More on the population genetics and introgression of loblolly pine (*Pinus taeda* L.) and shortleaf pine (*P. echinata* Mill.)

John F. Stewart[1], Yanyan Liu[1], C. G. Tauer[1], & C. Dana Nelson[2]

Loblolly pine (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) are both important timber species native to and grown in the southeastern United States. While both species have allopatric ranges, most of their ranges overlap in a sympatric range. The Mississippi River acts as a distinct boundary between eastern and western populations of these two species. The two species can be crossed with each other, either artificially (Scrheiner, 1937) or in nature (Zobel, 1953; Xu et al, 2008A).

Previous studies have used morphology, isoenzymes, and amplified fragment length polymorphism (AFLP) markers to measure the hybrid character of the two pine species. One promising isoenzyme locus, the isocitrate dehydrogenase locus (*IDH*) has shown promise as a good marker for hybrid identification. However, a recent study showed that it is inadequate (Xu et al, 2008A) for hybrid determination. In this study, we used short sequence repeat (SSR) markers, expressed sequence tag short sequence repeat (EST-SSR) markers, and the *IDH* isoenzyme marker to measure the hybrid character of trees from the Southwide Southern Pine Seed Source Study, a representation of most of the ranges of loblolly pine and shortleaf pine.

Materials and Methods

Needle Tissue Samples and DNA Extraction

Needle samples taken from 205 trees planted in the Southwide Southern Pine Seed Source Study, a range-wide provenance test established in 1953, were used in this study. These samples represent pine populations existing before widespread forest management in the southeastern United States. DNA was extracted from the loblolly and shortleaf pine samples using a modified CTAB protocol (Doyle & Doyle, 1988).

Microsatellite and Isocitrate Dehydrogenase (*IDH*) Markers

The 25 microsatellite markers used in this study were selected from markers previously determined to be polymorphic in shortleaf pine and loblolly pine by Nelson et al (2007). Additionally, 17 expressed sequence tag primer pairs were used from Chagne et al (2004). The DNA samples were analyzed with the LI-COR 4300 DNA Analyzer (LI-COR Biosciences, Lincoln, NE.) Running conditions were 1500 V, 40 W, and 40 mA. The *IDH* isoenzyme data used in this study were from Xu et al (2008A).

Population Genetics and Hybrid Analysis

[1] Natural Resources Ecology and Management, Oklahoma State University

[2] USDA Forest Service

General population genetics analysis was performed with the software GenAlEx 6.2 (Peakall & Smouse, 2006.) PhiPT and genetic distances were calculated with this software. Structure 2.2 (Pritchard et al, 2000) was used to determine hybrid identities.

Results and Discussion

Genetic Diversity

We used SSR, EST-SSR, and *IDH* isoenzyme data to calculate molecular variance, Phi-PT (Φ_{PT}), a measure similar to Wright's fixation index, F_{ST}. We also calculated Φ_{PT} from the AFLP data from Xu et al (2008B). These results are consistent with prior studies of shortleaf pine and loblolly pine. See Table 1 for details.

Table 1: Summary of Data. Metapopulation indicates which group of the samples is being discussed. Total trees indicates the total number of trees in the metapopulation. Number of hybrids shows the count of hybrids determined by Structure 2.2, using both the codominant data from this study and the AFLP data first presented in Xu et al (2008A). Φ_{PT} was also calculated for each group, using codominant data from this study, as well as the AFLP data from Xu et al (2008B).

Metapopulation	Total Trees	Number of Hybrids (%) calculated from codominant data	Number of Hybrids calculated from AFLP data	Φ_{PT} calculated from codominant data	Φ_{PT} calculated from AFLP data
All Pines	205	9 (4.4%)	13 (6.3%)	n/a[1]	n/a[1]
Loblolly Pine	112	5 (4.5%)	3 (2.7%)	0.061	0.059
Loblolly Pine West	22	2 (9.1%)	1 (4.6%)	0.010	0.019
Loblolly Pine East	90	3 (3.3%)	2 (2.2%)	0.054	0.054
Shortleaf Pine	93	4 (4.3%)	10 (10.8%)	0.085	0.057
Shortleaf Pine West	43	4 (9.3%)	7 (16.3%)	0.090	0.035
Shortleaf Pine East	50	0 (0%)	3 (6.0%)	0.082	0.077

[1] **Note that since loblolly pine and shortleaf pine are different species the Φ_{PT} calculated for all pines is a largely meaningless number.**

Genetic Distance and Geographic Distance

A comparison of geographic and genetic distances among populations of shortleaf pine and loblolly pine yield weak correlations-- for loblolly pine, $R^2 = 0.43$ and for shortleaf pine, $R^2 = 0.17$. The populations from Ashley County, AR, had unusually high genetic distances so they were removed prior to the correlation calculation. It is intuitive that genetic distance should correlate with geographic distance to some degree, and it appears to do so in this case. Importantly, most pollination events occur with local pollen, and pollen loses its viability over distance, so while it is possible for trees to be sired by distant parents, it is unlikely. Since the two species' migration north in the Pliocene epoch, apparently not enough time has transpired to allow for gene mixing across all subpopulations.

Hybrid Pines

Of the 205 pines in this study, 9 were determined to be hybrids (Table 1), using threshold values (estimated proportion loblolly genome) of greater than 0.0469 and less than 0.9531. These results are consistent with prior studies, with more hybrids existing west of the Mississippi River than east of it. However, on a tree-by-tree basis, our results are in disagreement with the AFLP analysis (Xu et al, 2008A). The codominant SSR analysis and the dominant AFLP analysis only agreed on 3 of the trees being hybrids.

We think that there are several important reasons to consider the codominant data in this study as more robust than that of Xu et al (2008A). First, codominant markers are better at detecting hybrids than are dominant markers, since they can detect heterozygotes, and the amount of data that is produced for a given locus is greater. Also, the two parents and two hybrid offspring used to select the AFLP markers in Xu et al (2008A) have now been shown to be unrelated causing some potentially informative AFLP markers to have been discarded. We are uncertain how this may have affected the results, but it may explain some of the differences.

Literature Cited

Chagne D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera MT, Vendramin GG, Garcia V, Frigerio JM, Echt C, Richardson T, and Plomion C. 2004. Cross-species transferability and mapping of genomic and cDNA SSRs in pines. Theoretical and Applied Genetics 109: 1204-1214.

Doyle JJ & Doyle J. 1988. Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Pritchard JK, Stephens M, & Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Schreiner EJ (1937) Improvement of forest trees. 1937. Yearbook of Agriculture, USDA pp1242-1279.

Xu S, Tauer CG, & Nelson CD. 2008A. Natural hybridization within seed sources of shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.) Tree Genetics and Genomics 4: 849-858.

Xu S, Tauer CG, & Nelson CD. 2008B. Genetic diversity within and among populations of shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.) Tree Genetics and Genomes 4: 859-868.

Zobel BJ. 1953. Are there natural loblolly-shortleaf pine hybrids? Journal of Forestry 51: 494-495.

Nelson CD, Josserand S, Echt CS, & Koppelman J. 2007. Loblolly pine SSR markers for shortleaf pine genetics. In: Kabrick JM, Dey DC, & Gwaze D, eds. Shortleaf pine restoration and

ecology in the Ozarks: Proceedings of a symposium; November 7-9, 2006; Springfield, MO. Gen. Tech. Rep. NRS-P-15. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: pp. 95-98.