THE POPULATION GENETICS OF LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*PINUS ECHINATA* MILL.) HYBRIDIZATION AND INTROGRESSION

By

JOHN F. STEWART

Bachelor of Science in Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts
2003

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2010

THE POPULATION GENETICS OF LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*PINUS ECHINATA* MILL.) HYBRIDIZATION AND INTROGRESSION

Dissertation Approved:

Dr. Charles G. Tauer

Dissertation Adviser

Dr. Liuling Yan

Dr. Yanqi Wu

Dr. Yinghua Huang

Outside Committee Member

Dr. Mark E. Payton

Dean of the Graduate College

ACKNOWLEDGMENTS

I thank Mr. Robert Heineman and the crew of the Kiamichi Forestry Field Research Unit of Oklahoma State University for their assistance in acquiring the samples for Chapter 1 and Chapter 3, as well as some of the samples for Chapter 3. I my wife Mary Tsien for her effort in making Figure 1-1 and Figure 2-1. I thank Dr. Shiqin Xu whose AFLP work laid the foundation for Chapter 1. I would like to thank all of the foresters who helped us gather the needles for this study: Scott Schlarbaum and David Griffin of the University of Tennessee; Roy Ward of the Ames Plantation; Thomas Hall of Pennsylvania State University; Russ Pohl of the Georgia Forestry Commission; Scott Merkle of the University of Georgia; William E. Dienst of the United States Forest Service in North Carolina; Gregory Powell and Dale Rye of the University of Florida; Larry Miller and I.N. Brown of Texas Agricultural and Mechanical University; Michael Mills of Claybourn Waters Company; Brad Claus of the Louisiana Department of Agriculture and Forestry; Randall Rousseau of Mississippi State University; David Gwaze of the Missouri Department of Conservation; and Robert Heinemann of the Kiamichi Forestry Research Station. I thank James Guldin of the United States Forest Service for aiding in finding collection sites and for consultation regarding the history of the region discussed in Chapter 3. Funding partly originated from the US Forest Service Southern Institute of Forest Genetics, Coop Agreement No SRS 05-CA-11330126-168. This study was supported by the Oklahoma State University Agricultural Experiment Station.

.

TABLE OF CONTENTS

Chapter Page

I. MICROSATELLITE VERSUS AFLP ANALYSES OF PRE-MANAGEMENT INTROGRESSION LEVELS IN LOBLOLLY PINE (PINUS TAEDA L.) AND SHORTLEAF PINE (P. ECHINATA MILL.) 1

Abstract 1
Introduction 2
Materials and Methods 3
Results 11
Discussion 16
Literature Cited 21

II. BIDIRECTIONAL INTROGRESSION BETWEEN LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*P. ECHINATA* MILL.) HAS INCREASED SINCE THE 1950S. 26

Abstract 26
Introduction 27
Materials and Methods 28
Results 34
Discussion 40
Literature Cited 46

III. HYBRIDIZATION IN NATURALLY REGENERATED SHORTLEAF PINE NEAR ARTIFICIALLY REGENERATED STANDS OF LOBLOLLY PINE. ..52

Abstract 52
Introduction 53
Materials and Methods 54
Results 59
Discussion 63
Literature Cited 66

LIST OF TABLES

Table Page

1-1 Loci 9
1-2: Metapopulation Φ_{PT} and Number of Hybrids estimated from SSR data and AFLP data. 15
1-3: Trees were identified as hybrids by Structure 2.2 analysis of the codominant SSR marker data and the AFLP data. 19
2-1: Marker Loci. The microsatellite marker loci used in this study. 32
2-2: Phi-PT and the hybrid count for the populations in this study, as well as the comparable populations from Stewart et al (2010). 35
2-3: Hybrids by collection location 37
3-1: Marker names and alleles. 57
3-2: X^2 tests. 61

LIST OF FIGURES

Figure Page

1-1: The Sources for all loblolly pine and shortleaf pine used in this study.....6
1-2: Loblolly Pine Genetic Distance vs. Geographic Distance.....13
1-3: Shortleaf Pine Genetic Distance vs. Geographic Distance.....14
2-1: The ranges of loblolly pine and shortleaf pine, as well as the sampling locations.....30
2-2: Genetic Distance vs. Geographic Distance.. 36
2-3: Number of pines by hybrid type.. 38
2-4: The weighted hybrid index score (Goodman, 1967) plotted on the y-axis against the values generated by Structure.. 40
3-1: Map of collection locations.. 55
3-2: The change in hybridization rates in each population.. 60
3-3: The correlation of distance from the loblolly plantations to the sample sites with the percent hybrids and the average Structure results.....62

CHAPTER I

MICROSATELLITE VERSUS AFLP ANALYSES OF PRE-MANAGEMENT INTROGRESSION LEVELS IN LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*P. ECHINATA* MILL.)

Stewart JF, Liu Y, Tauer CG, and Nelson CD

Abstract

Loblolly pine and shortleaf pine are known to naturally hybridize. In this study, we used 42 microsatellite markers and isocitrate dehydrogenase isozyme to create genetic profiles of 202 loblolly and shortleaf pine trees grown from seed collected in the 1950s for the Southwide Southern Pine Seed Source Study. Estimated PhiPT (Φ_{PT}) was low in both loblolly (0.061) and shortleaf (0.080) pines, indicating that most of the diversity is accounted for within seed sources. However, both loblolly and shortleaf pines showed significant correlations between seed sources' genetic and geographic distances, with R^2 of 0.43 and 0.17, respectively. The hybridization rate was 4.0%, with more hybrids west of the Mississippi River (8.1%) than east of the river (2.1%). Additionally, about the same proportion of both species (4.5% of loblolly and 3.3% of shortleaf pine) were identified as hybrids. These results are consistent with prior studies on these two species but do contrast with the results from an amplified fragment length polymorphism (AFLP) analysis of the same samples. For example, the AFLP study concluded that 6.3% of the trees were hybrids, or 1.4 times higher than determined by this study. Of the 12 hybrids identified in

the AFLP study, 6 were not identified as hybrids here, and of the 8 hybrids identified here, only 4 were identified in the AFLP study. Although similar in overall results, we suggest the microsatellite analysis is more convincing than the AFLP analysis, because microsatellites provide more information per genetic locus than do AFLPs.

Introduction

(The majority of the laboratory work, data analysis, and the primary writing was performed by John Stewart, author of this dissertation. Permission to reprint this work was acquired from Springer, License Number 2544840281434.)

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. Loblolly pine has an allopatric range along the Atlantic coast of Virginia, south to northern Florida and along the Gulf Coast to eastern Texas. Shortleaf pine has an allopatric range from the central Appalachian Mountains west to the Ozark Mountains of Missouri. In between, there is a large sympatric range. The Mississippi River Basin interrupts both ranges, making a clear geographic boundary between the eastern and western populations of both species (Figure 1). Loblolly pine and shortleaf pine can be crossed with each other (Schreiner 1937). In nature, the flowering times of the two species are known to occasionally overlap, depending on the particular weather of any given year (Dorman and Barber 1956), and naturally occurring hybrids have been observed (Hare and Switzer 1969; Zobel 1953).

Studies of hybrids between loblolly pine and shortleaf pine were first reported using morphological characters (Mergen et al 1965; Cotton et al 1975; Hicks 1973; Abbot 1974). These characters are easy to observe and often seem to show strong correlations to the hybrid state, but environmental factors easily alter individual phenotypes, making morphological traits inconsistent. Also, only three states were clearly observable: loblolly pine, shortleaf pine, and the F_1 hybrid. Backcrosses were generally not identifiable with morphological measurements, due to their inherent variation and mean trending nearer one parent or the other. Studies using isoenzymes have yielded

more discrete data (Huneycutt and Askew 1989; Raja et al 1997; Chen et al 2004; Edwards-Burke et al 1997). The isoenzyme isocitrate dehydrogenase (IDH) was reported by Huneycutt and Askew, 1989, to be a predictor of hybrids between shortleaf pine and loblolly pine, since there are two alleles at the IDH locus that appeared to show complete correlation with species identities. However, the IDH locus is of limited use for identifying later generation hybrids and backcrosses because of normal recombination and Mendelian segregation.

More recent studies on loblolly pine and shortleaf pine hybrids have focused on DNA molecular markers. Xu et al (2008a and 2008b) reported on hybridization in a study of samples from the Southwide Southern Pines Seed Source Study (SSPSSS) using amplified fragment length polymorphism (AFLP) markers combined with the IDH marker. A majority of the hybrids identified in this study were not heterozygous at the IDH locus, while some clearly pure species samples were heterozygous, meaning that the IDH locus is insufficient for the identification of hybrids between loblolly pine and shortleaf pine. Additionally, all trees identified as having hybrid origin were not F_1 hybrids, indicating that introgression has been occurring for multiple generations. In the study reported here, the same collection of trees that Xu et al (2008a and 2008b) analyzed with dominant AFLP markers was characterized with codominant microsatellite (or short sequence repeats, SSRs) markers developed from both genomic and expressed DNA.

Materials and Methods

Leaf Tissue Samples

Leaf samples were collected from trees planted in the Southwide Southern Pines Seed Source Study (SSPSSS). The SSPSSS plantings contain trees from seed collected across the southeastern United States in 1951 and 1952, representing pine populations in existence prior to widespread forest management. The loblolly pine samples originate from the following, SSPSSS sources, counties and states (number of trees sampled in parentheses):

303, Onslow County, NC (10 trees)

307, Newberry County, SC (10 trees)

311, Clarke County, GA (10 trees)

317, Clay County, AL (11 trees)

321, Prentiss County, MS (10 trees)

323, Livingston Parish, LA (10 trees)

327, Clark County, AR (11 trees)

329, Hardeman County, TN (10 trees)

331, Spalding County, GA (10 trees.).

Two other populations of equivalent age were collected, one from McCurtain County, OK (11 trees) and the second from Hernando and Citrus Counties, FL (10 trees). The shortleaf pine samples originated from the following SSPSSS sources, counties and states (number of trees sampled in parentheses):

401, 451, Franklin County, PA (4 trees in one sample, 10 trees in another, collected in 1951 and 1955, respectively)

419, Lafayette County, MS (5 trees)

421, St. Helena Parish, LA (5 trees)

423, Angelina County, TX (7 trees)

433, Dent County, MO (8 trees)

435, Morgan County, TN (9 trees)

461, Clarke County, GA (8 trees)

475, Cherokee County, TX (10 trees)

477, Pushmataha and McCurtain Counties, OK (8 trees)

481, Ashley County, AR (7 trees)

487, Anderson County, TN (9 trees).

Only seven of the original ten trees in population 481 from Xu et al (2008A and 2008B) were used in this study, because the DNA quality had degraded.

Map locations of the pine populations sampled are shown in Figure 1-1. Shortleaf pine Z15 and loblolly pine SE631 and two of their contol-pollinated hybrid progeny were used as controls. Needles of Z15 were provided by Dr. Bruce Bongarten, formerly Warnell School of Forest Resources, University of Georgia and currently College of Environmental Science and Forestry, State University of New York, Syracuse, NY, USA. The control loblolly pine tree SE631 (also designated GFC-631) is from the west central piedmont of Georgia. The Z15 x SE631 hybrids and SE631 samples were obtained from Mr. Larry Lott at the Southern Institute of Forest Genetics, Saucier, MS, USA.

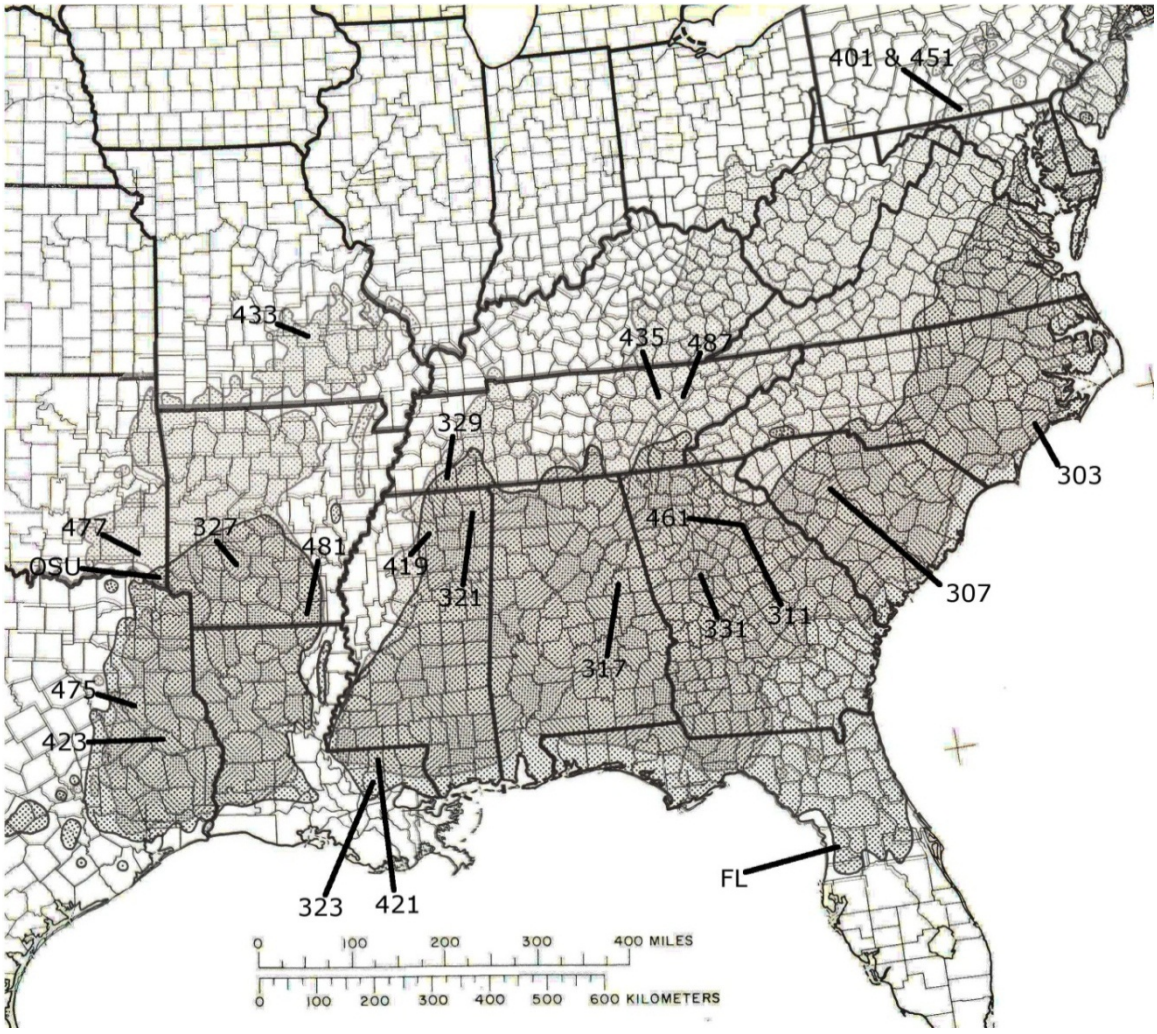

Figure 1-1: The Sources for all loblolly pine and shortleaf pine used in this study. The map was adapted from Little (1971). Lighter gray areas in the north make up the allopatric range for shortleaf pine; whereas, lighter gray areas in the south make up the allopatric range for loblolly pine. The large dark gray areas are the sympatric ranges shared by the two species. Population names beginning with the digit "4" represent shortleaf pine populations, and those beginning with the digit "3," as well as OSU and FL, represent loblolly pine populations.

