Candidate Genes for Lignin Structure Identified Though Genome Wide Association of Naturally Variant *Populus*

Nathan Bryant*[1] (nbryant3@vols.utk.edu), Jin Zhang[2], Kai Feng[2], Jin-Gui Chen[2], Wellington Muchero[2], Chang Geun Yoo[3,] Timothy J. Tschaplinski[2], Yunqiao Pu[2], Arthur J. Ragauskas[1,2], and **Gerald A. Tuskan**[2]

[1]Department of Chemical and Biomolecular Engineering, University of Tennessee, [2]Center for Bioenergy Innovation, [3]Department of Chemical Engineering, State University of New York College of Environmental Science and Forestry

cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI addresses strategic barriers to the current bioeconomy in the areas of 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition, and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to biofuels using CBP with cotreatment at high rates, titers and yield in combination with catalytic upgrading into drop-in hydrocarbon fuel blendstocks.

Understanding and controlling lignin variation will be a key technology in implementing biorefineries[1]. Further, genetic attribution of plant phenotypes is a barrier to the engineering and improvement on plant performance. To address these barriers, a genome wide association study (GWAS) was conducted to identify the underlying genetic basis of lignin variation. In this study, nuclear magnetic resonance (NMR) was utilized to elucidate twelve structural phenotypes of lignin across over 400 unique genotypes of three-year-old, naturally variant *Populus trichocarpa*. Population phenotyping revealed an average S/G ratio of 2.70 ± 0.45 and an average β-O-4 linkage content of 62.0% ± 3.18%. The population distribution of p-hydroxybenzoate (PB) content was unique in that it is best described by a lognormal distribution, whereas other phenotypes were better modeled by a normal distribution. PB content was also tended to be higher and trend with S/G ratio in samples with a low (<2.70) S/G ratio. The genotype-to-phenotype analysis identified 80 promising candidate genes strongly associated with at least one lignin phenotype. The candidate genes identified by this GWAS analysis can serve as guidance for future transgenesis targets. Several of these identified candidate genes have not previously been associated with lignin biosynthesis. Among the identified candidate genes include several from the 26S proteasome/ubiquitin pathway, including PBD1 (20S proteasome beta subunit D1), ATL13 (RING/U-box superfamily protein), and XBAT32 (RING-type E3 ligase). One candidate gene identified by this analysis, dirigent protein DIR23, was selected for overexpression in transgenic *Populus*. Another dirigent protein was shown to influence stereo- and regioselectivity of monolignol binding in lignans. Preliminary analysis of DIR23 overexpression lines indicates increased β-O-4 linkage content and lower molecular weight compared to wild-type. While additional analyses are required to confirm these results, these early findings serve as a promising indication that this GWAS analysis has identified novel candidate genes associated with lignin

phenotypes. The functional characterization of candidate genes identified in this study could provide new targets for the control of lignin biosynthesis.

Reference

1. U.S. Dept. of Energy (2016). *Biorefinery Optimization Workshop Summary Report*.

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