic distribution, genetic diversity, amenability to genetic engineering, and abundant genomic resources. The phenolic cell wall polymer lignin constitutes a significant barrier to biomass conversion but, at the same time, it is essential to normal plant growth and development. Recent advances in our understanding of how lignin monomers are synthesized provide us with an opportunity to modify the content and composition of the lignin polymer. The research to be conducted will enable us to rationally assess the cost savings that could result from using genetically engineered poplar, instead of corn, as a feedstock for producing biofuels.

OBJECTIVES: 1) Generation of transgenic poplar up- or down-regulated for four enzymes known to impact lignin quantity and quality; 2) Development of metabolic profiling methods for poplar and their application to greenhouse- and field-grown wild-type and transgenic plants; 3) Morphometric analysis of transgenic lines grown in field plots; and 4) Cell wall deconstruction analysis of wild-type and lignin-modified transgenic lines.

APPROACH: Obj. 1) The expression of four enzymes in the lignin biosynthetic pathway will be up- and/or down-regulated. For each DNA construct, poplar cDNA will be synthesized from young shoot RNA using reverse transcriptase and PCR-amplified with gene-specific primers developed based on conserved regions within the genes’ sequences identified from the poplar genome, the Arabidopsis genome, and other plant sequences. All constructs will be transformed into clone NM-6 (*Populus nigra x P. maximowiczii*) using an Agrobacterium-mediated transformation protocol. Obj. 2) Transformants will be tested for changes in lignin composition by a battery of lignin analyses (i.e., Klason lignin, pyrolysis GC-MS, and DFRC analysis). At the same time, HPLC and GC-MS will be used to assay total cell extracts and cell-wall hydrolysates from these plants to determine whether perturbations in phenylpropanoid pathway gene expression have led to alterations in free and/or cell wall-esterified phenolic compounds. Obj. 3) Morphometric analyses will be conducted on all lines in the field trial to ensure the transgenes have no deleterious effects on phenotype. All plants will be visually examined at least twice during the first and second growing seasons, including at least
Systematic Modification of Monolignol Pathway Gene Expression for Improved Lignocellulose Utilization

Richard Dixon (radixon@noble.org) and Fang Chen
Samuel Roberts Noble Foundation, Ardmore, Oklahoma

NON-TECHNICAL SUMMARY: Plant cell walls are made of three types of sugar polymer, cellulose, hemicellulose and pectin, and, as the cell wall develops, these are reinforced by lignin, a polymer of phenylpropane units (monolignols) that is recalcitrant to degradation. There are two stages involved in bioethanol production from lignocellulosic biomass: hydrolysis of the cell wall polysaccharides to their component hexose and pentose sugars, derived from cellulose/hemicellulose and hemicellulose respectively, and subsequent fermentation of the sugars to ethanol. The presence of lignin reduces access of enzymes and chemicals to hemicellulose and cellulose, thus reducing the efficiency of hydrolysis.

OBJECTIVES: The objectives of this proposal are 1) to determine which features of the lignocellulosic material (lignin content, lignin composition or other factors) are most detrimental to the fermentation of biomass to ethanol and 2) to develop the crop plant alfalfa (Medicago sativa) as a model system for genomic studies on biomass utilization.

APPROACH: Obj. 1. We have already generated transgenic alfalfa lines independently down-regulated in most (ten) of the enzymatic steps believed to be required for monolignol biosynthesis. Lignin content and composition have been determined in most of these lines (and cover a broader range of values than could be found in natural populations). The chemical analyses of the lignins will be completed, and the plant materials subjected to cell wall hydrolysis (acid and enzymatic) and fermentation. Yields of released sugars and bioethanol will be measured. We can then determine which features of the lignin polymer (content, composition, linkage types, etc) are most detrimental to sugar release and fermentation during bioethanol production, and design the optimal strategy for genetic modification of the plant feedstock for biofuel processing.

