Local adaptation in *Populus trichocarpa* Torr. & Gray

Hari B. Chhetri[1], Gancho T. Slavov[2], David Macaya-Sanz[1], Nesatalu Hiese[3], Gerald A. Tuskan[4], and Stephen P. DiFazio[1]

[1]West Virginia University, Morgantown, WV; Scion, [2]Bay of Plenty, New Zealand; [3] Nagaland Science & Technology Council, Kohima – 797001, Nagaland, India; [4]Oak Ridge National Laboratory, Oak Ridge, TN

Introduction:

- Some plants have higher fitness in their native environments compared to foreign environments, which is termed as local adaptation.
- Understanding factors affecting local adaptation and the genetic mechanisms underlying local adaptation is key for accelerated domestication, conservation and management of *P. trichocarpa* and ultimately developing this species as a biofuel crop.

Objectives:

- Is climate shaping the adaptive variation in *P. trichocarpa*?
- What are the loci conferring local adaptation in *P. trichocarpa*?
- How much of the adaptive variation at the genomic level is explained by the climate?
- How much of the adaptive variation in phenotypes is due to climate and genetics (genomic variation?

Experimental Approach:

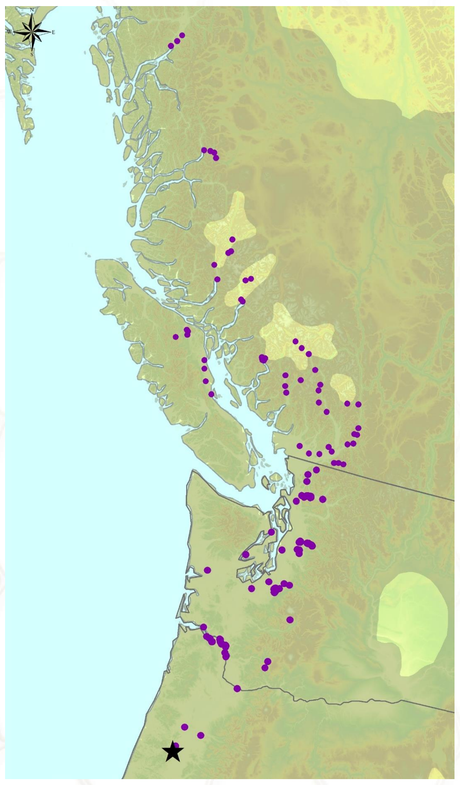


Figure 1. Source locations of 869 *P. trichocarpa* genotypes sampled in this study (purple dots). The trees were grown in a common garden in Corvallis, Oregon (black star).

- All phenotypic data were collected from 2010-2015 in the common garden in Corvallis, Oregon (Table 1) (Evans et al., 2014; Chhetri et al., 2019).
- Climate data from 1990 to 2010 were collected from the source location of the trees.
- Genotype-environment association (GEA) analyses were performed with 6.74 million SNPs with univariate (Zhou and Stephens., 2010) as well as the multivariate (Zhou and Stephens, 2014) GEMMA.
- Multivariate methods called Redundancy Analysis (RDA) was used to explain the variation in the response matrix constrained by the predictor variables.

Table 1. Broad-sense heritability estimates (H2) of morphological, physiological and phenological traits in *P. trichocarpa.*

Trait	H^2	GenoV[a]	N[b]	Total trees[c]
Height (HT)	0.363	0.152	876	2378 (851)
Leaf area (LA)	0.344	872.790	794	1056 (262)
Leaf aspect ratio (AR)	0.462	0.033	794	1056 (262)
Petiole diameter (PD)	0.297	0.147	839	1124 (285)
Petiole length (PL)	0.561	70.732	839	1124 (285)
Specific leaf area (SL)	0.371	101.465	784	1010 (226)
Stomatal density (SD)	0.500	215.654	813	1064 (251)
Bud flush	0.825	0.453	877	2442 (2426)
Bud set	0.628	0.253	841	2173 (2133)
Carbon isotope (CI)	0.363	0.240	681	759 (78)
SPAD2014 (SP)	0.310	6.420	839	1124 (285)

