Association Genetics of Water Relations and Growth Phenotypes in Loblolly Pine

Patrick Cumbie[1], Ross Whetten[1], Barry Goldfarb[1],

Andrew Eckert[2], Jill Wegrzyn[3], and David Neale[3]

As a part of the NSF funded *Allele Discovery of Economic Pine Traits II* (ADEPT2) our objective was to discover genotype-phenotype associations which explained variation in water use efficiency in loblolly pine (*Pinus taeda* L.) using an association genetics approach. The association population consisted of 500 unrelated genotypes from the North Carolina State University Cooperative Tree Improvement Program and the Western Gulf Forest Tree Improvement Program. Our aim was to find significant associations from among 3900 SNPs and several phenotypes including carbon isotope discrimination and assessments of growth. Carbon isotope discrimination ($\delta^{13}C$) has been observed to show moderate to high levels of heritability in several conifer species and is correlated with stomatal conductance and photosynthetic capacity (Johnsen et al. 1999; Brendel et al. 2002; Baltunis et al. 2008). Recent work by Gonzalez-Martinez et al. (2008) suggested that association genetics may be a successful method to discover genes influencing a complex trait such as carbon isotope discrimination.

Phenotypic data were collected on 450 clonally replicated genotypes (rooted cuttings) that were planted in a closely spaced trial and grown for 2 seasons before harvest. The population consisted of offspring from natural stand selections representing the natural range of loblolly pine, with one offspring per selection used in the population (Figure 1). Foliar tissue was collected from the second flush of the second year's growth at the end of the second growing season. Heights were recorded after the second growing season. Samples were sent to the Cornell University Stable Isotope Laboratory (COIL) for foliar carbon isotope discrimination and nitrogen concentration analysis. A two-stage approach was taken for association testing. Phenotypic data was analyzed with a mixed model approach which incorporated a spatial AR1 x AR1 auto-regressive adjustment to remove additional environmental error. BLUP predictions for clones were then used as phenotypes for association testing. Mixed model analysis of phenotypes was performed in the ASReml 2.0 statistical software package (Gilmour et al. 2006). Association testing

[1] Department of Forestry and Environmental Resources, North Carolina State University, Box 8002 Raleigh, NC 27695-8002, USA

[2] Section of Evolution and Ecology, University of California at Davis, Davis, CA 95616, USA

[3] Department of Plant Sciences, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA

was performed in TASSEL using a generalized linear model which incorporated population structure estimates in the model (Bradbury et al. 2007) and the q-value was used to adjust for multiple testing (Storey 2003).

