GENETIC DIVERSITY AND HYBRIDIZATION IN NATURAL STANDS OF SHORTLEAF PINE (*PINUS ECHINATA* MILL.) AND LOBLOLLY PINE (*PINUS TAEDA* L.)

By

SHIQIN XU

Bachelor of Science in Vegetable
Huazhong Agricultural University
Wuhan, Hubei, P. R. of China
1994

Master of Science in Vegetable
Huazhong Agricultural University
Wuhan, Hubei, P. R. of China
1997

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirement for the Degree of
DOCTOR OF PHILOSOPHY
December, 2006

GENETIC DIVERSITY AND HYBRIDIZATION IN NATURAL STANDS OF SHORTLEAF PINE (*PINUS ECHINATA* MILL.) AND LOBLOLLY PINE (*PINUS TAEDA* L.)

Dissertation Approved:

Dr. Charles G. Tauer
Dissertation Adviser

Dr. Bjorn Martin

Dr. Yinghua Huang

Dr. David Porter

A. Gordon Emslie
Dean of the Graduate College

ACKNOWLEDGMENT

First of all, I would like to express my sincere gratitude to my advisor, Dr. Charles G. Tauer, for his valuable research guidance and effort in the past four years. I want to thank him for having shared his expertise and writing talents with me, and for having helped me to grow professionally.

I would also like to thank my committee members, Dr. Bjorn Martin, Dr. Yinghua Huang, and Dr. David R. Porter, for spending their time and sharing their knowledge and expertise with me.

I would like to thank the work-study students: Erick Warren, Haley Smith and Shannon Adams. They helped me to finish extracting DNA, extract the seeds from the cones and with data input in my study. I also would like to thank my colleague in the Forestry Genetics Laboratory: John Stewart for his help and good suggestions. Especially I would like to thank the colleagues in Dr. Yinghua Huang's laboratory: Dr. Yanqi Wu, Angela Phillips and Lindsey Hollaway. I would like to thank Dr. Mark Payton for helping me in interpreting the analysis results. I appreciate the help from my friends: Drs. Jun Yang, Jiwang Chen, Xinkun Wang, Xi Xiong, Quan Zhang, Pingsheng Luoguan, Zheng Zou, Mrs. Ying Zhang, Xiaoping Guo and Louis Martin, for their help, good suggestions, friendship and encouragement.

Special thanks go out to my family, especially my husband, my parents and my parents-in-law, who have been supporting me through the highs and lows in my pursuit of this doctorate degree, and I am deeply indebted to them.

Finally, I would also like to acknowledge Dr Dana Nelson, USDA Forest service, Southern Institute of Forest Genetics, Saucier, MS, Oklahoma Agriculture Experiment Station and Forestry Department of OSU, for their financial support of this study.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Loblolly pine (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) have important economic significance in the southeast United States. Both species can be used for construction lumber, plywood, and many other products. Loblolly pine and shortleaf pine have broad geographic ranges, a large part of which is sympatric.

Since loblolly pine grows faster than shortleaf pine for at least the first 30 years following establishment, more and more shortleaf pine has been replaced with improved loblolly pine. The USDA Forest Service is one of only a few organizations which regenerate shortleaf pine, usually relying on natural regeneration. As a result, the shortleaf pine stands naturally regenerated by the Forest Service are becoming surrounded by more and more loblolly pine.

Previous studies (Raja et al., 1998; Chen et al., 2004) found a high level (about 15%) of hybridization between these two species in shortleaf populations in west-central Arkansas. Edwards and Hamrick (1995) found the hybridization level between these two species in shortleaf populations located west of the Mississippi River to be 4.6% and 1.1% east of the River. But the level of hybridization in shortleaf and loblolly pine populations throughout their ranges is largely unknown. If there is a consistently high hybridization level between these two species across their ranges, or in part of their ranges, the effect of such a high hybridization level on species integrity in the long term is unknown. A second issue is whether the hybridization level is increasing with

naturally regenerated shortleaf pine being surrounded by expanding loblolly pine plantings. This study will provide a reference or base level for addressing these questions since the samples collected in this study were from Southwide Southern Pine Seed Source Study (SSPSSS) plantings and the trees are from seeds collected in 1951 and 1952, when man's influence due to management was minimal.

This study has three separate chapters. In Chapter 1, the genetic diversity and structure of natural shortleaf pine populations were analyzed. In Chapter 2, the genetic diversity and structure of natural loblolly pine were studied. In Chapter 3, the hybridization level between shortleaf pine and loblolly pine in natural populations was studied.

References

Chen, J. W., Tauer, C. G., Bai, G., Huang, Y., Payton, M. E., and Holley, A. G. 2004. Bidirectional introgression between P*inus taeda* and P*inus echinata*: Evidence from morphological and molecular data. Can. J. For. Res. 34: 2508-2516.