[a]Genotypic variance component; [b]Number of genotypes
[c]Number of ramets sampled, with replicates in parentheses

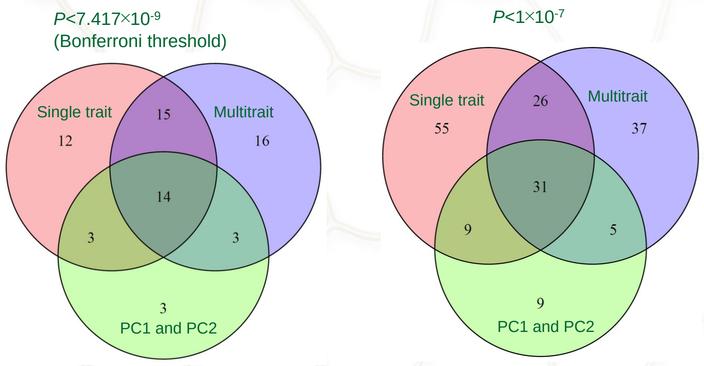


Figure 2. Genes detected by genotype environment association (GEA) across the methods – single trait, multitrait and PC-based.

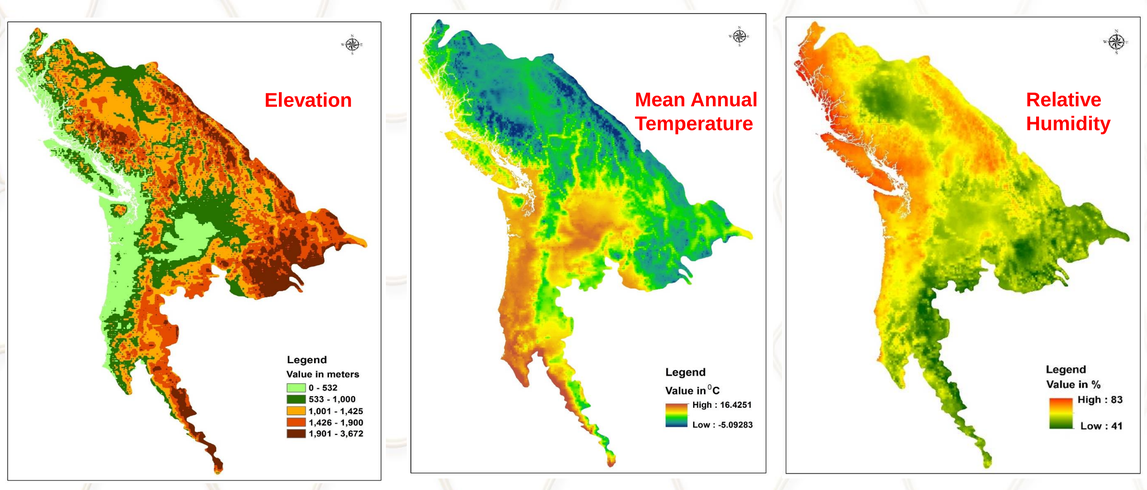


Figure 3. Geography and climate maps for the range of *P. trichocarpa* distribution.