DNA Isolation

DNA was isolated from the shortleaf pine samples, the SE631 loblolly pine control, the Z15 shortleaf pine control, and the Z15 x SE631 hybrids Hy1 and Hy2 using a modified CTAB protocol

(Doyle and Doyle 1988). Loblolly pine DNA was prepared using the Qiagen DNEasy Plant Minikit (Qiagen, Valencia, CA). For the CTAB protocol ten grams of frozen needles were ground into a fine powder in a mortar and pestle with liquid nitrogen. The powdered tissue was gently suspended in 100 mL 4°C CTAB extraction buffer (50 mM Tris, pH 8.0, 5 mM EDTA, 0.35 M sorbital, 10% PEG 4000, 0.1% BSA, and 0.1% β-mercaptoethanol.) The suspension was filtered through 4 layers of cheese cloth and 1 layer of miracloth with a Buchner funnel. The organelles were pelleted in the JA-14 rotor at 9000 RPM for 15 minutes at 4°C. The pellet was resuspended in 4°C CTAB wash buffer (50 mM Tris, pH 8.0, 25 mM EDTA, 0.35 M sorbital, and 0.1% β-mercaptoethanol) and brought to room temperature before being transferred to a 50 mL Oakridge tube. One-fifth volume of 5% sarkosyl was added, and the tube contents were mixed gently by inversion and allowed to rest at room temperature for 15 minutes. One-seventh volume of 5 M NaCl was added to the tube, and the tube contents were mixed gently by inversion. One-tenth volume of 8.6% CTAB wash buffer with 0.7 M NaCl was added, and the tube content was gently mixed by inversion. The tube was incubated at 60°C for 15 minutes, and an equal volume of 24:1 chlorform/octanol was added before the tube contents were gently mixed by inversion. The tube was centrifuged at 8000 RPM for 10 minutes at room temperature, and the aqueous phase was transferred to a fresh 50 mL tube. Twice the volume of cold 95% ethanol was added to the aqueous solution, and it was mixed by inversion. The tube was centrifuged at 8000 RPM for 10 minutes at room temperature to pellet the DNA. The supernatant was poured off, and 20 mL of 4°C 76% ethanol, 10 mM ammonium acetate was added. The tube was left on the bench-top for 20 minutes. The ethanol/ammonium acetate was poured off, and the pellet was allowed to dry. The DNA pellet was resuspended in 150 μL TE (10 mM Tris, pH 8.0, 1 mM EDTA.)

Microsatellite Markers

Forty-two microsatellite markers were used in this study— 25 had been previously confirmed to be polymorphic in shortleaf pine (Nelson et al 2007), 15 are new markers developed for loblolly pine (Echt, et al., in preparation, markers prefixed with PtSIFG) and 2 were tested and selected here from those provided in Chagne et al. 2004 (markers prefixed with SsrPt). Of the 42 markers, 24 were cloned and sequenced from genomic DNA (prefixes PtTX and RIPt) and 18 were cloned and sequenced as expressed sequence tags (prefixes RPTest, PtSIFG, and SstPt). The 42 microsatellite markers are described in Table 1-1.

The forward primer of each primer pair had an extension with the M-13 sequence, and all PCR reactions were run with LI-COR M-13 Forward (-29)/IRDye (700 or 800) primers (LI-COR Biosciences, Lincoln, NE.) Reactions of 10 μL were composed of 0.2 μL genomic DNA (approx. 0.1 μg/μL), 1 μL dNTP mix (20 mM), 1 μL 25 mM MgCl2, 0.05 μL Promega GoTaq polymerase (5 u/ μL; Promega Corporation, Madison, WI), 2 μL 5x Promega GoTaq PCR buffer (Promega Corporation, Madison, WI), 0.5 μL forward primer (1.25 μM), 0.5 μL reverse primer (5 μM), 1 μL M-13 Forward (-29)/IRDye(700 or 800) primer (0.5 μM), and 4 mL H2O. The reaction was started by heating to 94°C for 10 minutes. Then, 25 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 45 seconds were run. Next, 8 cycles of 94°C for 30 seconds, 51°C for 45 seconds and 72°C for 45 seconds were run. A final extension run at 72°C for 10 minutes completed the run. Each reaction was then mixed with 10 μL stopping buffer (96% Formamide, 20 mM EDTA, and 0.8% bromophenol blue) and heated to 94°C for 10 minutes.

The reactions were analyzed with the LI-COR 4300 DNA Analyzer (LI-COR Biosciences, Lincoln, NE.) A 25 cm x 25 cm plate and 0.25 mm spacers with 6.5% KB-Plus gel matrix (LI-COR Biosciences, Lincoln, NE) was used with a 64-well comb. TBE buffer (89 mM Tris, pH 8.3, 2 mM EDTA, and 89 mM boric acid) was used for electrophoresis buffer. In each well, 0.3 μL sample was added, and 0.3 μL IRDye 50-350 bp or 50-700 bp size standard (LI-COR Biosciences, Lincoln, NE) was loaded into some lanes. Running conditions were 1500 V, 40 W, and 40 mA.

Table 1-1: Loci. The following loci were used in this study. The numbers of alleles are the gross counts of alleles for each locus, whether it appeared one or many times. Numbers of major alleles are the counts of alleles that appeared with a frequency of at least 0.05. The last two columns show the number of major alleles at each locus that had statistically different frequencies in each species than would be expected if they were evenly distributed between the two species.

Locus	Number of alleles	Number of major alleles (frequency ≥ 0.05)	Number of major alleles in which the X^2 test had a p-value <0.05 and were biased for loblolly pine	Number of major alleles in which the X^2 test had a p-value <0.05 and were biased for shortleaf pine
RPtest9	19	4	2	2
RIPt0031	25	8	3	2
RIPt0079	15	5	2	2
RIPt0126	34	6	2	2
RIPt0165	20	7	3	0
RIPt0211	13	6	2	0
RIPt0367	29	5	3	2
RIPt0369	18	8	4	2
RIPt0388	13	6	4	2
RIPt0467	24	6	2	3
RIPt0567	20	7	4	2
RIPt0619	17	5	3	1
RIPt0629	12	5	2	2
RIPt0852	18	4	3	1
RIPt0968	24	6	4	2
RIPt0984	19	8	5	2
PtTX2123	5	4	2	2
PtTX3011	24	4	2	1
PtTX3013	13	3	1	2
PtTX3034	15	8	4	2
PtTX3052	15	3	1	1
PtTX4093	23	5	2	2
PtTX4181	37	6	3	2
PtTX4205	20	6	3	1
PtTX4228	22	7	4	3
PtSIFG_1318	5	2	1	1

PtSIFG_1008	5	1	1	0
PtSIFG_1185	6	2	1	1
PtSIFG_1166	2	2	1	1
PtSIFG_0493	5	2	0	0
PtSIFG_0440	4	2	1	0
PtSIFG_0265	2	2	1	1
PtSIFG_0587	3	2	1	1
SsrPt_ctg4487b	6	3	2	1
SsrPt_BF778306	3	2	0	0
PtSIFG_1018	3	2	0	0
PtSIFG_0424	4	3	1	2
PtSIFG_1295	3	2	1	1
PtSIFG_0437	4	3	1	0
PtSIFG_1190	6	2	1	1
PtSIFG_1207	8	4	2	1
PtSIFG_0371	10	2	1	1
IDH	2	2	1	1

Isocitrate Dehydrogenase Analysis

Data for the isocitrate dehydrogenase isoenzyme (IDH) locus were reported in Xu et al (2008a) and provided by S. Xu for use in this study. Results of the current codominant marker (i.e., SSR and IDH) analyses were compared with the dominant marker (i.e., AFLP) results reported by Xu et al. 2008a. As mentioned above Xu et al (2008a) also used the IDH data.

Population Genetics and Hybrid Analysis

General population genetic analyses were performed with the software GenAlEx 6.2 (Peakall and Smouse 2006), including calculation of phiPT (Φ_{PT}) and genetic distances.

Chi-square tests were performed for each allele with a frequency of 0.05 or greater. These tests compared the number of times the allele appeared in each species to the number of times it would be expected to appear if it had no bias for one species or the other. An allele bias index was calculated for each allele that was present at an overall frequency of 0.05 or greater. The allele bias index was developed with respect to loblolly pine and was defined as AL/(AL+AS) where AL= allele

frequency in loblolly pine and AS = allele frequency in shortleaf pine. A loblolly allele bias index of 1 indicates an allele that only appears in loblolly pine, and an index of 0 indicates an allele that only appears in shortleaf pine.

Structure version 2.3.2 (Pritchard et al 2000; Falush et al 2003) was used to determine hybrid identities as described by Xu et al. 2008a. Structure uses a Markov Chain Monte Carlo formula to calculate the likelihood that an individual is a member of each of k populations, where k is determined by the user. In this study, k was set equal to 2 as previously found appropriate using an analytical method with AFLP data (Xu et al 2008a). The Structure conditions used were an admixture ancestry model with correlated allele frequencies. Twenty thousand burn-in repetitions and 50,000 repetitions after burn-in were used. Hybrids were reported when predicted genome proportion levels were between 0.9531 and 0.0469, about what is expected for trees in a third backcross generation (Xu et al 2008a).

Results

Genetic Diversity

The molecular variability of both shortleaf pine and loblolly pine is primarily within population variation. PhiPT (Φ_{PT}), similar to F_{ST}, is a measure of the proportion of variation among populations to variation among and within populations. Using both genomic and EST derived SSR marker data and the IDH data from Xu et al (2008a), we calculated Φ_{PT} values for different metapopulations of the two species (Table 1-2). At the species level, shortleaf pine Φ_{PT} was 0.080 and loblolly pine Φ_{PT} was 0.061. Within species at the west vs. east of the Mississippi River level, a trend of more outcrossing in the west than in the east was detected in loblolly pine, $\Phi_{PT} = 0.010$ in the west and $\Phi_{PT} = 0.054$ in the east. In shortleaf pine, the two measures yielded roughly equivalent numbers—western Φ_{PT} was 0.076 and eastern Φ_{PT} was 0.082. Table 1-2 presents the Φ_{PT} values collected in this study, as well as for the AFLP data collected by Xu et al (2008b). AFLP generated

Φ_{PT} values (Xu et al 2008b) were very similar for loblolly pine, .059 overall and 0.019 and 0.054 for western and eastern populations, respectively. The AFLP values of the shortleaf pine populations differed more from the SSR analysis than they did for the loblolly pine populations. AFLP data yielded a similar Φ_{PT} for the eastern shortleaf populations, 0.077. However, the overall population's Φ_{PT} and that of the western population were different from the SSR analyses, 0.057 and 0.035, respectively.

Genetic Distance and Geographic Distance

The relation between geographic distances and genetic distances among populations was variable by species but plausibly real. The R^2 for the correlation of genetic to geographic distances for loblolly pine was 0.433 (Figure 1-2) while it was 0.168 for shortleaf pine. The AFLP results showed much lower R^2 for the geographic-genetic distance correlations—0.020 for loblolly pine and 0.053 for shortleaf pine (Figures 1-2 and 1-3).

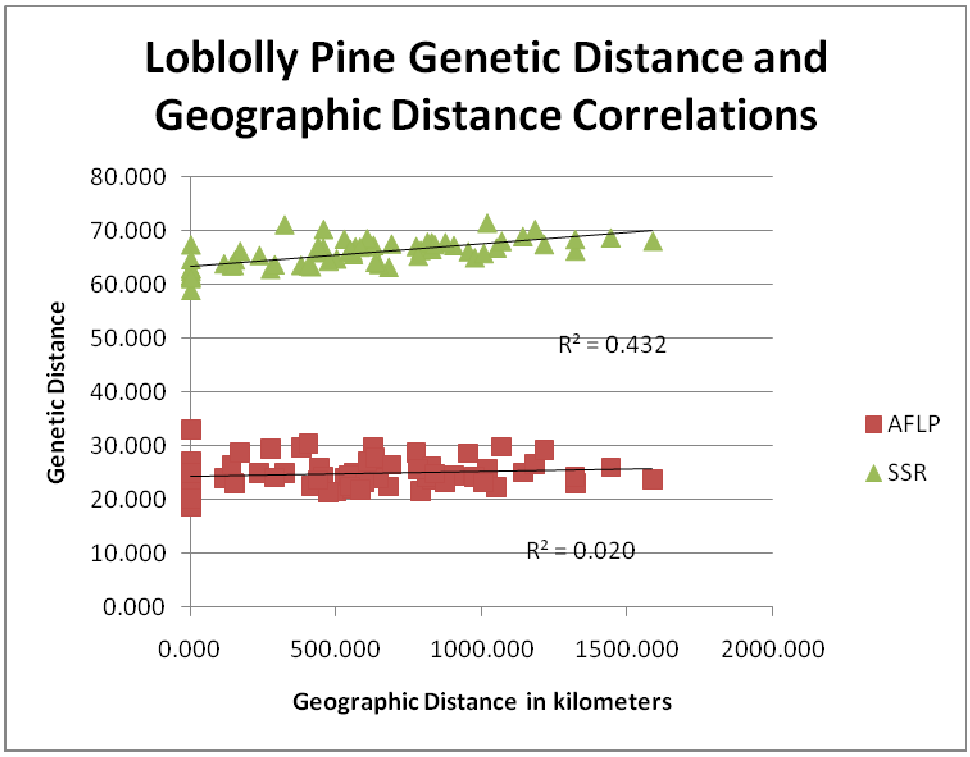

Figure 1-2: Loblolly Pine Genetic Distance vs. Geographic Distance. The genetic distances and geographic distances among all pairs of loblolly pine populations were compared. The combined markers indicate the SSR markers, EST-SSR markers, and the IDH marker used in this study. The AFLP data comes from a recalculation of genetic distances using the data first presented in Xu et al (2008a).

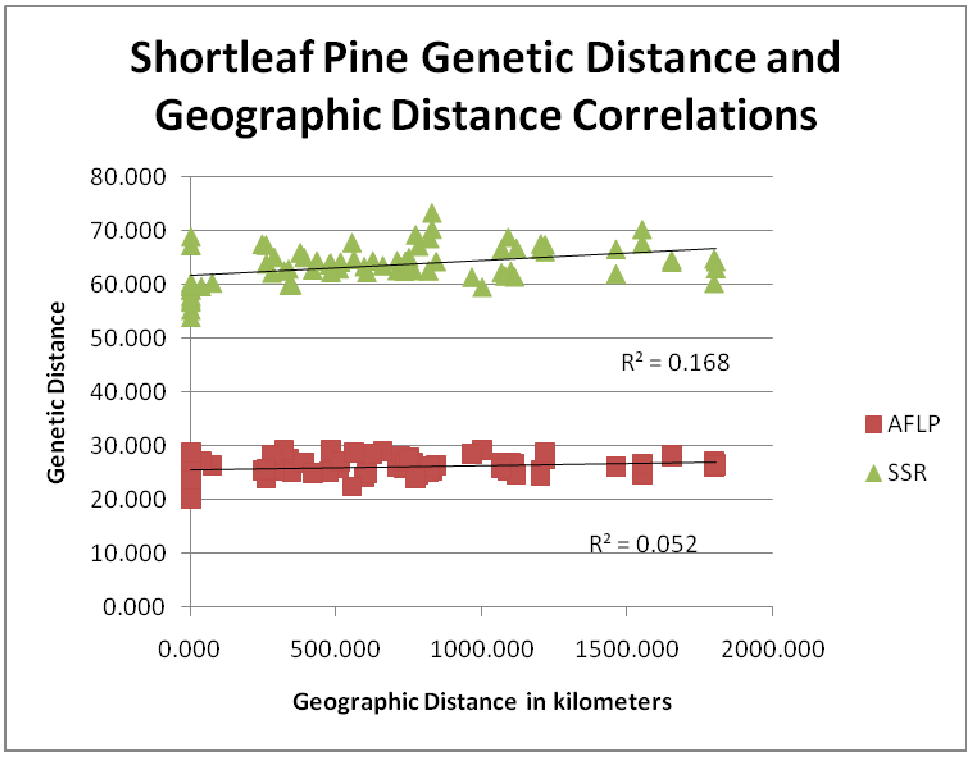

Figure 1-3: Shortleaf Pine Genetic Distance vs. Geographic Distance. The genetic distances and geographic distances among all pairs of shortleaf pine populations (sans shortleaf pine from Ashley County, Arkansas) were compared. The combined markers indicate the SSR markers, EST-SSR markers, and the IDH marker used in this study. The AFLP data comes from a recalculation of genetic distances using the data first presented in Xu et al (2008a).

Chi-Square Test and Bias Indices

In total, 575 different alleles were observed in the 43 codominant markers, an average of 13.4 alleles per locus. There were 182 major alleles for all of the loci, or 4.2 major alleles per locus. Eighty-seven major alleles showed bias for loblolly pine, 2.0 such alleles per locus, and 56 major alleles were biased for shortleaf pine, 1.3 alleles per locus. Thirty-nine major alleles were unbiased, for an average of 0.9 unbiased major alleles per locus. See Table 1-1 for details about each locus.

Hybrid Pines

Of the 202 pines examined, 8 were determined to be hybrids (Table 1-2). More hybrids were detected in the western populations (8.1%) than in the eastern populations (2.1%), and both species were similar in the proportion of individuals in stands sampled that have hybrid character (4.5% in loblolly pine populations and 3.3% in shortleaf pine populations.) In shortleaf pines, there were no hybrids detected in the east, but 7.5% of the western shortleaf pines were hybrids, indicating a much stronger tendency to cross with loblolly pine in the west. Loblolly pines showed a similar, though less pronounced trend. Eastern loblolly pines showed 3.3% hybrids, and 9.1% of western loblolly pines were hybrids. Of the four trees that were heterozygous at the IDH locus, only one of them was identified as a hybrid by SSR (Table 1-3). This tree (433-2) was also identified as a hybrid by AFLP analysis. The other three trees heterozygous at the IDH locus were not identified as hybrids by either SSR or AFLP analysis. None of the trees appear to be F_1 hybrids.