Obj. 2. Using genomic approaches (DNA microarray and informatics), we will discover additional genes necessary for lignin accumulation in alfalfa. These will be evaluated by down-regulation in transgenic plants as described above We will develop approaches for non-biased discovery of genes impacting lignocellulose processing in Medicago truncatula, a model legume closely related to alfalfa, utilizing a large population of plants generated at the Noble Foundation that harbor transposon insertions. These lines will be screened for altered lignin properties by near infrared reflectance spectroscopy and simple staining procedures.
Sorghum Biomass/Feedstock Genomics Research for Bioenergy

William Rooney¹ (wlr@tamu.edu), John Mullet, Steve Kresovich, Doreen Ware, and P. Klein

¹Texas A&M University, College Station, Texas; ²Cornell University, Ithaca, New York; and ³Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

NON-TECHNICAL SUMMARY: Future production of renewable transportation fuels will require a consistent supply of biomass produced specifically for biofuel production. There will likely be many sources of biomass and species will be selected for their ecological fit, and their production and processing capability. Sorghum (Sorghum bicolor L. Moench) has the potential to be one of the species dedicated to biomass production because of its high productivity, drought tolerance, established production systems, and its genetic diversity. Research activities to develop sorghums for biomass production have been limited in the past. The purpose of this research is to use traditional and biotechnological approaches to produce sorghum genotypes with the genetic potential for use in bioenergy production.

OBJECTIVES: Within this grant period our specific objectives are to: (1) annotate genes, pathways and regulatory networks identified in the sorghum genome sequence that are important for biomass generation, and (2) identify, map and clarify the function of trait loci that modulate accumulation and quality of biomass in sorghum.

APPROACH: (Obj. 1) Genes encoding proteins involved in biochemical pathways important for biomass generation, and plant composition related to biofuel production (i.e., starch, lignin, sugar, cellulose and hemicellulose), will be identified and projected onto biochemical pathways using the database MetaCyc. The pathway projections will provide a baseline of information on sorghum genes involved in biochemical pathways related to biomass/drought tolerance thus aiding our downstream analysis of QTL and traits. Moreover, the information on sorghum biochemical pathways in Gramene can be readily compared to information on other cereals and other organisms via Gramene’s comparative mapping tools. This will help identify gaps in our current knowledge of sorghum biochemistry and help identify pathways and genes that may be useful to deploy in sorghum for biomass/bioenergy generation. (Obj. 2) The goals of objective 2 will be met in two approaches. First, grain, biomass, and carbohydrate yields will be measured in a population consisting of 175 recombinant inbred lines (RILs) (F₅:6) from the cross of BTx623 (a high yielding early flowering grain sorghum) × Rio (a high biomass sweet sorghum). Plant growth parameters will be analyzed to obtain a baseline for downstream meta-analysis. These include plant height, flowering time and tillering, traits that likely modulate carbohydrate partitioning in various tissues and total biomass. Traits that affect grain yield, biomass (i.e. the tissue harvest index and distribution of grain, stem, and leaf weight), the composition of structural and non-structural carbohydrates, and the overall energy gain of the plant will be evaluated. A genetic map of this population will be created and based on this map, QTL analysis will be carried out using QTL Cartographer, Mapmaker/QTL, or a similar analysis program. Second, tall, late flowering forage sorghum hybrids have the highest potential for total biomass generation. We are identifying genotypes that have high yield potential and excellent combining potential. In parallel, a RIL population will be developed by crossing tall, photoperiod sensitive late flowering genotypes that vary in biomass accumulation to explore the genetic basis of biomass accumulation and composition traits in this material. The population will be analyzed for variation in growth characteristics (growth rate and partitioning of growth) during the extended vegetative phase, for variation in lodging, total biomass and components of biomass.
related to biofuel production as described above for the BTx623 × Rio population. This information will build a baseline of data on biomass production in tall, late flowering sorghums that have the highest potential for biomass production.

Streamlined Method for Biomass Whole-Cell-Wall Structural Profiling

John Ralph (jralph@wisc.edu), F. Lu, B. Sundberg, and S. Mansfield

U.S. Dairy Forage Research Center, USDA-Agricultural Research Service, University of Wisconsin, Madison, Wisconsin

NON-TECHNICAL SUMMARY: In wide-ranging research aimed at altering plant cell wall characteristics, by conventional breeding or modern genetic methods, one of the biggest problems is in delineating the effects on the cell wall. Plant cell walls are a complex conglomerate of a variety of polysaccharides and lignin. Each component alone is complex, and their interactions are only poorly characterized and understood. The most common approach has been to isolate and purify components and to characterize them in detail using a variety of methods. Such studies will always be necessary. As studies in which lignin-biosynthetic-pathway enzymes were targeted have made abundantly clear, simple compositional analysis is not sufficient. Some plants with only minor compositional changes have drastically altered chemical structure that belies the important alterations that can be made in processes ranging from natural digestibility in ruminant animals to industrial chemical pulping.

