Identifying growth-coupled genetic loci using integrated metabolic modeling in populus

Debolina Sarkar,*[1] (debolina@psu.edu), Jin Zhang[2], Kai Feng[2], David Kainer[2], Timothy Tschaplinski[2], Dan Jacobson[2], Wellington Muchero[2], Costas D. Maranas[1] and **Gerald A. Tuskan**[2]

[1]Department of Chemical Engineering, Pennsylvania State University, University Park, PA; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, TN

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.

Designing feedstock transgenics with a combination of desired characteristics has been a focus of plant studies aimed at improving growth yield, feedstock composition, and nutrient-use efficiency. Although a small fraction of commercially available genetically modified crops present 'stacked' traits, only a handful of products have been developed that modify two or more genes. In this work, we identify enzymatic genes and their optimal combinations that accord a growth benefit in *P. trichocarpa* by integrated metabolic modeling. Previously, a large number of single nucleotide polymorphisms (SNP) have been identified in populations of *Populus trichocarpa* [1]. A previously developed SNP role assignment tool ('**SNPeffect**'[2]) was first used to identify functional SNPs and their effect on pathway flux. SNPeffect constructs mechanistic genotype-to-phenotype relationships by integrating the biochemical knowledge encoded in metabolic models with all available omics, phenotype, and sequence data. Inferred SNP activities were used to constrain corresponding gene activities, and then predict growth-coupled genetic interventions. As expected, our results indicate that growth is a complex polygenic trait governed by carbon and energy partitioning. The predicted intervention sets are associated with experimentally characterized growth-determining genes and also suggest putative ones. Predicted genes belong to pathways such as amino-acid metabolism, nucleotide biosynthesis, and cell wall and lignin biosynthesis, in line with breeding strategies that target pathways governing carbon allocation, and cellular ATP and reductant production. We envision this framework to be used in future for targeting additional agronomically-relevant traits such as robust growth yield, nutrient-use efficiency, and poplar feedstock composition. Beyond the design of stacked transgenics, we expected this work to have extensions in genetically assisted breeding of improved non-transgenic lines. This work contributes to the understanding of the molecular mechanisms governing growth and plant productivity, in order to further exploit them and develop a sustainable bioenergy feedstock.

References:

1. Tuskan, G. A., Muchero, W., Tschaplinski, T. J. & Ragauskas, A. J. Population-level approaches reveal novel aspects of lignin biosynthesis, content, composition and structure. *Current Opinion in Biotechnology* vol. 56 250–257 (2019).
2. Sarkar, D. & Maranas, C. D. SNPeffect: identifying functional roles of SNPs using metabolic networks. *Plant J.* **103**, 512–531 (2020).

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.

Support for the Poplar GWAS dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER) via the Bioenergy Science Center (BESC) under Contract No. DE-PS02-06ER64304. The Poplar GWAS Project used resources of the Oak Ridge Leadership Computing Facility and the Compute and Data Environment for Science at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.