Table 1-2: Metapopulation Φ_{PT} and Number of Hybrids estimated from SSR data and AFLP data. Metapopulation indicates which group of samples is represented. Total trees indicates the total number of trees in the metapopulation. Number of hybrids shows the count of hybrids determined by Structure 2.3.2, using either the codominant data from this study or the AFLP data first presented in Xu et al (2008a). ΦPT was also calculated for each group, using the same data sets.

Metapopulation	Total Number of Trees in this Study, Total Trees in Xu et al (2008a)	Number of Hybrids (%) calculated from SSR data	Number of Hybrids (%) calculated from AFLP data	Φ_{PT} calculated from SSR data	Φ_{PT} calculated from AFLP data
All Pines	202, 205	8 (4.0%)	12 (5.8%)	n/a[1]	n/a[1]
Loblolly Pine	112, 112	5 (4.5%)	4 (3.6%)	0.061	0.059
Loblolly Pine East vs. West	112, 112	n/a	n/a	0.049	0.039
Loblolly Pine West	22, 22	2 (9.1%)	1 (4.5%)	0.010	0.019
Loblolly Pine East	90, 90	3 (3.3%)	3 (2.2%)	0.054	0.054

Shortleaf Pine	90, 93	3 (3.3%)	8 (8.6%)	0.080	0.057
Shortleaf Pine East vs. West	90, 93	n/a	n/a	0.014	0.010
Shortleaf Pine West	40, 43	3 (7.5%)	5 (11.6%)	0.076	0.035
Shortleaf Pine East	50, 50	0 (0%)	3 (6.0%)	0.082	0.077

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

Discussion

The genetic diversities (Φ_{PT} = 0.061 for loblolly pine and Φ_{PT} = 0.080 for shortleaf pine) measured in this study were in accordance with previous estimates in loblolly pine and shortleaf pine. Edwards and Hamrick (1995) calculated shortleaf pine's G_{ST} as 0.085 using isoenzymes. G_{ST}, or genetic differentiation, is a measure similar to F_{ST} and Φ_{PT}. Raja (1997) measured F_{ST} for shortleaf pine collected from sites representing the species' entire range as 0.089, and Schmidtling et al (1999) reported F_{ST} to be 0.066 among populations of loblolly pine from throughout the species' range. Outcrossing species like loblolly pine and shortleaf pine are expected to have low $F_{ST}/G_{ST}/\Phi_{PT}$ values. These two species are clearly very diverse, and both species maintain their diversity largely within subpopulations. Φ_{PT} was similar in shortleaf pine (0.080) and in loblolly pine (0.061). Hedrick (1999) predicted that the high levels of microsatellite heterozygosity observed in these species, especially microsatellites with many alleles, could artificially lower G_{ST} relative to G_{ST} measured by other marker less variable types. This prediction is not a concern in this study, because Φ_{PT} was similar to previous measurements, and it was similar to the AFLP results.

Whereas the difference in Φ_{PT} between shortleaf pine populations in the east and west was small, Φ_{PT} was much lower in the west than it was in the east (0.010 and 0.054, respectively) for

loblolly pines. There are two possibilities that could account for this. First, the sample size of western loblolly populations is relatively small with 22 individual trees in two subpopulations. Second, there is greater hybridization detected in western loblolly pines than in eastern loblolly pines. More hybridization would add diversity to the populations, but it is worth noting that while there are two populations, all three hybrids in the western loblolly pine populations are from the Oklahoma population. On the surface, then, one might take the opposite point of view, that an asymmetric distribution of hybrids would increase Φ_{PT}. However, if there is more hybridization in the west in general, then we can explain this trend through a greater degree of gene flow among all western populations. Gene flow and Φ_{PT}¬ are approximately inversely related, and thus, Φ_{PT}¬ ¬¬should be smaller in the west. Notably, loblolly pines also have two refugia from two migration events in the Pleistocene, one east of the Mississippi River and one west of it (Schmidtling 2007). However, it is unlikely that biological differences in these large metapopulations would have as much of an impact on Φ_{PT} as the climate-related differences between the regions do.

When using codominant markers, genetic distance showed a significant correlation with geographic distance for both loblolly pine and shortleaf pine, whereas the AFLP analysis did not yield a significant correlation (Figures 1-3 & 1-4). Loblolly pine seeds, and the very similar shortleaf pine seeds, are capable of long-distance travel. Still, most seeds remain in the locality of the maternal trees (Williams et al 2006). Also, as prolific outcrossers, pines are able to spread genes across their ranges slowly but extensively, but pine pollen loses its viability over distance, primarily due to ultraviolet light (Bohrerova et al 2009). In pines, both the pollen and the seeds are wind-dispersed. Evidence for directional gene flow along the path of prevailing winds from west to east has been detected in loblolly pine (Al-Rahab'ah and Williams 2002). This effect should limit the uniform dispersal of loblolly pine and shortleaf pine genes. Thus, while genes will move through the ranges of loblolly pine and shortleaf pine, they do not move rapidly or evenly, a phenomenon which may explain the detectable correlation between genetic distance and geographic distance. The AFLP analysis may not have shown a significant correlation between genetic distance and geographic

distance, because there are only two possible alleles per locus, and it is a dominant marker type. These factors cause AFLP markers to show less diversity than microsatellite markers.

The combination of the X^2 test and the bias index measurements showed that there is a great deal of diversity in the markers used in this study. While there were many minor alleles, most major alleles were biased for one of the species. These biased loci were clearly important for Structure to assign individual trees to each species during analysis. These markers could also be further developed as a diagnostic tool for finding hybrid trees in nature. Such tress would have mixtures of biased alleles.

We saw similar trends in hybridization levels as have been observed in the past. Using isoenzymes, Edwards and Hamrick (1995) reported 4.6% of shortleaf pines west of the Mississippi River were hybrids, and 1.1% of shortleaf pines east of the Mississippi River were hybrids. Our data (7.5% and 0%, respectively) is in rough agreement with that study. Certainly, the trend of more hybrids west of the Mississippi river is reinforced in this study. This trend is likely best explained by greater year to year climatic fluctuations in the west, causing the pollen shedding and pollen reception times of the two species to occasionally overlap.

The results in this study are not in full agreement with a previous study using the same sample set but a different type of genetic markers (SSR vs. AFLP). The current SSR-based study showed differences in the overall rate of hybridization (4.0% vs. 6.3 %) as well as some differences in the individual trees identified as hybrids (Table 3.) Of the 8 trees identified as hybrids in this study, 4 of them were not determined to be hybrids in the AFLP-based study (Xu et al 2008b). Likewise, of the 12 trees assigned to be hybrids in that study, 6 were designated non-hybrid pines in this study. (Three of the 13 AFLP identified hybrids were not in this study due to poor DNA quality.) Because SSR markers are co-dominant and multi-allelic, and thus highly informative, we suggest they provide better predictions of hybrid status than the dominant, bi-allelic AFLP markers, especially given that we utilized a relatively large set of SSRs derived from both coding and non-coding DNA.

Table 1-3: Trees were identified as hybrids by Structure 2.2 analysis of the codominant SSR marker data and the AFLP data. Individual indicates which trees were identified as hybrids by Structure 2.3.2 analysis of the SSR marker data. AFLP structure values were generated by Structure 2.2. Values nearer to 1 indicate that the tree is more often assigned to the status of loblolly pine by the simulations, and values nearer to 0 indicate that the tree is more often assigned to the status of shortleaf pine. Values between the thresholds of 0.0469 and 0.9531 are considered hybrids. Fields with such values are marked with an "H" to indicate a hybrid.

Individual	Structure vales; SSR markers	Structure values; AFLP markers	IDH genotype	West or East of Mississippi River	County / State of Origin
303-3*	0.950H	0.900H	LL	E	Onslow / North Carolina
311-3	0.880H	0.997	LL	E	Clarke / Georgia
321-9*	0.674H	0.965H	LL	E	Prentiss / Mississippi
327-8	0.932H	0.999	LL	W	Clark / Arkansas
OSU-8	0.826H	0.995	LL	W	McCurtain / Oklahoma
433-2*	0.127H	0.265H	LS	W	Dent / Missouri
477-2	0.062H	0.004	SS	W	Pushmataha, McCurtain / Oklahoma
477-8*	0.112H	0.099H	SS	W	Pushmataha, McCurtain / Oklahoma
481-8‡	N/A	0.002	SS	W	Ashley / Arkansas
307-4	0.998	0.691H	LL	E	Clarke / Georgia
327-10	0.998	0.802H	LL	W	Clark / Arkansas
451-6	0.001	0.050H	SS	E	Franklin / Pennsylvania
451-7	0.009	0.070H	SS	E	Franklin / Pennsylvania
487-6	0.003	0.124H	SS	E	Anderson / Tennessee

481-5	0.002	0.089H	SS	W	Ashley / Arkansas
481-7‡	N/A	0.290H	SS	W	Ashley / Arkansas
481-9‡	N/A	0.214H	SS	W	Ashley / Arkansas
321-4†	0.998	0.994	LS	E	Prentiss / Mississippi
327-2†	0.996	0.989	LS	W	Clark / Arkansas
433-1†	0.002	0.008	LS	W	Dent / Missouri

* The SSR analysis and the AFLP analysis are in agreement that the individual is a hybrid.

† The individual was heterozygous at the IDH locus but not classified as a hybrid by either Structure analysis.

‡ The individuals were omitted from this study do to many missing data points.

The IDH locus, long considered a good test for detecting hybrids of loblolly pine and shortleaf pine, did not appear to be accurate in this study as also found by Xu et al (2008a). Only one of four IDH heterozygotes was classified as hybrid by the SSR or AFLP markers. Overall, most individuals in both species are homozygous for what appears to be species-specific alleles, and heterozygotes appear to be uncommon individuals found in various stands. We think that they are evidence for remnant genetic material left over from ancient hybrids that have been backcrossed too many times to be considered hybrids today. The observation is consistent with Edwards-Burke et al. (1997) in which the chloroplast DNA evidence showed that all IDH heterozygous backcrosses were sired by one species (shortleaf pine). Unexpectedly, the one IDH heterozygous individual (433-2) that was found to be a hybrid was from Dent County, Missouri, part of shortleaf pine's allopatric range. We do not have an explanation for this phenomenon, but each tree is unique with its own life story, so 433-2 may just be an aberration or a recent hybrid with a planted loblolly pine in its lineage. The allele that is strongly loblolly pine-biased could be from an ancient backcross, or it could exist in shortleaf pines as a rare allele.

Of further interest, the SSR markers used in this study have shown that the two hybrid offspring of Z15 and SE631 were not actual progeny of these two parents, although they probably are hybrid trees (data not shown). In Xu et al. (2008b), these four trees were used to provide an inheritance check as an additional criterion for selecting AFLP primers and bands for data collection. The origin of this sample identity mistake is unclear and may be lost to the ages. This error does not detract from the results of Xu et al (2008b), since the data were submitted to analyses consisting of complex models that incorporate the data de novo (i.e., without reliance on pedigree or seed source information) (Pritchard et al 2000; Falush et al 2003). However, the error may help to explain the discrepancy between results from AFLP and SSR analyses as it reduced the number of AFLP bands included in the final data set.

In conclusion, this study produced a baseline hybridization rate for shortleaf pine and loblolly pine in the southeastern United States. In the future, we will use these same SSR markers and analysis on loblolly pine and shortleaf pine samples of current natural regeneration collected from the same counties as were collected in the 1950's for the SSPSSS. We will compare the rates of hybridization, as well as basic population genetic parameters and structure results. We hope to be able to deduce how human activity in the ranges of these two species has changed the introgression patterns and populations structural dynamics.

Literature Citations

Abbot JE. 1974. Introgressive hybridization between shortleaf and loblolly pine in southeast Oklahoma. Oklahoma State University Master's Thesis, 31pp

Al-Rahab'ah MA, Williams CG. 2002. Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. Forest Ecol Manag 163: 263-271

Auckland L, Bui T, Zhou Y, Shepherd M, WilliamsC.2002. Conifer Microsatellite Handbook Corporate Press, Raleigh, N.C.

Bohrerova Z, Bohrer G, Cho KD, Bolch MA, Linden KG. 2009. Determining the viability response of pine pollen to atmospheric conditions during long-distance dispersal. Ecol App 19: 656-667

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

Chen JW, Tauer CG, Bai G, Huang Y, Payton ME, Holley AG. 2004. Bidirectional introgression between *Pinus taeda* and *Pinus echinata*: Evidence from morphological and molecular data. Can J Forest Res 34: 2508-2516

Cotton MH, Hicks RR, Jr., Flake RH. 1975. Morphological variability among loblolly and shortleaf pines of east Texas with reference to natural hybridization. Castanea 40: 309-319

Dorman KW, Barber JC. 1956. Time of flowering and seed ripening in southern pines. USDA Forest Service, Southeastern Forest Experiment Station, Station Paper 72.

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

Edwards-Burke MA, Hamrick JL, Price RA. 1997. Frequency and direction of hybridization in sympatric populations of *Pinus taeda* (Mill) and *P. echinata* (Pinaceae). Am J of Bot 84: 879-886

Edwards MA, Hamrick JL. 1995. Genetic variation in shortleaf pine, *Pinus echinata* Mill. (Pinaceae). Forest Gen 2: 21-28

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587

Hare RC, Switzer GL. 1969. Introgression with shortleaf pine may explain rust resistance in western loblolly pine. USDA Forest Service, Research Note SO-88. Southern Forest Experiment Station, New Orleans, LA. pp2

Hedrick PW. 1999. Perspective: Highly variable loci and their interpretation in Evolution and Conservation. Evolution 53: 313-318

Hicks RR Jr. 1973. Evaluation of morphological characters for use in identifying loblolly pine, shortleaf pine, and loblolly x shortleaf hybrids. Castanea 38: 182-189

Huneycutt M, Askew G. 1989. Electrophoretic identification of loblolly pine-shortleaf pine hybrids. Silvae Genet 38: 3-4

Little EL. 1971. Atlas of United States Trees, Volume 1. Conifers and Important Hardwoods. USDA Forest Service Miscellaneous Publication No. 1146. Washington, D.C. 200 Maps.

Mergen F, Stairs GR, & Snyder EB. 1965. Natural and controlled loblolly x shortleaf pine hybrids in Mississippi. Forest Sci 11: 306-314

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

Peakall, R, Smouse, PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6, 288-295

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

Raja RG, Tauer CG, Wittwer RF, Huang YH. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (*Pinus echinata*). Can J Forest Res 27: 740-749

Schmidtling RC. 2007. Genetic variation in the southern pines: evolution, migration, and adaptation following the Pleistocene. In: Kabrick, John M.; Dey, Daniel C.; Gwaze, David, eds. Shortleaf pine restoration and ecology in the Ozarks: proceedings of a symposium; 2006 November 7-9; Springfield, MO. Gen. Tech. Rep. NRS-P-15. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 28-32

Schmidtling RC, Carroll E, LaFarge T. 1999. Allozyme diversity of selected and natural loblolly pine populations. Silvae Genet 48: 35-45

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

Williams CG, LaDeau SL, Oren R, Katal GG. 2006. Modeling dispersal distances: implications for transgenic Pinus Taeda. Ecol Appl 16: 117-124

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 Genet and Genomes 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? J Forestry 51: 494-495

CHAPTER II

BIDIRECTIONAL INTROGRESSION BETWEEN LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*P. ECHINATA* MILL.) HAS INCREASED SINCE THE 1950S.

Stewart JF, Liu Y, Tauer CG, and Nelson CD

Abstract

Loblolly pine and shortleaf pine are known to hybridize. In this study we used 25 microsatellite markers to determine the hybrid status of 172 loblolly pine trees and 154 shortleaf pine trees sampled across the ranges of the two species in the southeastern United States. Estimated Φ_{PT} for current day samples was higher in both species—0.148 in loblolly pine and 0.174 in shortleaf pine—than for trees planted from seed collected from the same locations in the 1950s for the Southwide Southern Pine Seed Source Study. This increase is likely due to anthropogenic causes such as habitat fragmentation. The proportion of hybrids rose dramatically in both species, as well: 34.9% hybrids in loblolly pine populations and 51.3% hybrids in shortleaf pine populations compared to rates of 4.5% and 3.3%, respectively, in the 1950s populations. Our results suggest that shortleaf pine and remnant loblolly pine are at risk to anthropomorphic introgression.

Introduction

(The majority of the laboratory work, data analysis, and the primary writing was performed by John Stewart, author of this dissertation.)