Edwards, M. A., and Hamrick, J. L. 1995. Genetic variation in shortleaf pine, P*inus echinata* Mill. (P*inaceae*). For. Genet. 2: 21-28.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. H. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (P*inus echinata*). Can. J. For. Res. 27: 740-749.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. 1998. Regeneration methods affect genetic variation and structure in Shortleaf Pine (Pinus echinata Mill.) For. Genet. 5: 171-178.

I. GENETIC DIVERSITY AND STRUCTURE IN NATURAL STANDS OF SHORTLEAF PINE (*PINUS ECHINATA* MILL.)

1.1 Abstract

Ninety-three shortleaf pine trees from 11 seed sources were sampled from Southwide Southern Pine Seed Source Study (SSPSSS) plantings in Oklahoma and Arkansas. These samples represent shortleaf pine from seed formed in 1951 and 1952, prior to extensive forest management throughout its geographic range. Eighteen primer pairs of the 48 screened produced AFLP markers at 794 loci in these samples. The AFLP markers were used to estimate genetic diversity and structure of the shortleaf pine populations. Throughout the species, shortleaf pine was polymorphic at 65.87% (p) of the 794 loci, and had 1.66 observed alleles (na) and 1.24 effective alleles (ne) per polymorphic locus. The average heterozygosity (h) was 0.15. Western populations were a little more diverse than eastern ones. They have higher p, h, na and ne than the eastern populations. Genetic structure analysis showed 19.71% of the genetic variation existing among the 11 subpopulations, and 80.29% of the genetic variation within populations. The high value of unbiased measures of genetic identity and low value of genetic distance for all pairwise comparisons indicted that the subpopulations have similar genetic structures. The high inter-population gene flow (Nm=2.04) may explain the high similarity among the subpopulations. High gene flow (Nm=25.11) existed between eastern and western populations. Throughout the shortleaf pine range there was no apparent relationship between geographic distance and genetic distance.

1.2 Introduction

Genetic diversity is believed to be related to adaptability, and adaptability is especially important to the long-term survival of plant species (Gemmill et al, 1998). Estimates of genetic diversity and population genetic structure provide important information about natural selection and gene flow forces which shape the evolutionary dynamics of natural populations (Tarayre and Thompson, 1997) and offer a valuable reference for conservation strategies and breeding programs (Ivey and Richards, 2001).

Shortleaf pine (*Pinus echinata* Mill.) is valued for construction lumber, plywood and paper. It comprises more than 22 percent of the standing volume of the four major southern pines and it occurs naturally in 22 states (Dorman, 1976). Shortleaf pine has the broadest geographic range of the southern pines (Figure 1.1) and appears from near sea level to 3,300 feet in the southern Appalachian Mountains. It is reasonable to assume that shortleaf pine possesses a large amount of genetic variation due to adaptation to a variety of habitats.

Tauer and McNew (1985) reported considerable genetic variation in shortleaf pine populations in the state of Oklahoma using morphological characters. They reported age ten stand means for height ranged from 6.0 m to 7.5 m, diameter at breast height (DBH) from 13.9 cm to 16.8 cm and volume/tree from 36.8 dm^3 to 53.8 dm^3 in their study. Edwards and Hamrick (1995) used 14 isoenzyme markers at 22 loci and reported a high level of genetic variation (91% polymorphic loci and 2.77 alleles per locus) in 18 shortleaf pine populations sampled across its geographic range. Raja et al. (1997) used 23 isoenzyme systems at 39 loci and also found a high level of genetic variation (87.2% polymorphic loci, 2.18 alleles per locus and 2.35 alleles per polymorphic locus) in 15

shortleaf pine populations covering much of its natural range. Schmidtling et al. (2005) explored shortleaf pine geographic variation in 22 populations across its range using cortical monoterpenes and reported that all of the major terpenes showed geographic differences.

Although morphological and biochemical methods, such as isoenzyme electrophoresis techniques and measure of terpenes content, are useful in studying genetic diversity in shortleaf pine, these methods have limits. For example, morphological characters of trees are easily affected by environmental factors and biochemical methods are time-consuming, labor-intensive, expensively and/or require large amounts of plant material. Since DNA-based markers may distinguish hybrids that can not be easily discriminated by their morphological, phenological or isozyme markers, the use of DNA markers to identify hybrids and study genetic structure has rapidly developed. Some researchers are developing AFLP markers for studying population genetics and classify hybrids in trees (Muluvi et al., 1999) because this technique requires no previous sequence knowledge, has good repeatability and can detect multiple loci. In this study, we used AFLPs as DNA markers to explore genetic diversity in natural shortleaf pine populations sampled across its range.

