High Resolution Analysis of Recombination Rates in *Populus trichocarpa*

Chanaka Roshan Abeyratne[1], David Macaya-Sanz[1], Gancho Slavov[2], Lee Gunter[3], Kathleen Haiby[4], Richard Shuren[4], Jay Chen[3], Daniel Jacobson[3], Brian Stanton[4], Stephen DiFazio,[1] (Stephen.DiFazio@mail.wvu.edu) and **Gerald A. Tuskan**[3]

[1]West Virginia University, Morgantown; [2]Scion, Rotorua, NZ; [3]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; and [4]GreenWood Resources Inc., Portland, OR

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

As part of the CBI goals, we have developed *Populus trichocarpa* as a model organism for enhancing feedstock characteristics. A major thrust of the project is to mine natural genetic variation in feedstock-relevant phenotypes using quantitative trait locus mapping in structured and unstructured populations. Accomplishing this goal requires a fine-scale understanding of meiotic recombination, which is one of the fundamental factors determining the outcome of natural and artificial selection. Here we describe an intensive mapping effort using whole genome resequencing of 829 progeny from 49 families to produce genetic maps representing recombination in fourteen parents. We show how recombination varies within and among genomes and between the sexes. This information is essential for understanding patterns of adaptive variation in populations as well as implementation of advanced marker-aided breeding approaches.

High resolution genetic maps for obligate outcrossing perennial trees such as *Populus trichocarpa* (black cottonwood) have been limited due to the high cost of genotyping and the challenge of establishing mapping populations that span multiple inbred generations in long-lived outcrossing trees (Yin, DiFazio, Gunter, Riemenschneider, & Tuskan, 2004). Most existing genetic maps for *Populus* have been constructed for two parents using their full-sibs or using maternal half-sibs for a limited number of fully informative molecular markers. Our work represents the first attempt at producing multiple linkage maps for both sexes with half-sib families in *P. trichocarpa*. We present seven genetic linkage maps for each sex and all 19 haploid linkage groups that allowed us to ask new questions related to sexual dimorphism as well as intra-specific variation in single generation cross-over (CO) rates across the genome.

We have re-sequenced the genomes of 49 families (N = 829 offspring), corresponding to a full factorial cross of seven females and males each. We have called biallelic SNP/INDELs using

GATK4. Using the pedigree information, we have revealed patterns of segregation in the genome and selected high-quality markers under Mendelian segregation. We have phased and imputed the progeny genomes and recovered the gametic haplotypes of the fourteen parents, allowing us to demarcate CO regions to a finer genomic scale. Using the R package Onemap (Margarido, Souza, & Garcia, 2007), we produced fourteen framework genetic maps that contain 1820 SNP/INDEL markers on average and cover 376.7 Mb of physical length across 19 chromosomes in the Stettler-14 male reference assembly. Resulting marker order is highly collinear with the physical position for all cases other than in LG-XVII for two of the male parents. Interestingly however, we observed considerable variation among individuals for observed recombination rates leading to an analysis of other covariates associated with this variation.

We observed a total of 38,846 CO events within 1658 gametes (829 offspring x 2 haplotypes of each parent) averaging approximately 1.2 CO events per linkage group. We observed male-biased CO rates in *P. trichocarpa* with female and male half-sib families contributing 18,355 and 20,491 COs respectively. Furthermore, we observed a chromosome scale interaction in several linkage groups that attenuated the high CO rate observed in males. However, statistical evidence for a female bias was not observed in any of the linkage groups. We used wavelet analysis to circumvent the issue of having to pre-select a window size to analyze sex-based differences in CO counts. This revealed a chromosome dependent pattern of sex differences at different resolutions.

We present evidence for 125 CO hotspot regions spread across the whole genome. Interestingly, there was minimal overlap of hotspot regions between males and females. We employed the MEME suite of software to identify DNA motif enrichment in 2,054 CO regions that were demarcated to a genomic region of 10 kbps or less and resulting motifs ranking within the top three were all A/T-rich repeat sequences. We identified several other motifs that showed striking similarity to DNA sequence motifs associated with CO's in maize, *Arabidopsis* and tomato as well (Zelkowski, Olson, Wang, & Pawlowski, 2019).

This analysis revealed the genome-wide recombination landscape and patterns of inheritance at a finer scale than has previously been possible in most plants. In *P. trichocarpa* this information could be used to identify Quantitative Trait Loci (QTL) for commercially important phenotypes and along with recombination rate estimates can be used to improve genomic prediction models. Given that *P. trichocarpa* is identified as a promising renewable feedstock for bioenergy and bioproducts, we believe that our findings will be useful for ongoing efforts to accelerate domestication of this and other feedstocks, as well as future studies that investigate broader questions such as evolutionary history, perennial development related to phenology, wood formation, vegetative propagation, and dioecy that cannot be studied using conventional plant model systems such as *Arabidopsis*, rice, or maize.

Margarido, G. R. A., Souza, A. P., & Garcia, A. A. F. (2007). OneMap: Software for genetic mapping in outcrossing species. *Hereditas*, *144*, 78–79.

Yin, T. M., DiFazio, S. P., Gunter, L. E., Riemenschneider, D., & Tuskan, G. A. (2004). Large-scale heterospecific segregation distortion in Populus revealed by a dense genetic map. *Theoretical and Applied Genetics*, *109*(3):451-463.

Zelkowski, M., Olson, M. A., Wang, M., & Pawlowski, W. (2019). Diversity and Determinants of Meiotic Recombination Landscapes. *Trends in Genetics*, *35*(5), 359–370.

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.