Loblolly pine (*Pinus taeda* L.) and shortleaf pine (*P. echinata* Mill.) are important forest species that have large ranges across the southeastern United States. Loblolly pine's allopatric range extends from the Atlantic coast of Virginia, south into northern Florida, and west to eastern Texas. Shortleaf pine's allopatric range extends from the central Appalachian Mountains to the Ozark Mountains in Missouri. Between the two allopatric ranges, a large sympatric range exists (Fig. 2-1). The two species have been crossed artificially (Schreiner 1937), and natural hybrids have been observed (Hare and Switzer 1969; Zobel 1953). It is thought that hybridization is normally prevented by the different flowering time in the two species, but when the climatic conditions are right, hybridization may occur (Dorman and Barber 1956).

Early studies of shortleaf pine x loblolly pine hybrids were conducted using morphological characters (Mergen et al 1965; Cotton et al 1975; Hicks 1973), but while easy to observe, these characters are not reliable for detecting backcrosses. They are also subject to environmental influence, reducing their reliability. Isoenzymes were later employed to detect hybrids with much apparent success (Huneycutt and Askew 1989; Raja et al 1997; Chen et al 2004; Edwards-Burke et al 1997). Recent studies have used molecular markers to identify hybrids. Xu et al (2008a, b) reported hybrids in study samples from the Southwide Southern Pine Seed Source Study (SSPSSS) using amplified fragment length polymorphism (AFLP) markers. That source material was grown from seed planted from seed source collections of loblolly pine and shortleaf pine made in the 1950s. Stewart et al (2010) followed up on that study using short sequence repeat (SSR) markers, also called microsatellite markers, to identify hybrids in the same source material. While the two studies were in agreement about the level of hybridization and population differentiation, they disagreed about which pine trees were actually hybrids.

In this study, we used microsatellites to characterize the hybrid status of trees collected from current stands from the same counties that were represented in Xu et al (2008a, b) and Stewart et al (2010). The goal of this study is to compare the rates of hybridization and introgression in modern stands to those from the 1950s. Population genetics statistics were generated, and morphological data were compared to the molecular data. From the 1950s to present, the rate of hybridization and introgression in both species has increased dramatically. Introgression can be a major threat to species, even leading to extinction, and increased introgression in many species has been connected to human activities (Wolf et al 2001; Rhymer and Simberloff 1996).

Materials and Methods

Source Material

Pine needles from both species were collected by foresters in the same counties as those collected for the studies by Xu et al (2008A), Xu et al (2008B), and Stewart et al (2010) i.e., the SSPSSS (Table 2-3.) Foresters were instructed to collect needles from young (approximately 10 years old or younger) naturally regenerated trees approximately 600 to 1000 feet apart from each other. The needles were cooled to 4˚C and shipped to our laboratory in Stillwater, OK. They were then stored at -70˚C.

Loblolly pine source material originated from the following locations. Populations with a superscript E are from counties east of the Mississippi River, and populations with the superscript W are from counties west of the Mississippi River. The numbers given are the number of trees sampled and the number of trees used after excluding individuals with poor data completion (i.e., trees with >80% of the data points in subsequent microsatellite analysis) as well as trees that were the wrong species for their populations:

NCLLE: Onslow County, North Carolina (20/20)

SCLLE: Newberry County, South Carolina (11/11)

GALLE (1-20): Clarke County, Georgia (20/(20)

TALE: Clay County, Alabama (11/11)

MSLLE: Prentiss County, Mississippi (20/17)

LALLE: Livingston Parish, Louisianna (10/7)

ARLLW: Clark County, Arkansas (20/19)

HarLE: Hardemon County, Tennessee (20/18)

GALLE (21-30): Spalding County, Georgia (10/10)

FLLLE: Hernando and Citrus Counties, Florida (20/16)

OKLLW: McCurtain County, Oklahoma (16/16)

Shortleaf pine source material originated from the following locations. Foresters collecting the material reported that naturally regenerated shortleaf pine saplings were difficult to find in some regions. Populations with a superscript E are from counties east of the Mississippi River, and populations with the superscript W are from counties west of the Mississippi River: The numbers given are the number of trees sampled and the number of trees used.

PASLE: Franklin County, Pennsylvania (25/23)

MSSLE: Lafayette County, Mississippi (20/13)

LASLE: Saint Helena Parish, Louisiana (10/7)

TXSLW: Cherokee County, Texas (20/17)

MOSLW: Dent County, Missouri (20/19)

MorSE: Morgan County, Tennessee (20/20)

GASLE: Clarke County, Georgia (13/9)

OKSLW: Pushmataha and McCurtain Counties, Oklahoma (20/19)

ARSLW: Ashley County, Arkansas (20/7)

AndSE: Anderson County, Tennessee (20/20)

The approximate sample locations are shown in Figure 2-1.

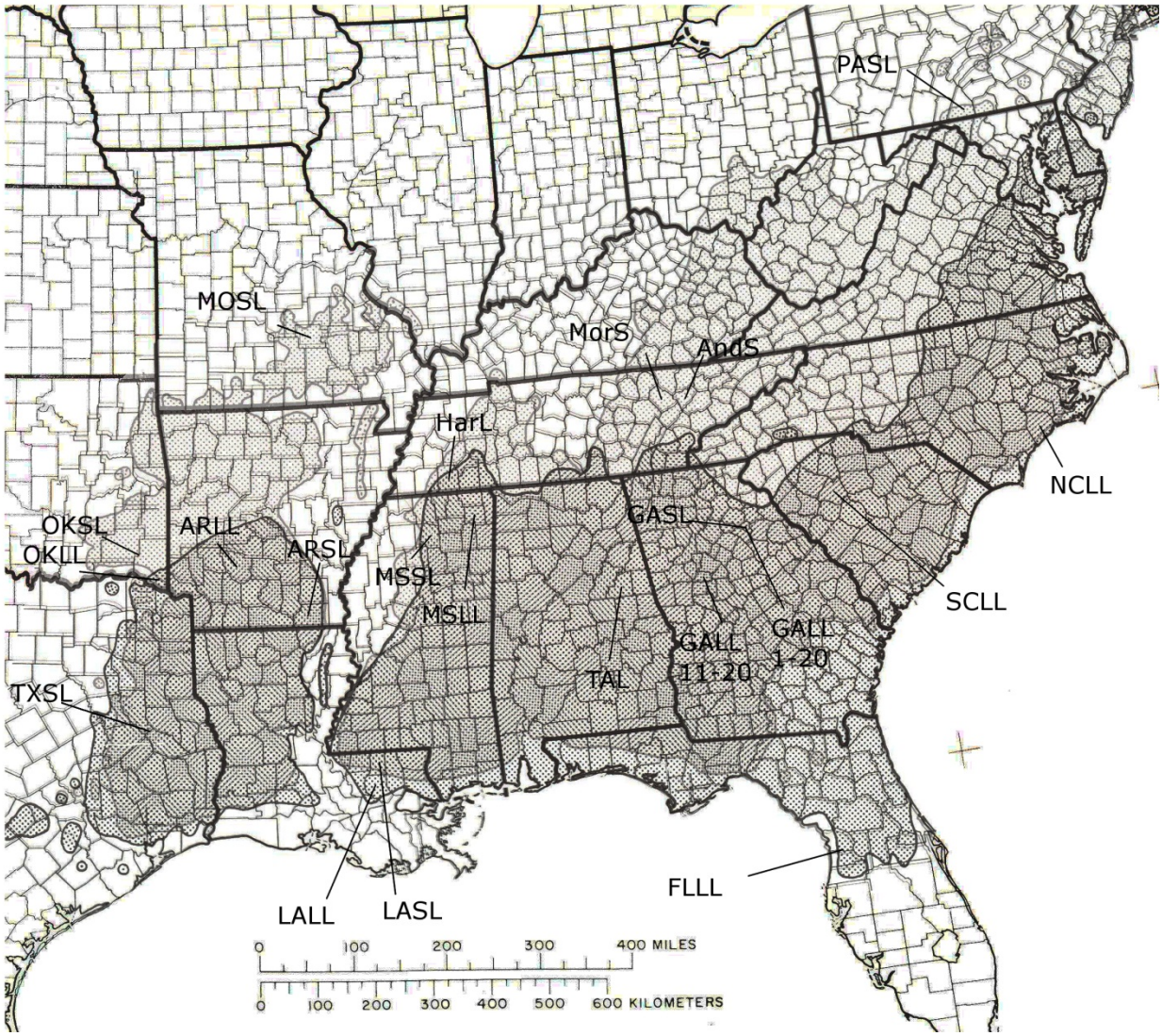

Figure 2-1: The ranges of loblolly pine and shortleaf pine, as well as the sampling locations. The dark color in the middle indicates the sympatric range of the two species, while the lighter colors indicate the allopatric ranges with shortleaf pine in the north and loblolly pine in the south. The map was adapted from Little (1971).

Morphological Measurements

Length of twenty needles of each tree were measured to the nearst millimeter, and 20 fascicles were scored for needle count, except for trees in ARLL, ARSL, MOSL, and half of

TXSL, TAL, and OKSL, in which 10 fascicles per tree were scored. We estimated average needle length and average needles per fascicle for each tree. Statistical analysis of these two variables was performed using the weighted hybrid index score (Goodman, 1967) which is a means of using multiple morphological measures to assign individuals to a species or a hybrid status. It is calculated as:

$$\sum_{i=1}^{m} \frac{(x_{ki}-u_{2i})^2}{\sigma_{ji}^2} - \frac{(x_{ki}-u_{1i})^2}{\sigma_{ji}^2},$$

where m is the number of characters observed, x_{ki} = mean of each value for character i of tree k, u_{2i} is the mean of character i for shortleaf, u_{1i} is the mean of character i for loblolly, and σ_{ji2} is the variance of character i in species j.

DNA

Pine DNA was prepared using the Qiagen DNEasy Plant Minikit (Qiagen, Valencia, CA), using a single final elution step with 150 μL TE (10 mM Tris, pH 8.0, 1 mM EDTA.)

Microsatellite Markers

Twenty-five microsatellite markers were used in this study; 11 had been previously confirmed to be polymorphic in shortleaf pine (Nelson et al 2007), 12 are new markers developed for loblolly pine (Echt, et al., in preparation, markers prefixed with PtSIFG) and 2 were tested and selected from those provided in Chagne et al. 2004 (markers prefixed with SsrPt). Of the 25 markers, 10 were cloned and sequenced from genomic DNA (prefixes PtTX and RIPt) and 15 were cloned and sequenced as expressed sequence tags (prefixes RPTest, PtSIFG, and SstPt). The 25 microsatellite markers are described in Table 2-1.

The forward primer of each primer pair had an extension with the M-13 sequence, and all PCR reactions were run with LI-COR M-13 Forward (-29)/IRDye (700 or 800) primers (LI-COR Biosciences, Lincoln, NE.) Reactions of 10 μL were composed of 0.2 μL genomic DNA

(approx. 0.1 μg/μL), 1 μL dNTP mix (20 mM), 1 μL 25 mM MgCl2, 0.05 μL Promega GoTaq polymerase (5 u/ μL; Promega Corporation, Madison, WI), 2 μL 5x Promega GoTaq PCR buffer (Promega Corporation, Madison, WI), 0.5 μL forward primer (1.25 μM), 0.5 μL reverse primer (5 μM), 1 μL M-13 Forward (-29)/IRDye(700 or 800) primer (0.5 μM), and 4 μl H2O. The reaction was started by heating to 94°C for 10 minutes. Then, 25 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 45 seconds were run. Next, 8 cycles of 94°C for 30 seconds, 51°C for 45 seconds and 72°C for 45 seconds were run. A final extension run at 72°C for 10 minutes completed the run. Each reaction was then mixed with 10 μL stopping buffer (96% Formamide, 20 mM EDTA, and 0.8% bromophenol blue) and heated to 94°C for 10 minutes.

The reactions were analyzed with the LI-COR 4300 DNA Analyzer (LI-COR Biosciences, Lincoln, NE.) A 25 cm x 25 cm plate and 0.25 mm spacers with 6.5% KB-Plus gel matrix (LI-COR Biosciences, Lincoln, NE) was used with a 64-well comb. TBE buffer (89 mM Tris, pH 8.3, 2 mM EDTA, and 89 mM boric acid) was used for electrophoresis. In each well, 0.3 μL sample was added, and 0.3 μL IRDye 50-350 bp or 50-700 bp size standard (LI-COR Biosciences, Lincoln, NE) were loaded into size marker lanes. Running conditions were 1500 V, 40 W, and 40 mA.

Table 2-1: Marker Loci. The microsatellite marker loci used in this study. The numbers of alleles are the gross counts of alleles for each locus, whether it appeared one or many times. Numbers of major alleles are the counts of alleles that appeared with a frequency of 0.05 or greater. Major alleles that were determined to be non-neutral by a X^2 test were tested by the bias index to determine to which species they were biased for.

Marker	Number of Alleles	Number of Major Alleles (Frequency ≥	Number of major alleles in which the	Number of major alleles in which the

		0.05)	X^2 test had a p-value <0.05 and were biased for loblolly pine	X^2 test had a p-value <0.05 and were biased for shortleaf pine
PtTX3052	18	4	1	2
PtTX3034	15	9	2	2
PtSIFG_0437	7	3	1	1
PtSIFG_1207	7	4	1	1
PtTX4205	19	5	1	0
RIPt0619	22	8	2	2
PtSIFG_0424	7	3	1	1
PtSIFG_0265	4	2	1	1
SsrPt_BF778306	5	2	0	0
PtSIFG_1190	7	4	2	1
PtSIFG_1295	7	3	2	1
PtTX3013	11	3	1	2
PtSIFG_1008	6	2	1	1
PtSIFG_1166	4	2	1	1
PtSIFG_0493	11	2	1	0
PtSIFG_1018	7	2	1	0
PtSIFG_0440	5	2	1	1
RPtest9	25	7	2	2
RIPt0079	15	4	1	1
RIPt0211	12	6	1	1
RIPt0388	17	7	2	2
RIPt0629	20	6	3	2
PtTX2123	6	5	1	2
SsrPt_ctg4487b	8	4	2	2
PtSIFG_0587	7	2	1	1

Population Genetics and Hybrid Analysis

General population genetic analyses were performed with the software GenAlEx 6.3 (Peakall and Smouse 2006), including calculation of phiPT (Φ_{PT}) and genetic distances. Chi-square tests were performed for each allele with a frequency of 0.05 or greater. These tests compared the number of times the allele appeared in each species to the number of times it would be expected to appear if it had no bias for one species or the other. An allele bias index was

calculated for each allele that was present at an overall frequency of 0.05 or greater. The allele bias index was developed with respect to loblolly pine and was defined as AL/(AL+AS) where AL = allele frequency in loblolly pine and AS = allele frequency in shortleaf pine. A loblolly allele bias index of 1 indicates an allele that only appears in loblolly pine, and an index of 0 indicates an allele that only appears in shortleaf pine.

Structure version 2.3.2 (Pritchard et al 2000; Falush et al 2003) was used to determine hybrid identities as described by Xu et al. (2008a). Structure uses a Markov Chain Monte Carlo formula to calculate the likelihood that an individual is a member of each of k populations, where k is determined by the user. In this study, k was set equal to 2 as previously found appropriate using Structure with AFLP data (Xu et al 2008a). The Structure conditions used were an admixture ancestry model with correlated allele frequencies. Twenty thousand burn-in repetitions and 50,000 repetitions after burn-in were used. Hybrids were reported when predicted genome proportion levels (*Q*) were between 0.9531 and 0.0469, about what is expected for trees in an F1 cross or a first through third backcross generation (Xu et al 2008a).

Results

Population Genetics

Phi-PT (Φ_{PT}) was calculated for both species and each metapopulation, i.e., each species west and east of the Mississippi River. This measure calculates the proportion of diversity that is accounted for as differences among populations. The remainder of that diversity ($\Phi_{PT} - 1$) is the diversity accounted for within populations. A higher Φ_{PT} value indicates more differentiation among the populations sampled in the study. In all but one metapopulation (Table 2-2), Φ_{PT} increased between the 1950s and the present. It was measured to be 0.148 in loblolly pine populations, 0.154 in eastern loblolly pine populations, and 0.076 in western loblolly pine populations, compared to 0.061, 0.054, and 0.010 in the respective SSPSSS loblolly pine

populations. Phi-PT was 0.174 in shortleaf pine populations, 0.189 in eastern shortleaf pine populations, and 0.064 in western shortleaf pine populations, whereas those 1950s values were 0.080, 0.082, and 0.076, respectively.

Table 2-2: Phi-PT and the hybrid count for the populations in this study, as well as the comparable populations from Stewart et al (2010).