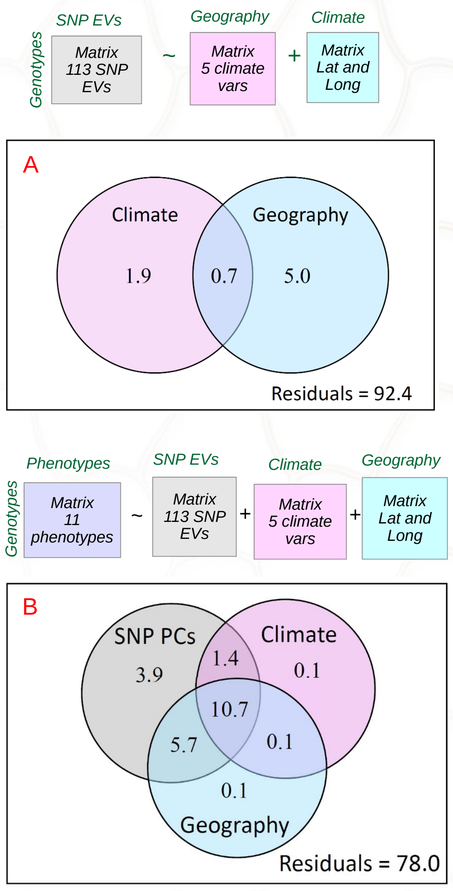


Figure 5. Partitioning of variance components in RDA analyses. (A) SNP ~ Clim + Geo model. 7.6% is of the total variation is explained by climate and geography (matrix of spatial variable). (B) Pheno ~ SNP + Clim + Geo model. 22% is of the total variation is explained by SNP EVs, climate and geography. More than 10 million SNPs with MAF≥0.01 were decomposed into 868 SNP eigenvectors (EVs) and the first 113 significant EVs were used in all RDA analyses.

Table 2. Pearson correlation between phenotypic traits and geo-climate variables and PCs of 26 geo-climate variables. Values >0.12 or <-0.12 are significant. Dark blue and dark red colors indicate high negative and high positive *P*-values.

Geo-climate*/Trait	Bud flush	Bud set	Height	Leaf area	Petiole length	Petiole diameter	Leaf aspect ratio	Specific leaf area	Stomatal density	Carbon isotope	SPAD
Lat	-0.51	0.61	-0.42	-0.15	-0.18	-0.10	0.21	-0.33	-0.17	0.03	0.18
Elev	-0.256	0.293	-0.229	-0.072	-0.149	-0.033	0.123	-0.131	-0.035	0.12	0.091
MAT	0.46	-0.56	0.43	0.19	0.24	0.12	-0.22	0.22	0.07	-0.09	-0.16
MWMT	0.36	-0.30	0.21	0.08	0.09	0.07	-0.15	0.08	0.09	-0.02	-0.09
TD	-0.33	0.56	-0.46	-0.23	-0.26	-0.12	0.21	-0.26	0.00	0.09	0.16
MAP	-0.34	0.25	-0.19	-0.01	-0.06	-0.02	0.03	-0.14	-0.03	-0.04	-0.01
SHM	0.32	-0.25	0.10	-0.09	-0.05	-0.05	-0.03	0.20	0.14	0.04	-0.05
RH	-0.37	0.15	-0.05	0.07	0.03	0.05	-0.02	-0.17	-0.23	0.05	0.05
PC1	0.48	-0.56	0.42	0.17	0.22	0.11	-0.21	0.24	0.08	-0.07	-0.15
PC2	0.16	0.05	-0.12	-0.18	-0.17	-0.10	0.10	0.09	0.16	0.07	0.04