How can the structural components of the cell wall be readily characterized? Although other methods have their place, and can be more rapid (e.g. NIR), the difficulty in interpretation of some spectral methods, or the destruction of structure by chemical methods, assures that key features of cell walls benefiting, for example, biomass production and conversion are lost. A promising recent approach is the dissolution of the whole cell wall and NMR analysis. We intend to extend the methodologies to provide rapid structural profiling of plant materials, aiming for a “screening rate” of 20-30 samples per day. Such methodology will be useful to plant researchers worldwide.

OBJECTIVES: To provide the plant cell wall and biomass research communities with improved methods for polysaccharide and lignin structural profiling, based on complete cell wall solubilization and NMR. The aim is to develop and streamline procedures to allow 20-30 samples per day to be profiled.

APPROACH: The following improvements to the Dissolution/NMR method will be sought: a. Provide the necessary database, via model compounds and isolated components, to characterize component polysaccharides and lignins in whole-cell–wall mixtures. b. Optimize milling conditions for the various biomass sample types and seek alternative solutions that require less rigorous milling. c. Attempt to develop improved rapid dissolution methods that can be performed directly in the NMR tube; determine solvent systems that do not interfere with the correlation contours from polysaccharide and lignin components. d. Develop NMR methods that allow the crucial HSQC NMR spectra to be acquired in under 1 hour (on the whole cell wall sample). e. Develop methods for databasing and quantifying the 2D NMR cell wall spectra. f. With collaborators, attempt to develop chemometrics methods that can be applied to 2D NMR data.

* Presenting author
Development of a Proteoglycan Chip for Plant Glycomics

Chris R. Somerville (crs@stanford.edu)
The Carnegie Institute of Washington, Washington, D.C.

NON-TECHNICAL SUMMARY: In order to develop plants for use as optimal feedstocks for biofuels production from cellulosic biomass, it will be necessary to understand how the polysaccharides that comprise the majority of plant biomass are made and deposited in cell walls. More than a thousand genes for proteins that may be involved in synthesis and assembly of plant cell walls are evident in the genomic DNA sequences of the higher plants for which whole genome sequences are available. In order to assign functions to such a large number of genes, high-throughput assay methods will be required. This project concerns the development of a novel assay method that may facilitate the assignment of function to most of the relevant proteins. The goal is to develop reagents and methods that will allow presentation of a large number of different oligosaccharide structures on a solid surface in a small area. In principle, this can be accomplished by using robots to print small amounts of material at high density on a suitably surface-modified slide (eg., a “glycochip”) in much the same way that DNA microarrays are made. The glycans presented in this way are expected to serve as acceptors for glycosyltransferases, methylases and acetytransferases, and as substrates for enzymes such as glycosidases that remove sugars from glycans. In this pilot project, we will focus on only those oligosaccharides that can be derived from plant proteoglycans containing O-linked glycans. The oligosaccharides will be prepared by using pure polysaccharide hydrolytic enzymes to fragment naturally occurring O-linked glycans which will then be purified by chromatographic methods. The glycochips produced in this way will be tested for their ability to act as acceptors in enzyme assays for glycosyltransferase enzymes from plants.

OBJECTIVES: To develop high throughput methods and reagents that will facilitate the assignment of function to large numbers of glycosyltransferases and other glycan modifying enzymes.

APPROACH: (1) Development of a series of transgenic plants that express synthetic peptides that become O-glycosylated in various ways and which have regions of sequence that facilitate purification of the glycopeptides; (2) purification of glycopeptides from transgenic plants; (3) determination of the structure of the glycans; (4) sequential cleavage of the glycans on each of the glycopeptides to produce a series of partial glycans; (5) production of glycochips by robotic spotting of the various glycopeptides onto chemically modified surfaces; (6) development of mass spectrometric methods for measuring the mass of glycopeptides in a microformat; (7) use of the glycochips to assay for glycosyltransferases activities in protein extracts from plants.