It has been suggested that the pineless expanse of the lower Mississippi River Valley acts as a barrier to gene flow between shortleaf pine populations west and east of the River, allowing these populations to evolve separately (Schmidtling et al., 2005). Also, paleoecological data (Delcourt et al., 1983) indicate that the west and east sides of the River have been separated by the Mississippi River plain from at least the end of the last glacial epoch and that the present day populations are progeny of the individuals

from separate glacial refugia. However, Edwards and Hamrick (1995) found the west and east populations had similar genetic variation using 14 isoenzyme systems. They did report a higher level of hybridization between shortleaf pine and loblolly pine in west populations (4.6%) than east populations (1.1%) based on *IDH* (Isocitrate dehydrogenase) data. Raja et al. (1997) found the west populations (H_0=0.167) were more diverse than east ones (H_0=0.044) at *IDH* locus. In this study AFLPs were used to explore differences between shortleaf pine populations from west and east of the Mississippi River.

Previous studies (Raja et al., 1998; Chen et al., 2004) found a high level (about 15%) of hybridization between these two species in shortleaf pine populations in west-central Arkansas. Edwards and Hamrick (1995) found the hybridization level between these two species in shortleaf populations located west of the Mississippi River to be 4.6% and 1.1% east of the River. However, the current level of hybridization in shortleaf pine and loblolly pine populations throughout their ranges is largely unknown. If there is a consistently high hybridization level between these two species across their ranges or in part of their ranges, what is the effect of such a high hybridization level on shortleaf pine's integrity in the long term?

Since loblolly pine grows faster than shortleaf pine for at least the first 30 years, more and more native shortleaf pine is being replaced with plantations of improved loblolly pine. The US Forest Service is one of only a few organizations that regenerate shortleaf pine, usually relying on natural regeneration. As a result, the shortleaf pine stands naturally regenerated by the Forest Service are being surrounded more and more

by loblolly pine. Is the hybridization level increasing with naturally regenerated shortleaf pine because of the expanding loblolly pine plantings?

The samples collected in this study are from Southwide Southern Pine Seed Source Study (SSPSSS) plantings in OK and AR, and the trees in these plantings are from seeds collected in 1951 and 1952, when man's influence due to management was minimal. Thus, this study estimates genetic variation found in natural populations of shortleaf pine approximately 50 years ago, and these results will provide a reference or base level data set for addressing the above questions concerning hybridization.

1.3 Materials and Methods

Shortleaf pine and loblolly pine have broad geographic ranges and large overlapping regions. To provide a base level for estimating the effect of shortleaf pine hybridization with loblolly pine on genetic variation in shortleaf pine in the long term, shortleaf pine was sampled from allopatric and sympatric populations as shown in Figure 1.1.