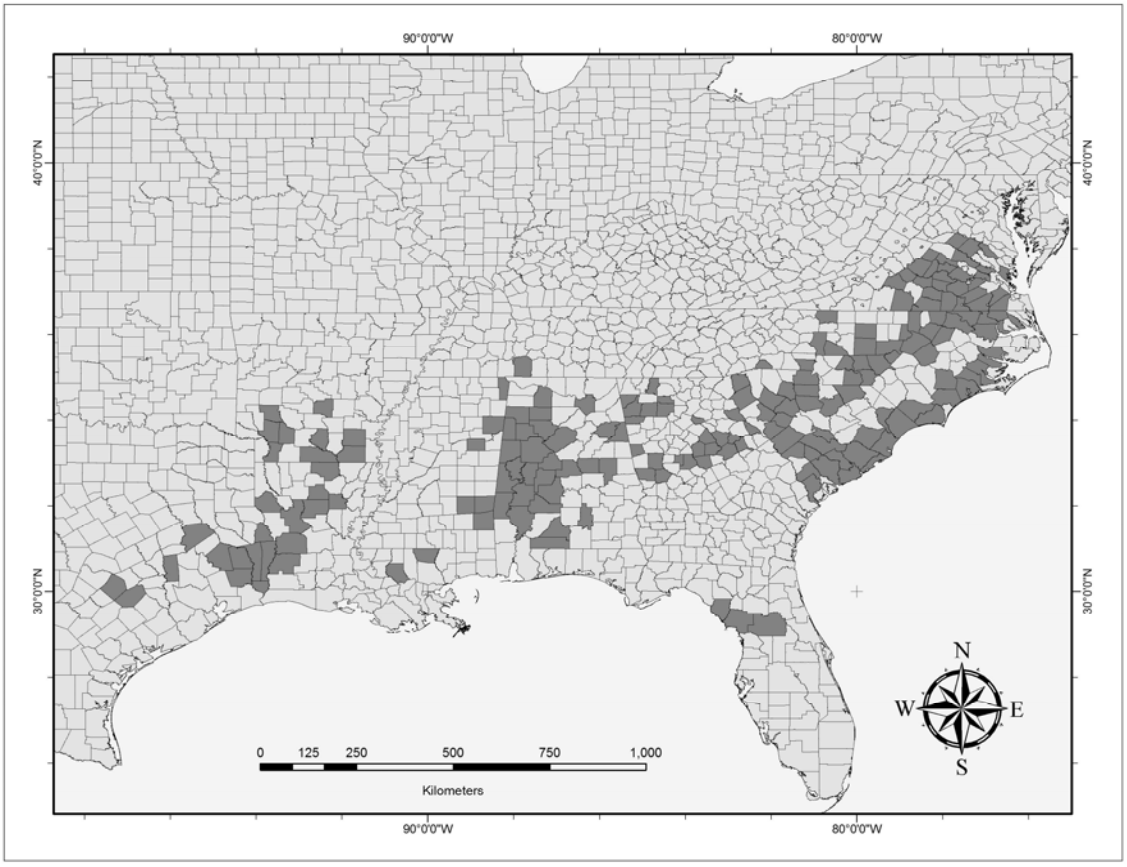

Figure 1. Counties of origin for genotypes incorporated in the loblolly pine association population.

Results showed significant genetic variation in $\delta^{13}C$ (H^2_c=0.48) and height at age two (H^2_c=0.48). Height and $\delta^{13}C$ were moderately correlated (r_g = 0.38) among unrelated clones (Figure 2). Baltunis et al. (2008) reported a low individual-tree narrow-sense heritability for $\delta^{13}C$ in a field-grown study of loblolly pine at three years of age. The association population in this study cannot partition additive from non-additive variation within the population, but the single site broad-sense heritability estimate for $\delta^{13}C$ was comparable to that found by Baltunis and was improved by the use of the AR1 by AR1 spatial model. Analyses using the generalized linear model for association testing revealed 55 significant SNPs associated with $\delta^{13}C$ at the $p<0.01$ significant level, while 32 SNPs were associated with height and 56 SNPs were associated with foliar nitrogen

concentration. After correction for multiple testing, there were 4, 0, and 5 SNPs significant using a 5% false discovery rate (q-value) for $\delta^{13}C$, height, and foliar nitrogen

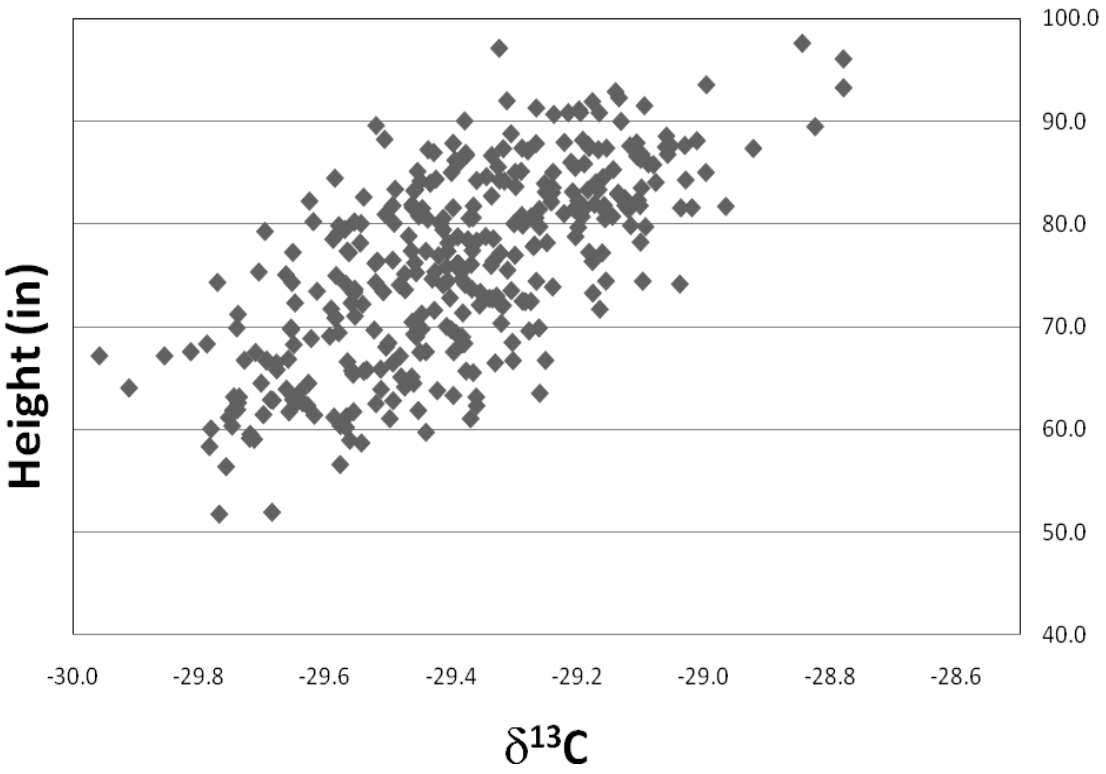

Figure 2. Plot of height and $\delta^{13}C$ for genotypes in the loblolly pine association population.