Metapopulation	**Φ_{PT} – present day sample**	**Φ_{PT} of Pines in Stewart et al (2010) – 1950's sample**	**Hybrids / Total Trees (% Hybrids) – present day sample**	**Hybrids / Total Trees (% Hybrids) in Stewart et al (2010) – 1950's sample**
Loblolly Pines	0.148	0.061	60/172 (34.9%)	5/112 (4.5%)
Loblolly East	0.154	0.054	52/137 (38.0%)	3/90 (2.2%)
Loblolly West	0.076	0.010	8/35 (22.9%)	2/22 (9.1%)
Shortleaf Pines	0.174	0.080	79/154 (51.3%)	3/90 (3.3%)
Shortleaf East	0.189	0.082	38/92 (41.3%)	0/50 (0%)
Shortleaf West	0.064	0.076	41/62 (66.1%)	3/40 (7.5%)

Genetic distances among populations and geographic distances among populations showed interesting correlation in both species (Figure 2-2). For loblolly pine, $R^2 = 0.35$, and for shortleaf pine, $R^2 = 0.45$.

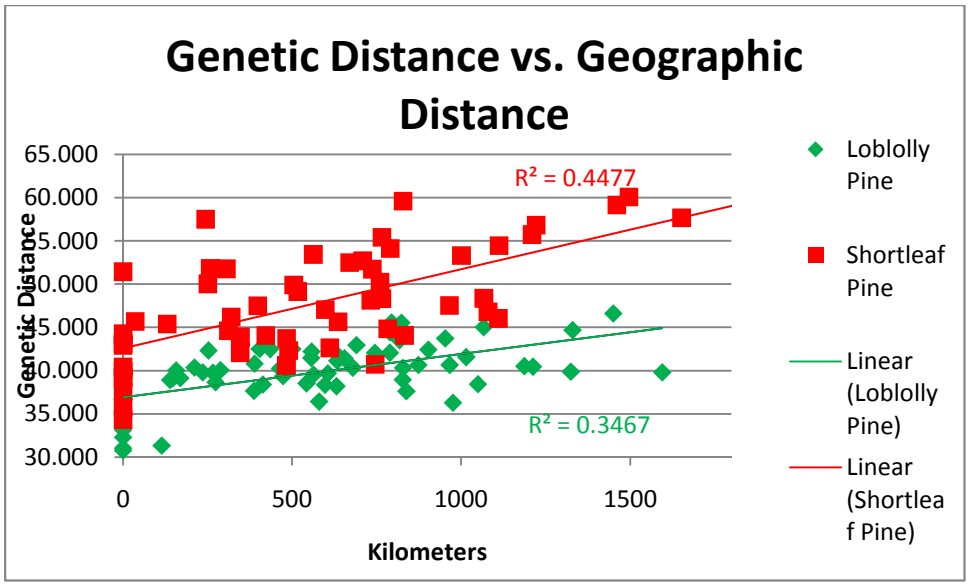

Figure 2-2: Genetic Distance vs. Geographic Distance. The above shows the genetic distances among populations plotted against the geographic distances among populations (in kilometers). The linear regression of genetic distance vs. geographic distance for each species was estimated as well. The linear regression equation for loblolly pine is y = 0.00500x + 30.78, and the R^2 value is 0.347. The linear regression equation for shortleaf pine is y = 0.00914x + 34.35, and the R^2 value is 0.448.

Introgression

Hybrids were identified through the use of Structure, a program that uses a model-based clustering method for inferring population structure. The program outputs values for each individual indicating how often it is assigned to each population. We set Structure to analyze our samples for two populations, and we used the *Q* values for loblolly pine, since those for shortleaf pine are equal to 1 - Q_{lob}. We interpret the *Q* values to represent the fraction of the trees' genomes that are loblolly pine. We defined F_1 hybrids as having *Q* values between 0.375 and 0.625. The remaining trees with *Q* values between 0.047 and 0.375 were considered to be hybrids backcrossed into shortleaf pine for up to 3 generations. Those trees with *Q* values between 0.625 and 0.953 were considered to be hybrids backcrossed into loblolly pine for up to 3

generations. These categories are useful estimates for degrees of introgression, and some individuals may be the products of different hybrids crossing with each other. Hybridization and introgression in both species was very high: 139 hybrids of 326 trees characterized, or 42.6%. In loblolly pine populations, 2.3% of the trees were F_1 hybrids, and in shortleaf pine populations, 7.1% of the trees were F_1 hybrids. Further details of the breakdown of hybrids can be found in the following tables: hybrids by metapopulation in Table 2-2, hybrids by population and hybrid degrees by metapopulation in Table 2-3.

Table 2-3: Hybrids by collection location. Each population is listed with the population that it shares a source county with in Stewart et al (2010). Hybrids shows the number of hybrids counted and Population indicates the total number of trees in that population after trees with too many missing data and trees of the wrong species were excluded from them.

Old Population Name	Hybrids / Population (%)	New Population Name	Hybrids / Population (%)
303	1/10 (10.0%)	NCLL	14/20 (70.0%)
307	0/10 (0%)	SCLL	8/11 (72.7%)
311	1/10 (10.0%)	GALL (1-20)	3/20 (15.0%)
317	0/11 (0%)	TAL	7/11 (63.6%)
321	1/10 (10%)	MSLL	4/17 (23.5%)
323	0/10 (0%)	LALL	2/7 (28.6%)
327	0/11 (0%)	ARLL	6/19 (31.6%)
329	0/10 (0%)	HarL	4/18 (22.2%)
331	0/10 (0%)	GALL (21-30)	0/10 (0%)
FL	0/10 (0%)	FLLL	3/16 (18.8%)
OSU	1/11 (9.1%)	OKLL	2/16 (12.5%)
401 & 451	2/14 (14.3%)	PASL	6/23 (26.1%)
419	0/5 (0%)	MSSL	3/13 (23.1%)
421	0/5 (0%)	LASL	6/7 (85.7%)
423&427	0/17 (0%)	TXSL	5/17 (29.4%)
433	1/8 (12.5%)	MOSL	15/19 (78.9%)
435	0/9 (0%)	MorS	5/20 (25.0%)
461	0/8 (0%)	GASL	9/9 (100%)
477	0/8 (0%)	OKSL	16/19 (84.2%)

481	1/7 (14.3%)	ARSL	5/7 (71.4%)
487	0/9 (0%)	AndS	8/20 (40.0%)

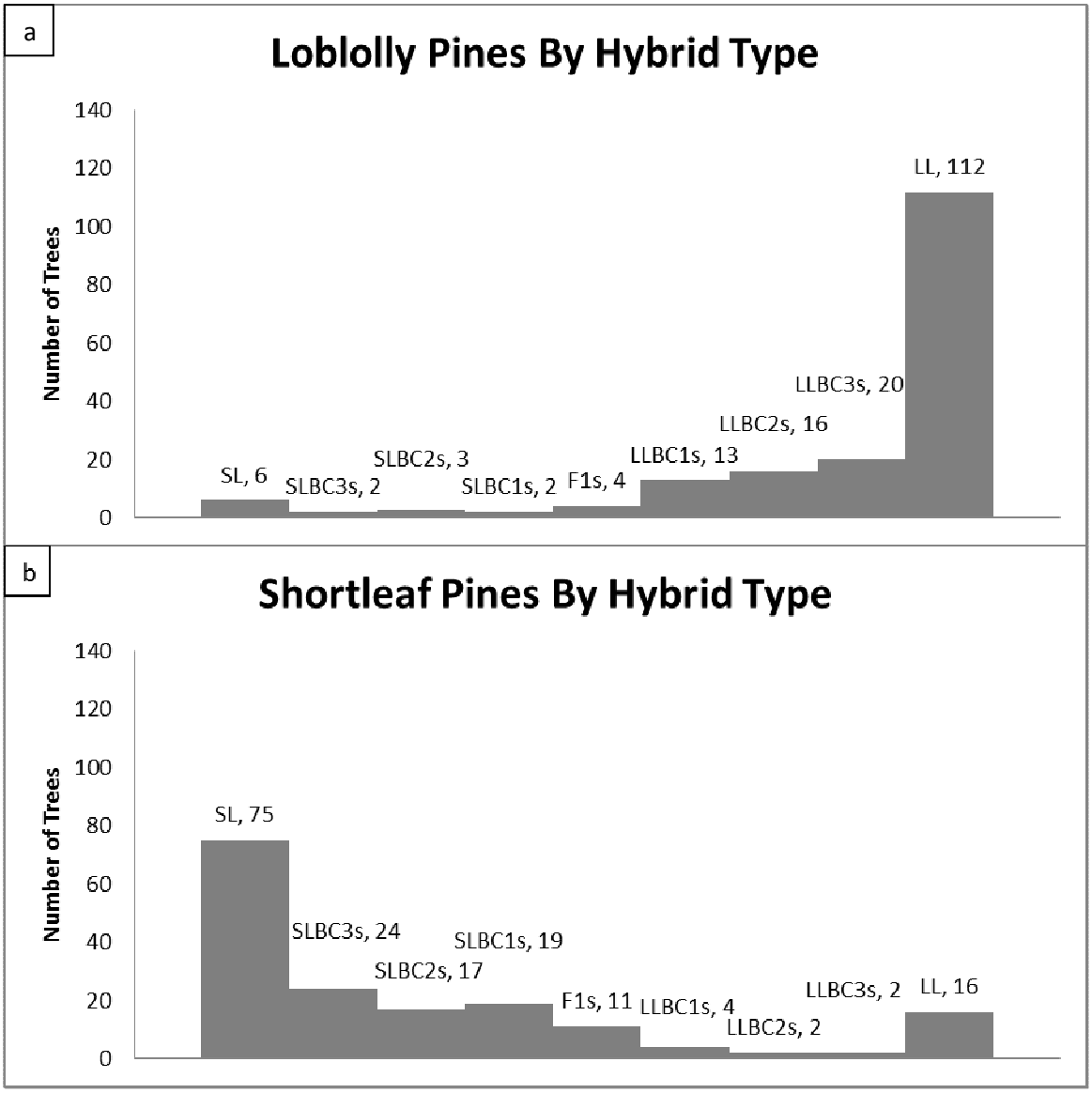

Figure 2-3: Number of pines by hybrid type. The specific categorizations of individuals based on their *Q* values derived using the program Structure is shown for each species: loblolly pine (LL) (a) and shortleaf pine (SL) (b). Trees were assigned to categories according to their *Q* values. While a tree with a value of about 0.5 could be an F_1 hybrid, it could also be an F_2 hybrid or some combination of other crosses. Likewise, all trees are assigned to these categories as basic

approximations of their hybrid status. F_1s are trees with *Q* values between 0.375 and 0.625. Loblolly pine backcrossed 1 generation trees (LLBC1s) are trees with *Q* values between 0.625 and 0.813, 2 generation trees (LLBC2s) are trees with *Q* values between 0.813 and 0.906, and 3 generation trees (LLBC3s) are trees with *Q* values between 0.906 and 0.953. Shortleaf pine backcrossed 1 generation trees (SLBC1s) are trees with *Q* values between 0.188 and 0.375, 2 generation trees (SLBC2s) are trees with *Q* values between 0.094 and 0.188, and 3 generations trees (SLBC3s) are trees with *Q* values between 0.047 and 0.094. Note that the trees included as shortleaf pines in (a) and those included as loblolly pines in (b) are not included in any of the hybrid analysis or tree totals. These individuals are likely examples of collection error, as these two species of often difficult to identify in nature, and sapling loblolly pine and shortleaf pine are especially challenging to differentiate from each other.

Morphological Measurements and Hybridization

The weighted hybrid index score showed good correlation with the Structure values (R^2 = 0.648; Figure 2-4), considering the morphological data were limited. The mean needle length and the mean needles per fascicle for each tree also indicated some correlation with the Structure values (R^2 = 0.471 and R^2 = 0.656, respectively).

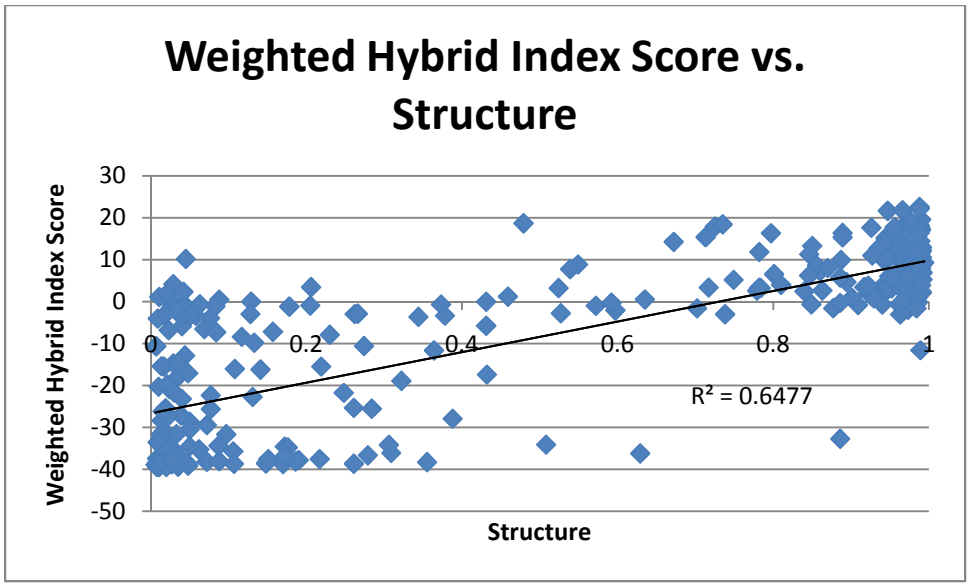

Figure 2-4: The weighted hybrid index score (Goodman, 1967) plotted on the y-axis against the values generated by Structure. These values correspond to the fraction of markers correspond to loblolly pine that the individual tree has. Numbers near zero indicate shortleaf pine, and numbers near one indicate loblolly pine.

Discussion

Φ_{PT} more than doubled in both species from the 1950s sample values (Stewart et al, 2010) to present day, increasing from 0.061 to 0.148 in loblolly pine and from 0.080 to 0.174 in shortleaf pine. In loblolly pine, the increase appeared in both the western and eastern populations, but in shortleaf pine, only the eastern populations showed an increase Φ_{PT}, while the western populations had a slight decrease Φ_{PT} (Table 2-2.) Fundamentally, these changes show a relative increase of the genetic diversity contribution of the differences among populations to the diversity contribution of the differences within populations. That is, the populations are becoming more distinct, or the diversity within populations is decreasing.

Several factors could lead to the increases in Φ_{PT} values, among them, the increase in hybridization, which is discussed below. Other factors might be habitat fragmentation, the pollen clouds generated by extensive plantations of cultivated loblolly pine, changes in selective

pressures due to severe drought, and changes in selective pressures due to the suppression of fire. Φ_{PT} and related measures of population differentiation such as F_{ST} and G_{ST} are approximately inversely related to gene flow. Habitat fragmentation reduces gene flow, so in time, fragmentation would also increase Φ_{PT}. The southeastern United States is a highly managed and developed landscape with loblolly plantations, cities, and other clearances that interrupt what was once a more contiguous forest. Studies on other wind-pollinated trees have shown that wind-blown pollen does not normally travel very far, making habitat fragmentation a more severe problem for wind-pollinated trees than for trees with animal pollinators. Knapp et al (2001) showed that acorn production in blue oaks was negatively impacted by the distance of the trees from each other on a scale of the tens of meters. Valley oak pollen successfully travels an average of about 50m (Sork et al, 2002). When comparing European beech populations in fragmented habitats to those in continuous habitats, Jump & Peñuelas (2006) found that the fixation index (F_{ST}, a measure analogous to Φ_{PT}) was higher in fragmented forests (0.029) than in continuous ones (0.010). Whereas those studies focused on angiosperms, Bohrerova et al (2009) found that loblolly pine pollen would not be viable after 24 hours in the air, primarily because of ultraviolet light, meaning that despite the long distances pollen may travel; it is only viable over a fraction of the distance travelled. Holsinger (1993) predicted that widely distributed species that dominate the ecosystems in their ranges—like loblolly pine and shortleaf pine—are at greater risk of genetic loss to habitat fragmentation than species that are naturally patchy across their ranges.

The extensive plantations of loblolly pines throughout the southeastern United States are sure to produce very large pollen clouds that can move into local stands of shortleaf pine and loblolly pine. Since plantation pines are as currently established necessarily more genetically homogenous, being siblings or even clones, they will contribute low-diversity pollen to the local pine populations. Given that different plantations have different open-pollinated loblolly pine families, each plantation could give its own genetic bias to local natural stands. The effect of this phenomenon could certainly increase Φ_{PT} across the ranges of loblolly pine and shortleaf pine.

Raja et al (1998) observed that artificial regeneration methods result in higher F_{ST} values than natural regeneration methods do in shortleaf pine.

