

The GWAS Times: A Time-Series Aware GWAS to Detect Natural Climate Adaptations in *Arabidopsis* and *Populus*

Jared Streich^{1*} (ju0@ornl.gov), David Kainer¹, Ashley Cliff¹, Jonathan Romero¹, Daniel Jacobson¹, and Gerald A. Tuskan¹

¹Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

<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. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Elite crop lines are often bred for specific traits and trialed in farming environments which are synthetic compared to natural conditions. Maintenance of crops due to environmental effects can be costly and drive down profits. The Center for Biofuel Innovation aims to increase cellulosic biomass for biofuels and chemical feedstocks made from lignin residuals by use and improvement of two plant species, Switchgrass (*Panicum virgatum*) and Poplar trees (*Populus trichocarpa*). To improve genomic selection of potential breeding lines we created a time-series aware association approach called GWATS (Genome Wide Association Time-series Studies) to detect climate adaptive alleles and demonstrated them on *Populus trichocarpa*. GWATS involves running a GWAS analysis using time-series phenotype data, in this case for each day of the year. BioClim has been an ideal data series for species distribution modeling and finding adaptive alleles but lacks a true seasonal component and leaves much of each seasonal period unexamined. Instead, from raw monthly climate data, we interpolated 365 daily values of 78 climate/environment layers using 498 locations of 970 *Populus trichocarpa* individuals from where our GWAS population were sourced using ~10M genome-wide SNPs. Geographically isolated alleles tend to coincide with adaptive alleles making them difficult to distinguish from false positives. However, true positive p-values become distinct from false positives by using a Fourier Transform to analyze co-variate variation and output similarity matrix calculations across the time-series. Further, using a machine learning algorithm called iRF (iterative Random Forest) we can filter complex climate phenotypes into epistatic interactions across multiple suites of alleles ranging as high as five or more orders of epistasis. Using GWATS we detected hundreds of candidate climate adaptive loci for Solar Radiation Stress, Temperature Stress, Aridity Stress, Light Quality and many more. From this analysis we expect to create a list of climate-ideal lines to breed both climate wide and climate specific ideotypes in addition to being optimized for CBI-target phenotypes to improve agricultural production of *Populus trichocarpa*.

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.