concentration, respectively (Table 1). The extreme values of the population (>1.5 standard deviations above and below the population mean) were compared for allele frequency differences. Only 5 out of the top 20 SNP loci associated with carbon isotope discrimination had significantly different allele frequencies at the $p<0.10$ level while only 1 was significantly different at the $p<0.05$ level using a chi-square test. These results indicate that SNP polymorphisms can be related to variation in carbon isotope discrimination and growth traits in loblolly pine.

Variation among clones was significant for $\delta^{13}C$, height, and foliar nitrogen concentration after two growing seasons. Positive associations were found between polymorphic SNP loci and phenotypic variation of complex traits. The discovery of SNPs associated with $\delta^{13}C$ phenotypic variation may be an opportunity for loblolly pine breeding programs to increase water use efficiency through the selection and breeding of individuals with rare alleles. SNPs associated with height and foliar nitrogen concentration could potentially

be used for improved growth rate. Future work will aim to further characterize differences in water relations and growth between the tails of the association population and to validate the SNPs found in this population in a different population.

Table 1. SNP loci associated with foliar nitrogen concentration (%N), $\delta^{13}C$, and height (Ht) after correction for multiple testing.

Trait	SNP_ID	qvalue	Annotation
%N	UMN-6338-01-99	0.0007	Predicted protein in Populus trichocarpa, phospholipase-like protein in *Arabidopsis*
%N	2-7865-01-156	0.0007	BAM1; ATP binding / kinase/ protein serine/threonine kinase in *Arabidopsis thaliana*
%N	2-1087-01-86	0.0026	conserved hypothetical protein [Ricinus communis]
%N	CL1074Contig1-03-101	0.0061	EF-hand, calcium binding motif, unknown protein in *Picea sitchensis*
%N	0-17195-01-417	0.0149	glutamate decarboxylase in *Pinus pinaster*
$\delta^{13}C$	0-17030-01-94	0.0019	hypothetical protein in *Vitis*
$\delta^{13}C$	2-1501-01-109	0.0274	Predicted protein similar to Pm27 in *Vitis vinifera*
$\delta^{13}C$	0-8304-02-414	0.0363	Predicted protein in *Vitis*, *Populus*; beta-glucosidase in *Ricinus*
$\delta^{13}C$	0-17543-01-196	0.0382	Nitrate transporter
$\delta^{13}C$	0-10921-01-353	0.0665	No significant hit
Ht	0-14415-01-190	0.0887	No significant hit

References

Baltunis, B. S., J. M. Davis, D. A. Huber and T. A. Martin (2008). "Inheritance of foliar stable carbon isotope discrimination and third-year height in Pinus taeda clones on contrasting sites in Florida and Georgia [electronic resource]." Tree genetics & genomes **4**: 797-807.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss and E. S. Buckler (2007). "TASSEL: software for association mapping of complex traits in diverse samples." Bioinformatics **23**(19): 2633-2635.

Brendel, O., D. Pot, C. Plomion, P. Rozenberg and J. M. Guehl (2002). "Genetic parameters and QTL analysis of delta13C and ring width in maritime pine." Plant, cell and environment **25**(8): 945-953.

Gilmour, A. R., B. J. Gogel, B. R. Cullis and R. Thompson (2006). ASReml User Guide Release 2.0. Hemel Hempstead, UK, VSN International Ltd**:** 320pp.

Gonzalez-Martinez, S. C., D. Huber, E. Ersoz, J. M. Davis and D. B. Neale (2008). "Association genetics in Pinus taeda L. II. Carbon isotope discrimination." Heredity **101**(1): 19-26.

Johnsen, K. H., L. B. Flanagan, D. A. Huber and J. E. Major (1999). "Genetic variation in growth, carbon isotope discrimination, and foliar N concentration in Picea mariana: analyses from a half-diallel mating design using field-grown trees." Canadian journal of forest research **29**(11): 1727-1735.

Storey, J. D. (2003). "The positive false discovery rate: A Bayesian interpretation and the q-value." Annals of Statistics **31**(6): 2013-2035.