KEYWORDS: Arabinogalactan, AGP, Extensin, Hydroxyproline-rich glycoprotein, Glycosyltransferase, Glycomodule, O-linked glycan, cell wall,
Biochemical Genomics of Wood Formation: O-Acyltransferases for Alteration of Lignocellulosic Property and Enhancement of Carbon Deposition in Poplar

Chang-Jun Liu (cliu@bnl.gov)
Brookhaven National Laboratory, Upton, New York

NON-TECHNICAL SUMMARY: The goal of GTL is to use newly acquired genomic data to better understand fundamental biological processes and enhance the translation of that scientific knowledge into new technologies for energy and environmental applications. Our project is going to characterize the O-acylation reactions participated in lignocellulosic biosyntheses using poplar genomics resources. O-acylation is involved in the cell-wall component polysaccharide and lignin biosyntheses. It affects the cell-wall sugar’s solubility and the lignocellulosic bio-digestibility. O-acylation is also responsible for the structural modification and sequestration of a variety of polyphenolic metabolites required for wood formation. Understanding the mechanism of O-acylation at molecular level implicates biotechnological applications in genetic modification of lignocellulosic structures to facilitate biomass to bioethanol conversion, and in improvement of feedstock biomass production.

OBJECTIVES: 1) genome-wide identify acyl-CoA dependent acyltransferase genes from poplar genomics database; 2) systemically explore the tissue specific and stress-responsible expression patterns of O-acyltransferase genes to identify the enzymes specifically involved in lignocellulosic biosynthesis; 3) systemically characterize the biochemical functions of acyltransferases responsible for polysaccharide acetylation, lignol biosynthesis and phenolic compound modification.

APPROACH: Obj. 1. tblastn algorithm will be applied to search poplar genomics resources (P. trichocarpa V1.0, http://genome.jgi-psf.org) by using the highly conserved sequence motifs (HXXXD and DFGWG) of acy-CoA dependent acyltransferases. In order to distinguish the potential soluble and membrane bound proteins, the encoded polypeptides of the identified gene candidates (at least ~50 gene models) will be subjected to the computational topology and post-translational modification analyses by using PSORT and SignalP web services to predict the protein sorting signal, subcellular localization site and the location of signal peptide cleavage sites in amino acid sequences. Obj. 2. The transcriptional profiling of putative acyltransferases will be analyzed both by “in silico” northern, based upon the high resolution poplar EST/microarray databases, and by QRT-PCR against mRNAs from different types of tissues (leaf, shoot, root, stem etc.) and the tissue sections representing different stages of developing wood, including the early expansion, late expansion, secondary cell wall formation, and programmed cell death (sapwood/heartwood) from poplars under normal growing conditions, the drought, salt stresses, and insect damage or physical wounding. Obj. 3. The recombinant proteins of the gene candidates that are highly expressed in wood-forming tissues will be produced using either E.coli, yeast, or Drosophila Gateway expression systems. Subsequently the combinatorial in vitro assays will be conducted by using different acyl-CoA donors and potential substrates including sugars, oligosaccharides, hydrolyzed and pre-deacylated pectin and xyloglucan, lignols, and other phenolics. Products detection and identification will be performed by LC-UV-ESI-MS, HPAEC-MS and MALDI-TOF-MS analyses.
Genomic Knowledgebase for Facilitating the Use of Woody Biomass for Fuel Ethanol Production

Vincent L. Chiang (vincent_chiang@ncsu.edu)
North Carolina State University, Raleigh, North Carolina

NON-TECHNICAL SUMMARY: Situation and Problem: (A) Wood in forest trees is a major, potential lignocellulosic material for ethanol. (B) Trees can potentially be modified with a genome-wide approach for traits to overcome virtually any major biomass conversion barrier to ethanol production. (C) Gene expression and regulation of plant traits suited to ethanol production is poorly understood. Purpose: (A) This project examines, at the genome level, gene expression and regulation of lignocellulosic formation in *Populus trichocarpa*, a targeted energy tree crop. (B) The purpose of this project is to establish a knowledgebase about the possible genes and transcription factor genes involved in lignocellulosic formation and those genes that may enable effective manipulation of lignocellulosic traits to facilitate ethanol production.

OBJECTIVES: We propose a 3-year project to accomplish the following four objectives. (1) Chemical, biochemical and morphological profiling of TW development in *Populus*. (2) Oligo-microarray profiling of transgenics and TW development in *Populus*. (3) In vitro functional analysis of putative *Populus* xylan synthase genes. (4) Gene functional analysis in transgenic *P. trichocarpa*.