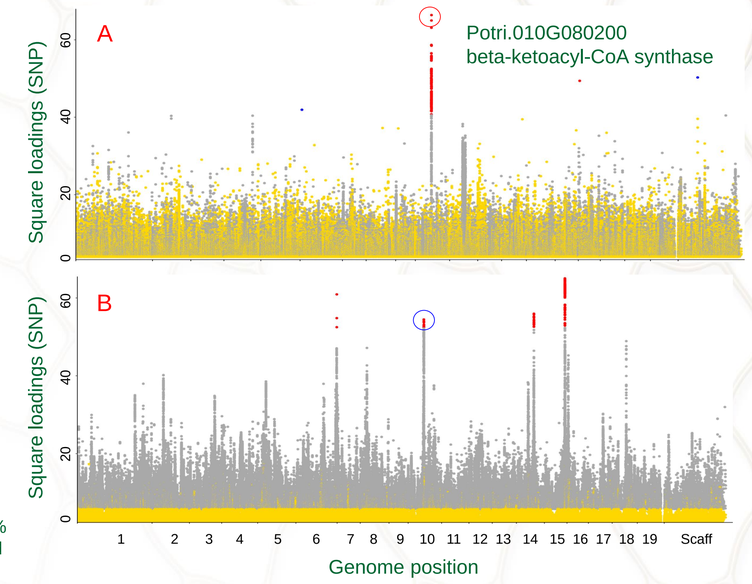


Figure 6. Square loadings of top 0.1% SNPs from one of the 5 significant eigenvectors (EV32) detected from the climate RDA analysis (A) and one of the 17 outlier eigenvectors (EV1) detected from the phenotypic RDA analysis (B). Numbers 1 to 19 represent chromosomes. Gray and yellow colors represent positive and negative square loadings, respectively. Red and blue indicate top 100 square loadings corresponding to positive and negative loadings, respectively. Climate RDA (EV32) identified the same gene (circled red) that was identified in GEA. Phenotypic RDA (EV1) also identified a similar gene, Potri.010G80300 (circled blue) just 100 bp upstream of the Potri.01G080200.

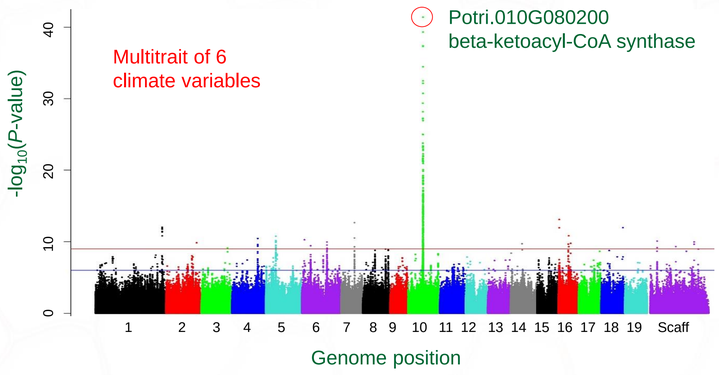
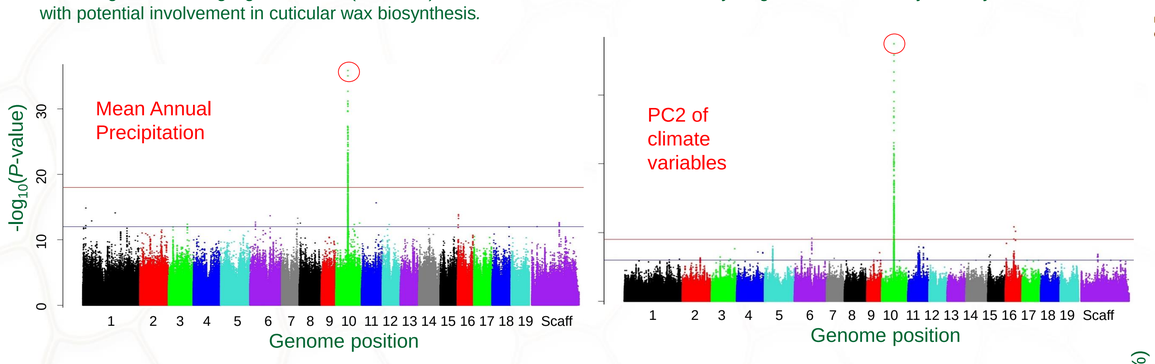
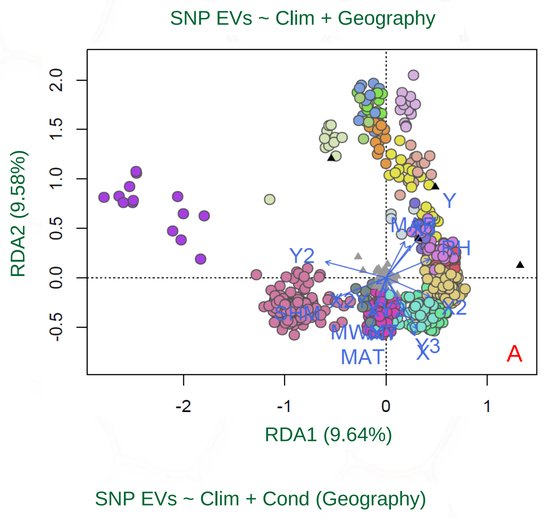