Recent droughts in forests of the southeastern United States have been fairly severe. The most recent drought was in 2006 and 2007, but the drought from 1998 to 2002 was more severe than average. However, there were far worse droughts prior to the 1900s, according to tree ring data (Seager et al, 2009). Given that 25 markers were used with the trees in this study, it is unlikely that any of them would have a strong linkage to major drought-tolerance genes. Also, droughts are normal cyclic phenomena that loblolly pine and shortleaf pine are adapted to survive. Thus, it is improbable that drought played an important role in the increase in the Φ_{PT} values we observed. However, the southeastern region is expected to become drier and prone to more droughts in the future, due to climate change (Seager et al, 2009), a consideration that must be taken into account for the management of loblolly pines and shortleaf pines.

For a long time, forest management practices emphasized fire suppression, since fire threatened existing tree stands, as well as built structures and other areas of economic importance. Removing fire from an ecosystem adapted to its presence upsets the balance of that ecosystem. Fire is important for shortleaf pine in particular. It prepares a seedbed for seed germination and establishment and can help to control hardwoods (Walker & Wiant, 1966). Shortleaf pine and loblolly pine are both considered fire resistant, but shortleaf pine is considerably more so than loblolly pine and seedlings can sprout after top kill (Wright & Bailey, 1982). Changes in the role of fire in the ecosystem may change the rate at which each species expands or fails to do so, as each species is an early succession species (Baker & Langdon, 1990; Lawson, 1990). Such changes could in time have an effect of isolating stands from each other and contribute to increasing Φ_{PT}.

Genetic distance and geographic distance showed some correlation for both species, but the level of correlation increased in shortleaf pine from R^2=0.168 in Stewart et al (2010) to 0.448 in this study, and the level of correlation decreased in loblolly pine from R^2=0.432 in Stewart et al

(2010) to 0.347 in this study. For shortleaf pine, the increase in correlation makes sense in light of the increase in Φ_{PT}. Greater differences among populations, as measured with Φ_{PT}, should be expected to be greater among populations further away from each other. In a study of European beech using microsatellites, populations in fragmented forests showed stronger correlation between geographic distances and genetic distances than those in continuous forests did: r=0.0401 and r=0.097, respectively (Jump & Peñuelas, 2006). The increased correlation of genetic distance and geographic distance is evidence that, at least for shortleaf pine, habitat fragmentation is affecting the species genetics. How can a decrease in the genetic distance with geographic distance correlation in loblolly pine be explained? It is possible that the increasing use of genetically improved loblolly pines in plantations could be a contributing factor. Since the same genetic sources are commercially available in many states, the geographic diversity of loblolly pine could decrease (Raja et al 1998). Also, there are two primary coop programs for planting loblolly pine: the Western Gulf Forest Tree Improvement Program and the North Carolina State University Cooperative Tree Improvement Program. These organizations control and their genetic improvement programs limit the diversity of seed utilized for plantings in the western and the eastern regions, respectively. By using these two seed sources, we can expect to see reduced genetic distances among the populations with genetic material from them.

In both pine species, the rate of hybridization increased dramatically from fifty years ago (Stewart et al, 2010) to present. There were more hybrids in the present study, as well as more trees identified as being products of introgression. This increase was seen in all metapopulations (Table 2-2) and nearly all populations (Table 2-3). Introgression is a known cause of extinction of species—or, to be more precise, genomes (Allendorf et al 2001). In general, hybridization can threaten a taxon in a wide variety of ways, through the generation of poorly adapted hybrids, the generation of hybrids with greater vigor than one or more of the contributing species, or the introgressive extinction of one or more species (Simberloff 1996). Discovering whether introgression is a natural process or anthropogenic is crucial to understanding how or whether to

manage the issue (Allendorf 2001). Given the timescale for change in introgression in this study (about 50 years), it is almost certain that the cause is, at least in large part, manmade in this case.

There are three main human causes for introgression: introduction of plants and animals, habitat fragmentation, and habitat modification (Allendorf and Luikart 2007). All three could have an impact on loblolly pine, shortleaf pine, and their hybrids. While there are no known exotic pines that have been introduced and are now hybridizing with these two species, loblolly pine is being planted outside of its range, as well as being planted as a replacement for lost/harvested shortleaf pine stands, and there is evidence that shortleaf pine genes have been introgressing into the allopatric loblolly pine populations. There are high levels of introgression of loblolly pine genes into shortleaf pines in the Pennsylvania and Missouri populations. This is direct evidence that human plantings of loblolly pine or shortleaf pine x loblolly pine hybrids in these areas have affected the genetic makeup of the naturally regenerating pines there. Likewise, there were hybrids detected in the Florida loblolly pine population which is quite far from the nearest native shortleaf pine stands.

As discussed above, habitat fragmentation is common in the southeastern United States. Not only can fragmentation lead to the isolation of populations, it can lead to the mixing of previously distinct gene pools (Rhymer and Simberloff 1996). In the case of loblolly pine and shortleaf pine, habitat fragmentation could cause a greater degree of cross-pollination. Ordinarily, shortleaf pine produces male and female strobili from late March into late April, depending on the climate, and often the male strobili will emerge two weeks ahead of that (Lawson and Edwin 1990). Loblolly pine releases pollen over a two to three week period beginning February to April, again depending on the local climate. The female cones are receptive for five to seven days during the same period (Schultz 1997). The time periods for pollen flight and female cone receptivity of both species may overlap occasionally, depending on the local weather, but if habitat fragmentation isolates local populations, then female cones may become more susceptible to accepting pollen from another species in order to insure that a full

complement of ovules are fertilized. The movement of seed and seedlings from some parts of the species' ranges to other parts for the purpose of artificial regeneration may have some effect on pollination and receptivity times, as well.

Habitat modification is occurring throughout the ranges of these two species. As both species are early successional pines, they will often invade the disturbed sites generated by human development, such as roadsides and abandoned farms. These disturbances, as well as other habitat modifications, can create a corridor for the two species to more often enter each other's habitat (Rhymer and Simberloff 1996). Development has led to the creation of a web of corridors for many species, changing their frequency of contact and encouraging introgression (Wolf et al 2001). These corridors and places of contact can become hybrid zones, or regions where two species often intercross to create hybrids (Buggs 2007). In the case of loblolly pine and shortleaf pine, a large sympatric range already existed, but the two species maintained somewhat different niches with loblolly pine occurring on wetter low-lying areas and shortleaf pine commonally on well-drained upland sites. Habitat modification could be creating corridors between these habitats through disturbed sites. One other important form of habitat modification for this case is the planting of loblolly pines in shortleaf pine habitats, often as replacement trees for lost/harvested shortleaf pine stands.

A fourth cause—one that could perhaps be considered a form of habitat modification—is anthropomorphic climate change. Global temperatures have been increasing, and the ranges of loblolly pine and shortleaf pine are not excluded from the increase in temperature. One way in which climate change could impact hybridization is in its effect on flowering times in loblolly pine and shortleaf pine. Changes in springtime temperatures could effectively make the two species' flowering times overlap, causing an increase in hybridization. Another way that climate change could impact hybridization is through its effect on hybrid zones. Climate change has already been implicated in the creation and movement of hybrid zones and introgression (Buggs 2007). As species migrate in response to changing climatic conditions, their movement can

intersect with related species, creating sympatric regions in which hybrids can be generated and can thrive.

In this study, we also compared a pair of morphological traits to the Structure results. We feel that the strong correlation between weighted hybridization index and the Structure results supports our conclusions. We would like to note that when the tissue collections were made, those collecting them were not instructed to collect mature or juvenile needles and they collected them at different times of the year, so the variability of the needle length data is quite high. However, the needles per fascicle should remain unchanged from needle development to shed, and the higher correlation of the Structure results with the needle per fascicle correlation confirms that.

The ecology of loblolly pine and shortleaf pine is rapidly changing, as human activity and forest management make their marks on the distribution of these two species. It appears that hybridization and introgression are phenomena with increasing effects on both pine species, and the future of these two species is difficult to ascertain. Through habitat modification, global warming, fire suppression, seed/seedling movement, and artificial regeneration, mankind is altering the genetic makeup of loblolly pine and shortleaf pine. While it is beyond the scope of this study, management practices regarding these two species need to be reexamined to determine their ecological efficacy.

Citations

Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problems with hybrids: setting conservation guidelines. Trends Ecol Evol 16: 613-622

Auckland L, Bui T, Zhou Y, Shepherd M, WilliamsC.2002. Conifer Microsatellite Handbook Corporate Press, Raleigh, N.C.

Baker JB, Langdon OG. 1990. Pinus taeda L. loblolly pine. In: Burns RM, Honkala BH. Silvics of North America. Volume 1: Conifers. Agric. Handb. 654. Washington, DC. U.S. Department of Agriculture, Forest Service: 497-512

Bohrerova Z, Bohrer G, Cho KD, Bolch MA, Linden KG. 2009. Determining the viability response of pine pollen to atmospheric conditions during long-distance dispersal. Ecol App 19: 656-667

Buggs RJA. 2007. Empirical study of hybrid zone movement. Heredity 99: 301-312

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

Chen JW, Tauer CG, Bai G, Huang Y, Payton ME, Holley AG. 2004. Bidirectional introgression between Pinus taeda and Pinus echinata: Evidence from morphological and molecular data. Can J Forest Res 34: 2508-2516

Cotton MH, Hicks RR, Jr., Flake RH. 1975. Morphological variability among loblolly and shortleaf pines of east Texas with reference to natural hybridization. Castanea 40: 309-319

Edwards-Burke MA, Hamrick JL, Price RA. 1997. Frequency and direction of hybridization in sympatric populations of Pinus taeda (Mill) and P. echinata (Pinaceae). Am J of Bot 84: 879-886

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587

Goodman MM. 1967. The identification of hybrid plants in segregating populations. Evolution 21: 334-340

Hare RC, Switzer GL. 1969. Introgression with shortleaf pine may explain rust resistance in western loblolly pine. USDA Forest Service, Research Note SO-88. Southern Forest Experiment Station, New Orleans, LA. pp2

Hicks RR Jr. 1973. Evaluation of morphological characters for use in identifying loblolly pine, shortleaf pine, and loblolly x shortleaf hybrids. Castanea 38: 182-189

Holsinger KE. 1993. The evolutionary dynamics of fragmented plant populations. In: Kingsolver JG, Kareiva PM, Huey RB. Biotic Interactions and Global Change. Stamford, CT. Sinauer Associates. pp. 198-216

Huneycutt M, Askew G. 1989. Electrophoretic identification of loblolly pine-shortleaf pine hybrids. Silvae Genet 38: 3-4

Jump AS, Peñuelas J. 2006. Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. PNAS 103: 8096-8100

Knapp EE, Goedde MA, Rice KJ. 2001. Pollen-limited reproduction in blue oak: implications for wind pollination in fragmented populations. Oecologia 128: 48-55

Lawson ER. 1990. Pinus echinata Mill. shortleaf pine. In: Burns RM, Honkala BH. Silvics of North America. Volume 1: Conifers. Agric. Handb. 654. Washington, DC. U.S. Department of Agriculture, Forest Service: 316-326

Little EL. 1971. Atlas of United States Trees, Volume 1. Conifers and Important Hardwoods. USDA Forest Service Miscellaneous Publication No. 1146. Washington, D.C. 200 Maps.

Mergen F, Stairs GR, & Snyder EB. 1965. Natural and controlled loblolly x shortleaf pine hybrids in Mississippi. Forest Sci 11: 306-314

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

Peakall, R, Smouse, PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6, 288-295

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

Raja RG, Tauer CG, Wittwer RF, Huang YH. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (Pinus echinata). Can J Forest Res 27: 740-749

Raja RG, Tauer CG, Wittwer RF, Huang YH. 1998. Regeneration methods affect genetic variation and structure in shortleaf pin (Pinus echinata Mill.) Forest Genetics 5: 171-178

Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. Annu Rev Ecol Syst 27: 83-109

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

Seager R, Tzanova A, Nakamura J. 2009. Drought in the southeastern United States: causes, variability over the last millennium, and the potential for future hydrodynamic change. J Climate 22: 5021-5045

Sork VL, Davis FW, Smouse PE, Apsit VJ, Dyer RJ, Fernandez JF, Kuhn B. 2002. Pollen movement in declining populations of California Valley oak, Quercus lobata: where have all the fathers gone? Mol Ecol 11: 1657-1668

Stewart JF, Liu Y, Tauer CG, Nelson CD. 2010. Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (Pinus taeda L.) and shortleaf pine (Pinus echinata Mill.) Tree Genet Genomes 6: 853-862

Walker LC, Wiant HV. 1966. Silviculture of shortleaf pine. Bull. No. 0. Nacogdoches, TX. Stephen F. Austin State College, School of Forestry. 59pp.

Wolf DE, Takebayashi N, Rieseberg LH. 2001. Predicting the risk of extinction through hybridization. Conserv Biol 15: 1039-1053

Wright HA, Bailey AW. 1982. Fire ecology: United States and southern Canada. New York. John Wiley and Sons. 501pp.

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 Genet Genomes 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 Genet Genomes 4: 859-868.

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

CHAPTER III

HYBRIDIZATION IN NATURALLY REGENERATED SHORTLEAF PINE NEAR ARTIFICIALLY REGENERATED STANDS OF LOBLOLLY PINE.

Abstract

Shortleaf pine is an important timber species in the southeastern United States that is known to hybridize with loblolly pine. In this study, we used 25 microsatellite markers to measure the levels of hybridization and introgression in populations of shortleaf pine saplings and adults from the Caney Creek Wilderness Area located in the Ouachita National Forest. Frequency of hybrids ranged from 9.2% to 24.0% among the populations sampled, and we found that the level of hybridization correlates with the distances to the nearest loblolly pine plantations. However, the rates of hybridization and introgression apparently remained the same amongst the parent populations and the sapling populations. We estimated Φ_{PT} to be 0.064 in these populations, indicating a population differentiation that is normal for the species. The inbreeding coefficient (F_{IS}) was low, but 20 of the 25 microsatellite markers failed the Hardy-Weinberg Equilibrium test, indicating that the populations are changing.

Introduction

(The majority of the laboratory work, data analysis, and the primary writing was performed by John Stewart, author of this dissertation.)

Shortleaf pine (*Pinus echinata* Mill.) is an important timber species found throughout the southeastern United States. Shortleaf pine primarily occurs naturally on dry upland near neutral pH soils. The species is known to hybridize with loblolly pine (*P. taeda* L.), another very important timber species in the southeastern United States (Hare and Switzer 1969, Zobel 1953). While the two species often occur together, loblolly pine generally occurs on lowland wet sites while shortleaf pine is generally found on the dry upland sites. Both species are known to be early succession colonizers (Baker and Langdon 1990, Waggoner 1975).

Natural hybridization and introgression between these species has been measured using morphological characters (Mergen et al 1965, Cotton et al 1975, Hicks 1973), but such techniques are not very definitive, since morphology varies so much by individual, and environment plays a large role in determining morphological expression. Additionally, the precise state of introgression is difficult to measure with morphological characters; taxonomists cannot judge whether individuals are F_1s, F_2s, backcrosses, etc. Several studies have more successfully used isozymes to determine hybrid character (Huneycutt and Askew 1989, Raja et al 1997, Chen et al 2004, Edwards-Burke et al 1997). This method has become less common, however, as DNA-based markers, like amplified fragment length polymorphisms (AFLP) (Xu et al 2008a) and microsatellites (Echt and May-Marquardt 1997), have been developed for pines. DNA-based markers have numerous advantages over isozymes. Each isozyme marker requires a different set of reagents to visualize it on a gel, whereas for DNA-based markers, only the primer sets need to be different. Unlike isozymes, DNA-based markers may come from any part of the genome, within or outside of expressed genes. In general, due to simpler sets of reagents and the abundance of reagents for DNA-based markers, these markers are also less expensive.

In the studies on hybridization between loblolly pine and shortleaf pine, more hybridization has been observed west of the Mississippi River than east of it (Edwards and Hamrick 1995, Xu et al 2008a, Stewart et al 2010). Stewart et al (2010) used microsatellites and found that 4.5% of loblolly pines grown from seed collected in the 1950s were hybrids with shortleaf pine. East of the Mississippi River, 3.3% were hybrids, while west of the river, 9.1% of the trees had hybrid character. Similarly, 3.3% of the shortleaf pines, which were also grown from seed collected in the 1950s, were hybrids, 0% and 7.5% of them being hybrids east and west of the Mississippi River, respectively. Most likely, the phenomenon of more hybridization in the west is due to the different climates of these two regions as they affect pollen shed and strobili receptivity. Factors leading to introgression in loblolly pine and shortleaf pine may include the distance between different stands. Here, using microsatellites, we measured levels of hybridization and introgression in naturally regenerating shortleaf pine stands in the Caney Creek Wilderness Area in Arkansas relative to proximity to extensive plantations of loblolly pine.