APPROACH: (1) Chemical, biochemical and morphological profiling of TW development in *Populus*: Wildtype and transgenics will be propagated for array characterization. For the TW system, we will profile cell wall trait changes at several different stages along the development of TW in Nisqually-1. These include cellulose, xylan, and lignin contents, lignin S/G ratios, xylan synthase activity, key lignin pathway gene transcript levels and enzyme activities, vessel/fiber ratios and TW fiber formation. These profiles reflecting changes due to preferential processes for the particular cell wall traits will guide microarray analyses to identify the involving genes and transcription factor genes and their contributions to these processes. (2) Oligo-microarray profiling of transgenics and TW development in *Populus*: RNAs from developing xylem of wildtype and selected transgenic *P. tremuloides* lines will be characterized by the updated full *Populus* transcriptome oligo-microarrays. RNAs from the developing xylem of the TW development stages with known cell wall trait/property profiles, will be characterized by the full genome microarrays. These RNAs will also be analyzed by miRNA oligo-microarrays designed with probes for detecting mature miRNAs that are mostly related to xylem development. Three biological replicates will be used in all array experiments. Differentially expressed genes will be determined and their transcript variation profiles between distinct transgenic levels or various TW developmental states will be correlated with the cell wall trait profiles to identify the proposed genes and genes encoding transcription regulators. (3) In vitro functional analysis of putative *Populus* xylan synthase genes: The array-selected and qRT-PCR confirmed putative xylan synthase genes will be expressed in our established Drosophila S2 cell system and the gene products will be characterized for biochemical functions. (4) Gene functional analysis in transgenic *P. trichocarpa*: We will select three transcription factor genes that may coordinate lignocellulosic accumulation and two miRNA genes that may regulate vessel and fiber cell development for overexpression in transgenic Nisqually-1 to determine their functions.
Genetic Dissection of the Lignocellulosic Pathway of Wheat to Improve Biomass Quality of Grasses as a Feedstock for Biofuels

Bikram Gill (bsgill@ksu.edu) and Wanlong Li
Kansas State University, Manhattan, Kansas

NON-TECHNICAL SUMMARY: As fossil fuel reserves dwindle, we are about to transition from nonrenewable energy to renewable bioenergy. Demand for ethanol is increasing steadily as an alternative fuel as well as an octane-boosting and pollution-reducing additive to gasoline. Keeping in step with the demand will require higher quantity and quality of biomass. In the Great Plains, cereal crops (wheat, corn and sorghum), as well as native grasses (switchgrass, bluestem) predominate and are important but as yet relatively untapped resources for bioenergy.

OBJECTIVES: 1) Investigate the expression of ~80 candidate genes for lignin biosynthesis, their enzymatic activities, and lignin content and composition in different organs at different stages of diploid wheat plant; 2) Silence these 80 genes individually by VIGS; 3) Identify knockout mutants of these genes using TILLING and 4) Characterize the silenced tissues and knockout mutants by metabolite profiling.

APPROACH: obj. 1. Plants will be grown at seedling stage (Feekes stages F1-2), leaf sheath/stem elongation stages F4-5 and heading stages F8-9. The leaf, sheath, stem and spike tissues will be assayed for the expression of the 80 genes by QRT-PCR to decide the developmental phase and tissue on which all experiments will be done. Obj. 2. Conserved sequences will be used to silence all 80 genes individually by VIGS and the silenced tissues will be verified by QRT-PCR and positively silenced samples will be subjected to metabolite profiling, enzymatic assays and histochemical staining to determine the consequences of the genetic block in the lignocellulosic pathway. Obj. 3. We will do TILLING to screen for mutations for genes with major effect on the lignin content based on VIGS results. The homozygous plants containing genetic lesions for lignocellulosic pathway genes will be phenotyped with regard to growth and development, lignin content and composition, and lignocellulose degradability. Obj. 4. We will determine the lignin content of wheat silenced tissues, knockout mutants and controls by the acetyl bromide method, lignin composition by thioacidolysis methods, soluble and wall-bound phenolic compounds by HPLC, and polysaccharide content and composition by the phenol-sulfuric acid method. We will also histochemically characterize these plant materials by Mauls and Wiesner staining.
Using Association Mapping to Identify Markers for Cell Wall Constituents and Biomass Yield in Alfalfa

Charles Brummer¹ (brummer@uga.edu), Kenneth J. Moore², and Jeff J. Doyle³

¹University of Georgia, Athens, Georgia; ²Iowa State University, Ames, Iowa; and ³Cornell University, Ithaca, New York

