203. Structural Polymorphisms as Causes of Heterosis in Populus

Upendra Kumar Devisetty1, Kori Ault1, Alejandro Riveros Walker2, Sandra Simon2, Stephen Difazio2, Austin Himes3, Brian Stanton3, Todd Rosenstiel4 and Steven H. Strauss1*
(Steve.Strauss@oregonstate.edu)

1 Oregon State University, Corvallis OR; 2 West Virginia University, Morgantown WV; 3Greenwood Resources, Portland, OR and 4Portland State University, Portland, OR

http://people.forestry.oregonstate.edu/steve-strauss/

Project Goals: The main objectives of the project are to characterize the extent of structural polymorphisms (SP) that exist between and within species in cottonwood (poplar) trees that are used to produce wood and bioenergy, and examine their relationships to phenotypes such as growth, stress and breeding efficiency. The ultimate goal is to identify the extent to which SPs-primarily insertions and deletions of genes-could improve hybrid breeding in Populus compared to alternative approaches such as SNPs (single nucleotide polymorphisms) using genomic selection methods. We are using phenotypic and genomic information from existing research trials as well as generating substantial new phenotypic and genomic data.

Project background: Heterosis refers to the superior performance and stress tolerance of highly heterozygous groups or individuals. It is widely known for its applicability in maize. Many perennial biomass programs also employ heterosis either in the form of polyploids (e.g., Miscanthus) or in the form of interspecific hybrids (e.g., Populus, Salix, Acacia and Eucalyptus). The proposed research will test the predictions of a fundamental theory of heterosis (dominance) in relation to environmental stress, and its application to marker- aided selection to improve the efficiency of Populus breeding for woody biomass production. We will first analyze SPs in three distinct populations that serve specific roles in the project: a) in a wild P. trichocarpa collection of 1,100 wild genotypes where SP effects can be finely dissected due to relatively low LD; b) in an outbred hybrid backcross collection of 337 genotypes of (P. trichocarpa x P. deltoides) × P. deltoides, to jointly examine the effects of between- as well as within-species variation in SPs; and c) in F1 hybrid families of P. deltoides × P. nigra (DN), P. deltoides × P. maximowiczii (DM) and P. deltoides × P. koreana (DK) comprised of 371 progeny in total, to evaluate the efficiency of detection and prediction impact of SPs within diverse hybrids. We will generate SP genotype data using bait-capture/resequencing, and test the degree to which SPs are associated with variation in productivity traits (stem height, diameter, wood density, chlorophyll content, chlorophyll fluorescence; and carbon assimilation and respiration under imposed heat stress). Finally we will test whether the selected SPs-putatively focusing on selection against homozygosity for gene-containing deletions can improve selection efficiency and economic gain within families over phenotypic and genomic selection approaches.

Progress to date: Currently, a high quality reference genome is available for P. trichocarpa, but not for other poplar species. We will therefore carry out an assembly to produce reference genomes for the other species in our project. The genome of P. deltoides (WY94) was first assembled using Illumina paired end reads of 100bp using ABYSS denovo assembler. The generated assembly has an N50 of 10,490 using k-mer of size 63, below our requirements for this project. Further work on genome assembly for this clone will await its genome assembly by JGI-as part of a separately funded Community Sequencing Project. The other three reference genomes needed for the project are being analyzed through access to data from collaborators (P. nigra), and de novo sequencing (P. maximowiczii and P. koreana). We have generated Illumina MiSeq (2x300bp) coverage totaling 12X for each of these species and are currently generating an additional 40X of Illumina HiSeq sequences (2x100bp). In addition, we will generate mate pair

libraries and/or long read sequencing (PACBio) to span large indels and create more contiguous scaffolds using hybrid assembly approaches.

For genotyping of 1,100 trees of P. trichocarpa, we have used the 15X resequencing data available from DOE to perform SP discovery. For genotyping of the hybrid and backcross populations, we will first use a sequence-capture assay devised to efficiently target the genome containing the most polymorphic gene containing SPs within and among species. We will sequence a discovery population of 24 individuals per species to a depth of 15X for discovery of segregating SPs, and will then devise a sequence capture assay for these loci. Thus far we have made progress in optimizing high-throughput library preparation protocol for this purpose.

We made substantial progress in phenotypic assessments in summer 2014. Working with GreenWood Resources, we established a new population of F1 hybrid families in Jefferson, Oregon in spring 2014 that had 98% survival. We made phenotypic measurements here and in two other populations. Preliminary phenotypic analyses indicated that average photosynthesis measured in both light-adapted and dark-adapted leaves was significantly different for genotypes at increasing temperatures for both association and F1 populations. The dark-adapted chlorophyll fluorescence varied significantly among F1 families but not among genotypes within families. The light-adapted chlorophyll fluorescence was significant for F1 families and is also significant for genotypes in two out of 6 families. In the P. trichocarpa association population we also assayed light- and dark-adapted fluorescence and identified loci underlying these traits using genome- wide association analyses with our existing SNP dataset.

In the coming year we plan to do more phenotyping, including chlorophyll fluorescence, gas exchange as a function of temperature, leaf area, specific leaf mass, stomatal density, and height/DBH. Currently we are working on indel discovery in the hybrid and backcross populations; optimizing the bait-capture probe design for their determination; and then combining high-throughput DNA library production with bait-capture to generate SP genotype data. Finally, we will be developing a genomic selection model to assess if the produced genomic data and selection of SPs could be used to incrementally improve molecular accelerate breeding.

This project is supported by the USDA National Institute of Food Agriculture - Institute of Bioenergy, Climate and Environment, competitive grant no. 2013-67009-21008