Figure 4. Single trait, multitrait and PC based association showing the association of geo-climate variables with the SNPs in chromosome 10. Color panels represent chromosome numbers 1 to 19 and scaffolds (last panel). Red horizontal line indicates Bonferroni correction threshold ($P = 7.417\times10^{-9}$) and blue horizontal line indicates suggestive association threshold ($P=1\times10^{-7}$). Nearest gene to the highlighted SNP/s (red circle) is Potri.010G080200, which is a very long chain beta-ketoacyl-CoA synthase with potential involvement in cuticular wax biosynthesis.





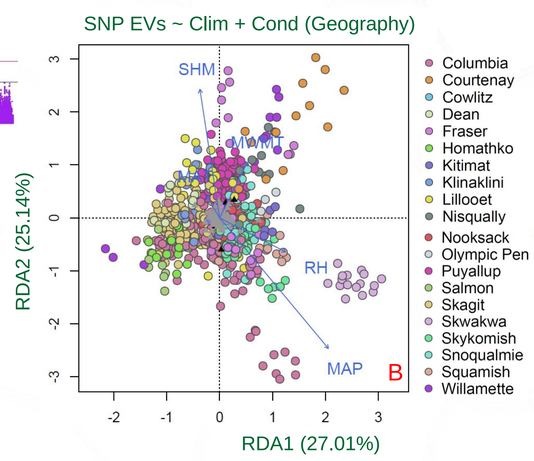


Figure 7. RDA triplots showing SNP eigenvectors constrained by the matrix of climate and geography variables (A) and climate variables with geography as covariates (B). Both RDA1 and RDA2 axes were significant in all RDA models. It appears that the effects of climate are confounded by geography. Factoring out the effect of geography reveals the effects of climate. Circles indicate 869 *P. trichocarpa* genotypes (color coded by river, the populations). Triangles indicate SNP eigenvectors with outlier eigenvectors (based on a 3 standard deviation cutoff selected from the tails of the distribution of RDA axes) in black. Blue arrows indicate the influence of climate variables on RDA axis.

Conclusions:

- Correlation of phenotypes with climate variables suggests a role of climate in shaping genetic variation.
- Climate alone significantly explained 1.9% of the variation in the genome.
- SNPs alone explained 3.9% of the variation in phenotypic traits.
- The RDA method allowed deconvolution of the effects of climate from geography.
- Several key genes identified from GEA and RDA are potentially good targets for breeding and genetic engineering.

References:

Chhetri, H. B., Macaya-Sanz, D., Kainer, D., Biswal, A. K., Evans, L. M., Chen, J. G., … DiFazio, S. P. (2019). Multitrait genome-wide association analysis of *Populus trichocarpa* identifies key polymorphisms controlling morphological and physiological traits. *New Phytologist*, *223*, 293–309. doi: 10.1111/nph.15777

Evans, L.M., G.T. Slavov, E. Rodgers-Melnick, J. Martin, P. Ranjan, et. al., 2014. Population genomics of Populus trichocarpa identifies signatures of selection and adaptive trait associations. Nature Gen. 46: 1089–1096.

Zhou, X., M. Stephens, 2012. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics. 44: 821-824.

Zhou, X., M. Stephens, 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nature Methods. 11(4): 407-409.

Acknowledgments:

We would like to thank many collaborators in the Center for Bioenergy Innovation who helped with plantation establishment and maintenance and collection of phenotypic and genotypic data. In particular, the personnel of Greenwood Resources and the Joint Genome Institute who could not be specifically listed here due to space constraints. We also thank Robin Paulman and the Appalachian Ecology Lab for the carbon isotope assays.

The BioEnergy Science Center (BESC) 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