NON-TECHNICAL SUMMARY: Alfalfa (Medicago sativa) is a potential biofuel crop because it produces high yield, its leaves can be used as a high value, high protein coproduct, it fixes atmospheric nitrogen, and it has beneficial effects on the environment. Improving alfalfa as a biofuel crop will entail breeding for increased biomass yield and altered cell wall composition. While traditional phenotypic selection can be successful, the perennial nature of alfalfa requires that a selection cycle lasts for several years. Decreasing the cycle time would increase genetic gain for all traits. This could be achieved using marker assisted selection for the traits of interest, but marker identification research conducted previously has not focused on representative alfalfa breeding populations nor has it examined wild germplasm as a source of new alleles to improve agronomically important traits. Our experiment will address these issues by studying both wild germplasm not typically used in alfalfa breeding programs and also a cultivated breeding population currently under selection. We will evaluate biomass yield and cell wall composition in the field. Concurrently, we will evaluate the genotype of each plant using genetic markers selected throughout the genome. We will also develop markers based on DNA sequence variation in genes of possible involvement in cell wall synthesis. Ultimately, this project will improve the efficiency of selection for enhanced bioenergy characteristics in alfalfa, produce numerous new markers at important candidate genes, and identify potentially useful alleles in wild germplasm.

OBJECTIVES: Our objectives are to use genomics approaches to identify chromosomal regions, and ultimately genes, controlling the two most important bioenergy traits, biomass yield and composition, and to develop genetic markers that can be used directly in applied plant breeding programs to improve the bioenergy qualities of alfalfa. We will pursue two complementary objectives to attain our goals: 1. Identify loci, and specific alleles, that control the concentration of alfalfa stem cell wall constituents and that are associated with biomass production using whole genome and candidate gene association mapping across a diverse set of natural diploid alfalfa accessions, and 2. Extend the analysis and methods used in the first objective to a tetraploid alfalfa breeding population currently under selection. As a result of this project, we (a) will have identified novel alleles in wild alfalfa germplasm that may be useful to improve cultivated alfalfa; (b) will have developed and used SNP markers in genes known to be involved in the biosynthesis of cell wall composition; (c) will be able to select individuals within a breeding population on the basis of these markers, and (d) will identify new alleles from wild germplasm useful for improving cultivated alfalfa. This experiment will provide the first estimate of linkage disequilibrium (LD) in alfalfa, both in a broad cross-section of wild diploid germplasm and in a practically important cultivated breeding population, both on a genome-wide and on an individual gene basis. Additionally, we will have applied association mapping to this important crop legume for the first time.

APPROACH: We will use association mapping to identify genome regions and candidate genes that are associated with biomass production and cell wall composition in both diploid and tetraploid alfalfa populations. We are proposing to begin by screening a broad diversity of diploid germplasm (three individuals from each of 96 plant introductions) in order to identify new genetic variation for
these traits that could be useful in alfalfa improvement. We will begin by analyzing diploid genotypes because they likely harbor a reservoir of unexploited genetic diversity and are more tractable for association mapping experiments than tetraploid genotypes. Subsequently we will extend the results to tetraploids. The tetraploid population we will examine is a breeding population currently under clonal selection at four locations, with 200 individuals being evaluated. As a breeding population, markers associated with traits could be immediately used in a recurrent selection program leading to the development of improved cultivars. Phenotypic analysis will be conducted based on field grown plant material clonally replicated to enable assessment of individual genotypes. In addition to biomass production and plant height measurements, we will conduct a through analysis of the stem cell wall composition of all entries. All plants will be genotyped throughout the genome with simple sequence repeat (SSR) markers, some of which will be selected based on their association with quantitative trait loci (QTL) for biomass yield, stem cell wall cellulose, hemicellulose, and lignin concentration, or agronomic traits that we have identified in other experiments. Concurrently, we will sequence portions of up to 100 genes that are candidate loci involved with cell wall biosynthesis. The sequencing will lead to the identification of single nucleotide polymorphisms (SNP), which we will develop into markers for those specific genes. All plants (288 diploid and 200 tetraploid) will be genotyped with the SNP markers. Association mapping will be conducted using the recently described mixed-model method that will account for underlying population structure within our two groups of genotypes (diploid and tetraploid), which will be analyzed separately. We will test for associations based on both genome-wide SSR molecular markers, as well as on SNP markers for 20 candidate genes, which will be developed from sequence data on 96 diploid and 20 tetraploid individuals.