Pleiotropic and Epistatic Network-Based Discovery: Integrated Networks for Target Gene Discovery

Deborah Weighill[1,2], Piet Jones[1,2], Manesh Shah[2], Priya Ranjan[2], Wellington Muchero[2], Jeremy Schmutz[4,5], Avinash Sreedasyam[5], David Macaya-Sanz[6], Robert Sykes[3], Nan Zhao[7], Madhavi Z. Martin[2], Stephen DiFazio[6], Timothy J. Tschaplinski[2], Daniel Jacobson[1,2*] (jacobsonda@ornl.gov) and **Gerald A. Tuskan**[2]

[1]The Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, [2]The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, [3]National Renewable Energy Laboratory, Golden, CO, [4]Department of Energy Joint Genome Institute, Walnut Creek, CA, [5]HudsonAlpha Institute for Biotechnology, Huntsville, AL, [6]Department of Biology, West Virginia University, Morgantown, [7]The University of Tennessee Institute of Agriculture, University of Tennessee, Knoxville

https://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 will address 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 specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. And CBI will maximize product value by** *in planta* **modifications and biological funneling of lignin to value-added chemicals.**

Biological organisms are complex systems that are composed of functional networks of interacting molecules and macro-molecules. Complex phenotypes are the result of orchestrated, hierarchical, heterogeneous collections of expressed genomic variants. However, the phenotypic effects of genomic variants are the result of historic selective pressure, current environmental conditions and epigenetic signals, and, as such, their co-occurrence can be seen as genome-wide correlations in a number of different manners. Lignin is a complex polymer which is a significant component of plant cell walls and can be viewed as a polygenic phenotype of high importance to biofuels initiatives. This study makes use of data derived from the re-sequenced genomes from over 800 different *Populus trichocarpa* genotypes in combination with metabolomic and pyMBMS data across this population, as well as co-expression and co-methylation networks in order to better understand and identify target genes involved in lignin biosynthesis/deconstruction. A Lines-of-Evidence (LOE) scoring system is being developed to integrate the information in the different layers and quantify the number of LOE linking genes to lignin-related-phenotypes across the

network layers. The resulting Genome Wide Association Study networks, integrated with Single Nucleotide Polymorphism (SNP) correlation, co-methylation and co-expression networks through the LOE scores are proving to be a powerful approach to determine the pleiotropic and epistatic relationships underlying cellular functions and, as such, the molecular basis for complex phenotypes.