Materials and Methods

Needle Source

Needles were collected from four locations in the Caney Creek Wilderness Area, Polk County, Arkansas, which is part of the Ouachita National Forest (Figure 3-1). Each collection location was on the corner of an approximate rectangle about 20 miles from east to west and 15 miles from north to south. Collection sites were thus labeled northwest (nw), southwest (sw), northeast (ne), and southeast (se). Sample sites were approximately quarter mile transects about 50 yards wide, along roads, except for the sw site, which was along a power line clearance. Twenty-five possible parent trees—trees that were at least 30 years old and labeled nwp, swp, nep, and sep, depending on the source location—were sampled from each location, and 100 naturally regenerated seedlings/saplings, which were labeled nw, sw, ne, and se, were sampled

from each location. No mature loblolly pine trees were observed in the collection areas. Needles were collected in fascicles and shipped on ice to Oklahoma State University (OSU), where they were stored at -80˚C. Needles per fascicle were counted for 35 fascicles per individual upon collection.

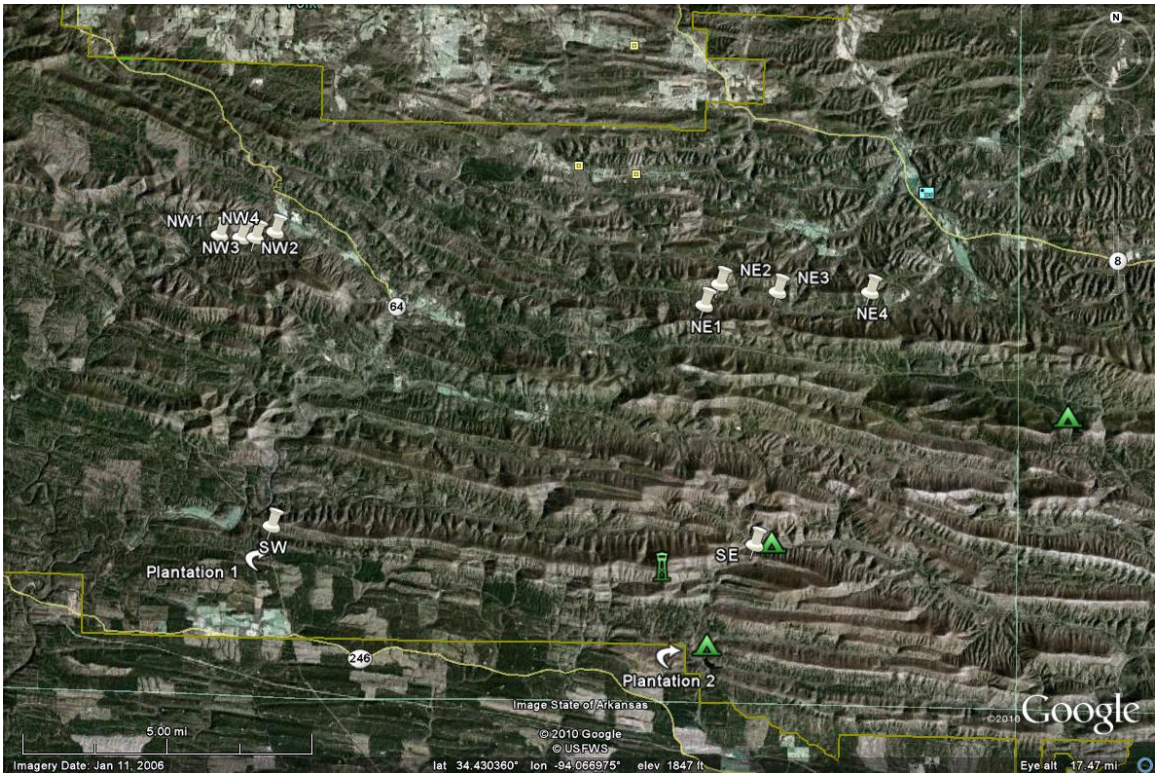

Figure 3-1: Map of collection locations. Needles from the parent populations and the sapling populations were collected from the locations shown above: NW (1-4), SW, NE (1-4), and SE. The NW and NE populations were collected over a larger area than those of the SW and SE populations. When calculating distances for the distance to the nearest plantation and the distances among populations, the means of each of the NW and NE populations were used. The site Plantation 1 is the nearest part of the plantation to the NW and SW locations, and the site Plantation 2 is the nearest part of the plantation to the NE and SE locations.

DNA Extraction

The DNA was extracted from the needle tissue through cutting the needles into small pieces and then using the Qiagen DNeasy 96 Plant Kit (Qiagen, Velencia, CA).

Microsatellite Markers

Twenty-five microsatellite markers were used in this study: of them, 11 had been previously confirmed to be polymorphic in shortleaf pine (Nelson et al 2007), 12 are new markers developed for loblolly pine (Echt, et al., in preparation, markers prefixed with PtSIFG) and 2 were tested and selected in the OSU lab from those provided in Chagne et al (2004) (markers prefixed with SsrPt). Of the 25 markers, 10 were cloned and sequenced from genomic DNA (prefixes PtTX and RIPt) and 15 were cloned and sequenced as expressed sequence tags (prefixes RPTest, PtSIFG, and SstPt). The 25 microsatellite markers are described in Table 3-1.
The forward primer of each primer pair had an extension with the M-13 sequence, and all PCR reactions were run with LI-COR M-13 Forward (-29)/IRDye (700 or 800) primers (LI-COR Biosciences, Lincoln, NE.) Reactions of 10 μL were composed of 0.2 μL genomic DNA (approx. 0.1 μg/μL), 1 μL dNTP mix (20 mM), 1 μL 25 mM MgCl2, 0.05 μL Promega GoTaq polymerase (5 u/ μL; Promega Corporation, Madison, WI), 2 μL 5x Promega GoTaq PCR buffer (Promega Corporation, Madison, WI), 0.5 μL forward primer (1.25 μM), 0.5 μL reverse primer (5 μM), 1 μL M-13 Forward (-29)/IRDye(700 or 800) primer (0.5 μM), and 4 μL H2O. The reaction was started by heating to 94°C for 10 minutes. Then, 25 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 45 seconds were run. Next, 8 cycles of 94°C for 30 seconds, 51°C for 45 seconds and 72°C for 45 seconds were run. A final extension run at 72°C for 10 minutes completed the run. Each reaction was then mixed with 10 μL stopping buffer (96% Formamide, 20 mM EDTA, and 0.8% bromophenol blue) and heated to 94°C for 10 minutes. The reactions were analyzed with the LI-COR 4300 DNA Analyzer (LI-COR Biosciences, Lincoln, NE.) A 25 cm x 25 cm plate and 0.25 mm spacers with 6.5% KB-Plus gel matrix (LI-

COR Biosciences, Lincoln, NE) was used with a 64-well comb. TBE buffer (89 mM Tris, pH 8.3, 2 mM EDTA, and 89 mM boric acid) was used for electrophoresis buffer. In each well 0.3 μL sample was added, and 0.3 μL IRDye 50-350 bp or 50-700 bp size standard (LI-COR Biosciences, Lincoln, NE) was loaded into the lanes between the sample lanes. Running conditions were 1500 V, 40 W, and 40 mA.

Table 3-1: Marker names and alleles. Markers are listed by their names (sources in the Materials and Methods section). The number of alleles refers to the gross count of different alleles scored in all individuals in this study. The number of major alleles refers to the number of alleles that have a frequency of 0.05 or greater for the marker. Expected heterozygosity (H_E) and observed heterozygosity (H_O) are also shown. The p-values from the Hardy-Weinberg Equilibrium tests are shown. Markers with values less than 0.05 are marked with an asterisk (*), meaning that they do not meet the Hardy Weinberg Equilibrium expectations.

Marker	**Number of Alleles**	**Number of Major Alleles (Frequency ≥ 0.05)**	**Expected Heterozygosity (H_E)**	**Observed Heterozygosity (H_O)**	**Hardy-Weinberg Equilibrium Test Probability**
PtTX3052	18	5	0.754	0.706	0.071
PtTX3034	15	6	0.803	0.677	0.000*
PtSIFG_0437	5	3	0.366	0.360	0.002*
PtSIFG_1207	8	5	0.698	0.621	0.000*
PtTX4205	17	5	0.795	0.606	0.000*
RIPt0619	21	7	0.858	0.488	0.000*
PtSIFG_0424	5	2	0.463	0.476	0.019*
PtSIFG_0265	7	3	0.270	0.229	0.000*
SsrPt_BF778306	3	1	0.101	0.101	0.820
PtSIFG_1190	6	3	0.374	0.322	0.001*
PtSIFG_1295	6	3	0.297	0.139	0.000*
PtTX3013	9	3	0.682	0.664	0.070
PtSIFG_1008	4	2	0.222	0.221	0.306
PtSIFG_1166	4	1	0.104	0.076	0.000*
PtSIFG_0493	9	2	0.478	0.419	0.000*
PtSIFG_1018	3	1	0.122	0.102	0.000*
PtSIFG_0440	3	2	0.369	0.356	0.821
RPtest9	30	4	0.767	0.701	0.000*
RIPt0079	17	6	0.746	0.628	0.000*
RIPt0211	15	4	0.778	0.634	0.000*
RIPt0388	18	5	0.805	0.641	0.000*

RIPt0629	21	6	0.777	0.484	0.000*
PtTX2123	6	3	0.557	0.539	0.009*
SsrPt_ctg4487b	5	2	0.382	0.225	0.000*
PtSIFG_0587	4	2	0.278	0.134	0.000*

Population Genetics and Hybrid Analysis

For all analysis, trees with more than 5 missing data points for the 25 markers were excluded from analysis. General population genetic analyses were performed with the software GenAlEx 6.3 (Peakall and Smouse 2006). Φ_{PT} was calculated using 999 iterations. Nei's genetic distance, the inbreeding coefficient (F_{IS}), expected heterozygosity (H_E), observed heterozygosity (H_O), and all Hardy-Weinberg Equilibrium statistics were also calculated using this software. Structure, version 2.3.2 (Pritchard et al 2000; Falush et al 2003), was used to determine hybrid identities as described by Xu et al. 2008a. Structure uses a Markov Chain Monte Carlo formula to calculate the likelihood that an individual is a member of each of k populations, where k is determined by the user. In this study, k was set equal to 2 so that the two populations generated by Structure would relate to the two species in this study. The Structure conditions used were an admixture ancestry model with correlated allele frequencies. Twenty thousand burn-in repetitions and 50,000 repetitions after burn-in were used. Structure yields *Q* values for each population, and in this study, the *Q* values for loblolly pine were used, since those for shortleaf pine are simply 1 minus the *Q* value for loblolly pine. Hybrids were reported when predicted genome proportion levels were between of 0.9531 and 0.0469, about what is expected for trees in a third backcross generation (Xu et al 2008a). Individuals were assigned to being F_1s and hybrids backcrossed the loblolly up to three times (*Q* = 0.375 to 0.953), loblolly pine (*Q* = 0.953 to 1), shortleaf pine (*Q* = 0 to 0.047), or hybrids backcrossed into shortleaf pine up to three times (*Q* = 0.047 to 0.375). It is recognized that individuals may have a more complex genetic history than these categories may indicate, but they are useful as a first approximation. Data representing loblolly pine and shortleaf pine trees from throughout the ranges of both species were included in the dataset to

ensure that each calculated population referred to the two species, instead of some other population structure component.

In order to test whether the proportion of hybrids in the sapling populations was different from the proportion of hybrids in the parent populations, a X^2 test was performed, using the proportions of individuals that were assigned to each hybrid category: F_1s and loblolly pine backcrosses, shortleaf pine backcrosses, and shortleaf pines. The proportions in the sapling populations were used as the test values, and the proportions in the parent populations were used as the expected values.

Results

Population Genetics

PhiPT (Φ_{PT}) was calculated across all 8 populations to be 0.064, on par with previous estimates of shortleaf pine population differentiation: Φ_{PT}, F_{ST}, and G_{ST} (Stewart et al 2010, Xu et al 2008a, Edwards and Hamrick 1995, Raja 1997). This means that of the genetic diversity in the Caney Creek Wilderness Area shortleaf pine trees, only 6.4% is among populations, while the remainder is within populations. The correlation (R^2) between geographic distances and genetic distances of the populations was 0.103, a number that is significant but low.

Average expected heterozygosity (H_E) for all populations was 0.514. For the parent populations, average H_E was 0.497, and for the sapling populations, average H_E was 0.531. Average observed heterozygosity (H_O) for all populations was 0.422, while average H_O for the parent populations was 0.409, and average H_O for the sapling populations was 0.435. The mean inbreeding coefficient (F_{IS}) was 0.176, indicated little inbreeding. H_E and H_O for each marker are also shown in Table 3-1. Of the 25 markers used in this study, 5 of them passed the Hardy-Weinberg Equilibrium test, showing that the population as a whole is in transition.

Hybridization and Introgression

All four populations of saplings showed no significant difference in the percentages of hybrids than their parent populations according to chi-square tests (Table 3-2). However, in all three of the differing populations (NW, SW, and SE), the percentage of hybrids decreased relative to the parental populations, though not significantly (Figure 3-2). One parent population (NWP) had one F_1 hybrid in it out of a population of 24, as did its offspring population (1 in 98). The sapling population SW had 2 F_1 hybrids out of 98 trees. The SW population was the closest population to the loblolly pine plantations.

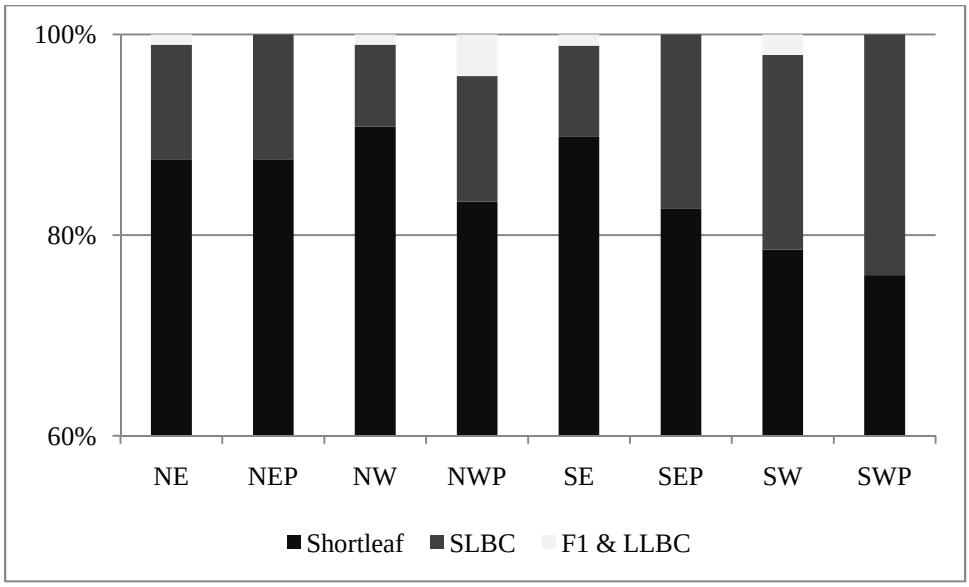

Figure 3-2: The change in hybridization rates in each population. NWP, NEP, SWP, and SEP are from representative parent populations, and NW, NE, SW, and SE are sapling populations. Loblolly pine and shortleaf pine from a separate study were included in the Structure run to provide a structural context for the two species and their hybrids. Categorization of the individuals is based on *Q* values generated by Structure. F1 & LLBC trees are trees with *Q* values between 0.375 and 0.953. Shortleaf backcrossed trees (SLBC) are trees with *Q* values between 0.047 and 0.375. Trees with *Q* values less than 0.094 are categorized as shortleaf pine.

There was a negative correlation between the geographic distance from the loblolly pine plantations to the south and the shortleaf populations sampled and the average Structure value for trees in those populations ($R^2 = 0.3017$) (Figure 3-3). Likewise, there was a negative correlation between geographic distance from the loblolly plantations to the shortleaf populations and the gross percentage of hybrids in each population ($R^2 = 0.3851$). The correlations were much stronger for the parent populations ($R^2 = 0.3912$ and R2 = 0.886, respectively).

Table 3-2: X^2 tests. Each sapling population was compared to its parent population with a X^2 test. The proportions of the categories of hybrid and non-hybrid types (shortleaf; shortleaf backcross; and F_1 and loblolly backcross) were tested. The proportions from the sapling populations were the tested values, and the proportions from the parent populations were the expected values. There were 2 degrees of freedom for each population and 11 degrees of freedom for the test of all saplings against the parent trees. In all cases, no sapling population was significantly different from the parent population.

Population	Chi-Square Test P-Value
NE	0.587
NW	0.977
SE	0.518
SW	0.127
All	0.364

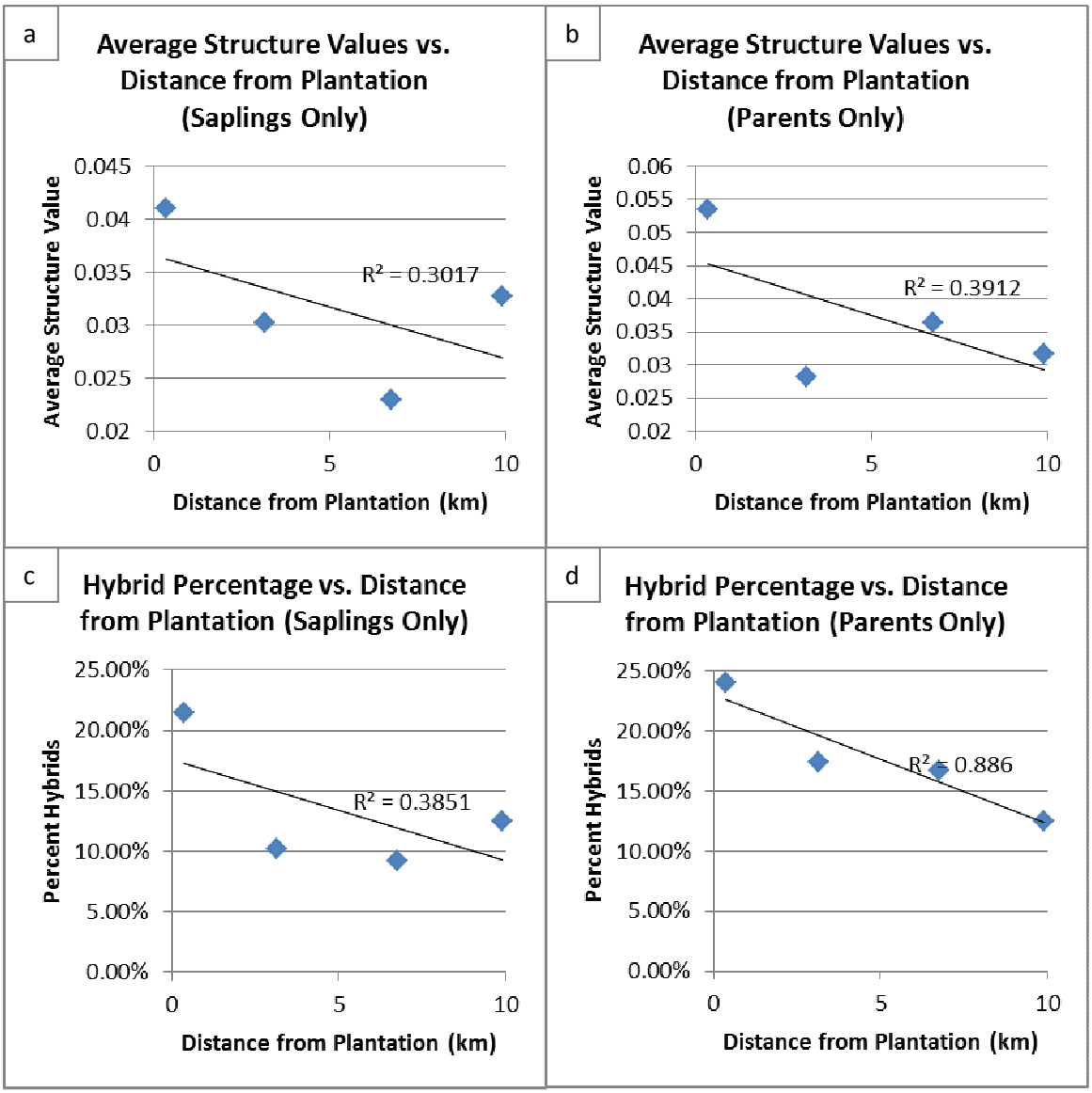

Figure 3-3: The correlation of distance from the loblolly plantations to the sample sites with the percent hybrids and the average Structure results. The correlation between the average Structure value of each population and the distance to the nearest point in the loblolly pine plantations is shown for the saplings (a) and the parents (b). Structure values of 0 indicate trees that are entirely shortleaf pine, and values of 1 indicate trees that are entirely loblolly pine. Values in between indicate different levels of hybridization. The averages were calculated as a mean of all values in a population. Hybrid percentages (the proportion of hybrids of any type) in the sapling populations showed a negative correlation with the distance from the nearest point of

the loblolly pine plantations (c). The hybrid percentages of the parents showed a strong negative correlation with distance from the nearest point of the loblolly pine plantations (d).

Morphology and Hybridization

On average, the pines in this study had 2.50 needles per fascicle with a standard deviation of 0.336, which is consistent with the common description of shortleaf pine having two or three needles per fascicle. The average needles per fascicle for each individual did not correlate with those individuals' structure values, either when all individuals were included (R^2=0.01) or when only hybrid individuals were included (R^2=0.06).

Discussion

Analysis of molecular variance calculated Φ_{PT} to be 0.064, which is similar to previous population differentiation measurements of shortleaf pine Φ_{PT}, F_{ST}, and G_{ST} (Stewart et al 2010, Xu et al 2008a, Edwards and Hamrick 1995, Raja et al 1997) as well as in other conifer species (Yeh et al 1985, Yow et al 1992, Niebling and Conkle 1990, Goncharenko et al 1992, Goncharenko et al 1993, Wu et al 1998). The Caney Creek Wilderness Area in the Ouachita National Forest is a relatively pristine shortleaf pine habitat, so unlike work done with current populations of shortleaf pine and loblolly pine across the ranges of both species (paper in progress) the shortleaf pines in the Caney Creek Wilderness have maintained expected levels of population differentiation. The correlation between genetic distance and geographic distance was significant but very low (R^2 = 0.103) likely because the distances among the populations were very small. If anything could explain the genetic distances observed among the populations, it is the different levels of hybridization among the populations relative to their distance from loblolly pine and subsequent levels of hybridization, which is discussed below.

The inbreeding coefficient (F_{IS}) for all populations was 0.176, indicating little inbreeding. This result is not unexpected, since pines, like most conifers, are outcrossing species, and the

Caney Creek Wilderness Area populations are maintained through natural regeneration and are quite extensive. However, since 20 of the 25 markers used in this study did not fulfill the predictions of the Hardy Weinberg Equilibrium model, one or more of the assumptions of the model is incorrect. Given the long distance that pine pollen can move, the likeliest assumption to be incorrect is that there is no immigration. Non-random mating is very likely to be true, as it is for most forest trees, especially wind-pollinated trees. It is probable that natural selection is occurring in these trees, as well.

We compared differences in hybridization and introgression among the populations in both time and space and found that both factors have likely played roles in the causation of interspecies crossing. There was a correlation between the distance from the nearest point in the loblolly pine plantations to the south and the level of hybridization and introgression in the shortleaf pine trees in the Caney Creek Wilderness. Correlation was stronger in the parent populations than in the sapling populations in both measures, average Structure value for the population and percent hybrids in each population.

A negative correlation between the distance of the stands from the loblolly pine plantations and the levels of hybridization and introgression is expected. Stands that are nearer to the plantations should receive a denser cloud of viable loblolly pine pollen in the spring. Pine pollen is capable of very long distance flight (Williams 2010). Also, despite its small size of about 50 μm (Williams 2008), loblolly pine pollen is highly resistant to desiccation and ultra-violet light (Bohrervora et al 2008). The distances observed in this study are short relative to the potential distance that loblolly pine pollen may travel.

Shortleaf pine and loblolly pine are receptive to and shed pollen at different times during the spring. Shortleaf pine is fertile between late March and late April (Larson and Edwin 1990), and loblolly pine is fertile between February and April (Schultz 1997). At any given location the timing separation is usually about two weeks. In both species, pollen is shed earlier in the south than in the north, and the northward position of shortleaf pine's range relative to that of loblolly

pine accounts for the two species not hybridizing more often. Currently, there are no data available for the pollen shedding and receptivity times of any loblolly pine x shortleaf pine hybrid, but it may be fair to assume that the timing would be intermediate to the parent species. However, studies have shown that hybrids between the two species are more common west of the Mississippi River than east of it, probably because climate differences between the two regions allow for more overlap in pollen shed and receptivity times (Edward and Hamrick 1995, Xu et al 2008a, Stewart et al 2010). Since the sampling site is west of the Mississippi river, hybridization is expected to be more common.

The shortleaf pine parental populations generally had more hybridization and introgression in them than their offspring did. The history of the site and the nature of the hybrids suggest that this might not be unexpected. The land south and upwind of, the Caney Creek Wilderness, the Athens Piedmont Plateau, was initially managed by Dierks Forest Industries (DFI) until 1969, when Weyerhaeuser Company purchased the land. Dierks utilized natural regeneration of the pine stands, which likely produced typical loblolly pine and shortleaf pine mixed stands found native in the Athens Piedmont Plateau (James Guldin, U.S. Forest Service, Southern Research Station, Hot Springs, AR, personal communication). In addition it is likely that Dierks favored loblolly pine in their stand management. Weyerhaeuser switched the management style to artificially regenerated loblolly pine plantations (J Guldin, personal communication). This history may explain the reduced level of hybridization in the younger trees in the Caney Creek Wilderness in a couple of ways. First, the mixed shortleaf pine-loblolly pine forests in the Athens Piedmont Plateau likely had many hybrids and backcrossed hybrids that may have shed pollen at times more concurrent with shortleaf pine than loblolly pine would have. These were replaced with pure loblolly pine, with possibly some of its genetic heritage coming from east of the Mississippi River. Second, shortleaf pine x loblolly pine hybrids appear to combine the faster growth rate of loblolly pine and the drought tolerance of shortleaf pine (paper in progress). For these reasons it may be that the parental population looked similar to the

current sapling population when it was in the sapling stage, but selective pressures favored the hybrids over time, and the percentage of hybrids will increase over time.

Citations

Auckland L, Bui T, Zhou Y, Shepherd M, WilliamsC.2002. Conifer Microsatellite Handbook Corporate Press, Raleigh, N.C.

Baker JB, Langdon OG. 1990. *Pinus taeda* L. loblolly pine. In: Burns RM, Honkala BH, technical coodinators. Silvics of North America. Volume 1. Conifers. Agricultural Handbook 654. Washington, DC: U.S. Department of Agriculture, Forest Service: 497-512

Bohrerova Z, Bohrer G, Cho KD, Bolch MA, Linden KG. 2008. Determining the viability response of pine pollen to atmospheric conditions during long-distance dispersal. Ecological Applications 19: 656-667

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

Chen JW, Tauer CG, Bai G, Huang Y, Payton ME, Holley AG. 2004. Bidirectional introgression between *Pinus taeda* and *Pinus echinata*: Evidence from morphological and molecular data. Can J Forest Res 34: 2508-2516

Cotton MH, Hicks RR, Jr., Flake RH. 1975. Morphological variability among loblolly and shortleaf pines of east Texas with reference to natural hybridization. Castanea 40: 309-319

Echt CS, May-Marquardt P. 1997. Survey of microsatellite DNA in pine. Genome: 40: 9-17

Edwards-Burke MA, Hamrick JL, Price RA. 1997. Frequency and direction of hybridization in sympatric populations of *Pinus taeda* (Mill) and *P. echinata* (Pinaceae). Am J of Bot 84: 879-886

Goncharenko GG, Padutov VE, Silin AE. 1992. Population structure, gene diversity, and differentiation in natural populations of cedar pines (*Pinus* subsect. *Cembrae*, Pinaceae) in the USSR. Plant Systematics and Evolution 182: 121-134

Goncharenko GG, Padutov VE, Silin AE. 1993. Allozyme variation in natural populations of Eurasian pines: I. Population structure, genetic variation, and differentiation in *Pinus pumila* (Pall.) Regel from Chukotsk and Sakhalin. Silvae gentica 42: 237-246

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587

Hicks RR Jr. 1973. Evaluation of morphological characters for use in identifying loblolly pine, shortleaf pine, and loblolly x shortleaf hybrids. Castanea 38: 182-189

Hare RC, Switzer GL. 1969. Introgression with shortleaf pine may explain rust resistance in western loblolly pine. USDA Forest Service, Research Note SO-88. Southern Forest Experiment Station, New Orleans, LA. pp2

Huneycutt M, Askew G. 1989. Electrophoretic identification of loblolly pine-shortleaf pine hybrids. Silvae Genet 38: 3-4

Mergen F, Stairs GR, & Snyder EB. 1965. Natural and controlled loblolly x shortleaf pine hybrids in Mississippi. Forest Sci 11: 306-314

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

Niebling CR, Conkle MT. 1990. Diversity of Washoe pine and comparisons with allozymes of ponderosa pine races. Canadian Journal of Forest Research 20: 298-308.

Peakall, R, Smouse, PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6, 288-295

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

Raja RG, Tauer CG, Wittwer RF, Huang YH. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (*Pinus echinata*). Can J Forest Res 27: 740-749

Stewart JF, Liu Y, Tauer CG, Nelson CD. 2010. Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) Tree Genet Genomes 6: 853-862

Waggoner GS. 1975. Eastern deciduous forest, Vol. 1: Southeastern evergreen and oak-pine region. Natural History Theme Studies No. 1, NPS 135. Washington, DC: U.S. Department of the Interior, Natuional Park Service. 206 p.

Williams CG. 2008. Aerobiology of *Pinus taeda* pollen clouds. Canadian Journal of Forest Research 2177-2188

Williams CG. 2010. Long-distance pine pollen still germinates after meso-scale dispersal. American Journal of Botany 97: 846-855

Wu J, Konstantin VK, Strauss SH. 1998. Abundant mitochondrial genome diversity, population differentiation and convergent evolution in pines. Genetics 150: 1605-1614

Yeh FC, El-Kassaby YA. 1980. Enzyme variation in natural populations of Sitka spruce (*Picea sitchensis*). I. Genetic variation patterns among trees from 10 IUFRO provenances. Canadian Journal of Forest Research 10: 415-422

Yeh FC, Cheliak WM, Dancik BP, Illingworth K, Trust DC, Pryhitka BA. 1985. Population differentiation in lodgepole pine, *Pinus contorta* spp. *latifolia*.: a discriminant analysis of allozyme variation. Canadian Journal of Genetics and Cytology 27: 210-218

Yow TH, Wagner MR, Wommack DE, Tuskan GA. 1992. Influence of selection for volume growth on the genetic variability of southwestern ponderosa pine. Silvae Genetica 41: 326-333

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 Genet and Genomes 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? J Forestry 51: 494-495

VITA

John Franklin Stewart

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE POPULATION GENETICS OF LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*PINUS ECHINATA* MILL.) HYBRIDIZATION AND INTROGRESSION

Major Field: Plant Science

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in your major at Oklahoma State University, Stillwater, Oklahoma in December, 2010.

Completed the requirements for the Bachelor of Science in your major at Massachusetts Institute of Technology, Cambridge, Massachusetts in 2003.

Experience:

Worked at Boston Biomedical Research Institute, Watertown, Massachusetts from 2003 to 2005.

Name: John Stewart Date of Degree: December, 2010

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: THE POPULATION GENETICS OF LOBLOLLY PINE (*PINUS TAEDA* L.) AND SHORTLEAF PINE (*PINUS ECHINATA* MILL.) HYBRIDIZATION AND INTROGRESSION

Pages in Study: 70 Candidate for the Degree of Doctor of Philosophy

Major Field: Plant Science

Scope and Method of Study: Molecular markers, population genetics, ecological genetics, hybridization, introgression, conservation

Findings and Conclusions: Results derived from loblolly pine and shortleaf pine cultivated from seeds collected from throughout their ranges in the 1950s shows that these two species have been forming natural hybrids for some time, and microsatellite results show how much introgression could be found in stands of these two species. However, these results are in some conflict with amplified fragment length polymorphism data acquired from the same data set. Both species have low Φ_{PT}, which is expected from wind-pollinated forest trees. Results from loblolly pine and shortleaf pine collected from naturally regenerating stands of today indicate that the rate of hybridization has increased in the last 50 years and that Φ_{PT} is increasing in both species. These results indicate that one or both of these species are at risk of extinction by introgression. However, there has been no significant change in the degree of introgression in the shortleaf pine stands of the Caney Creek Wilderness Area, and Φ_{PT} is normal for the populations found there, despite these stands living so close to cultivated loblolly pine plantations.

ADVISER'S APPROVAL: Charles G. Tauer