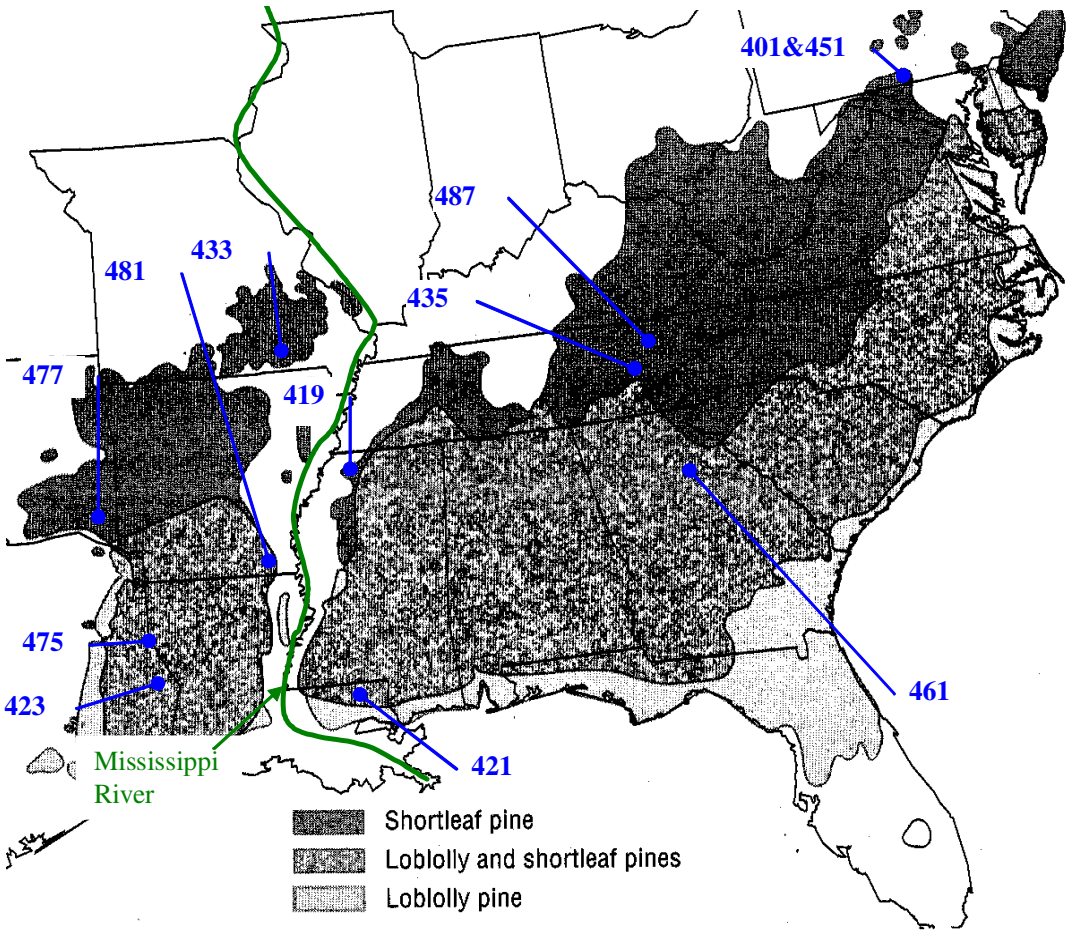

Figure 1.1 The origin of the seed source samples and natural range of shortleaf pine and loblolly pine

The numbers are seed source IDs of samples.

Shortleaf pine samples were collected from 12 seed sources in 1951 and 1952. The seed sources were created by collecting cones from 20 or more trees in each area and the resulting seeds were mixed. The locations and sample sizes of the seed sources sampled in this research are shown in Figure 1.1 and Table 1.1.

Table 1.1 The origin and sample sizes of the shortleaf pine sources sampled in this study

Source ID	State	County	No of tress
401*	PA	Franklin	4
419	MS	Lafayette	5
421	LA	St. Helena	5
423	TX	Angelina	7
433	MO	Dent	8
435	TN	Morgan	9
451*	PA	Franklin	10
461	GA	Clarke	8
475	TX	Cherokee	10
477	OK	Pushmataha & McCurtain	8
481	Ark	Ashley	10
487	TN	Anderson	9

(*401 belongs to the samples whose seeds originally collected in 1951 and 451 to the samples whose seeds originally collected in 1955, they were considered as a single source for analysis)

Needles from 93 shortleaf pine trees of the SSPSSS were sampled. These materials were collected by Oklahoma State University Kiamichi Forest Resources Center personnel, Idabel, OK 74745, USA.

When using the 4300 DNA Analyzer from LI-COR for AFLP analysis, only 64 samples can be loaded in one gel. Consequently, the remaining 29 samples had to be loaded in a second gel. To ensure the same locus was scored for all 93 samples, loblolly pine 631, shortleaf pine Z15, and two hybrids between them were used as standards or check lanes. The shortleaf pine parent Z15, was provided by Bruce Bongarten, Warnell School of Forest Resource, University of Georgia. Z15 is from North Carolina. Loblolly

pine parent 631 and the artifical hybrids (F1) between Z15 x 631 were supplied by Dana Nelson, USDA Forest Service, Southern Institute of Forest Genetics, Saucier, MS, USA. Loblolly pine 631 is from the west central piedmont of Georgia County, GA.

Needles were placed in plastic bags and kept cool with blue ice in a cooler during overnight shipment. Upon arrival in the laboratory, the needles were frozen at -80^{0}C for later use.

1.3.1 AFLP Analysis

Total DNA was extracted from needles using a modified CTAB protocol (Doyle and Doyle, 1988) used by our laboratory, as follows: Ten grams frozen needles were put into a mortar which contained a generous amount of liquid nitrogen (covered all needles). The needles were ground to a fine powder adding liquid nitrogen as needed to keep tissue frozen. The fine powder was poured into a 200 ml tube containing 100 mls cold CTAB extraction buffer (the CTAB extraction buffer has 50 mM Tris, 5 mM EDTA, 0.35 M sorbital, 10% PEG 4000, 0.1% BSA and 0.1% β-mercaptoethanol; BSA and β-mercaptoethanol were added just before using. The pH of the CTAB extraction buffer was 8.0 at 4^{0}C). The tube was shaken gently until all the fine powder was well suspended. The mixture was filtered through four layers of cheese cloth with one layer of miracloth underneath (a Buchner funnel, vacuum flask and vacuum were used). The organelles were pelleted in the JA-14 rotor at 9000 RPM for 15 minutes at 4^{0}C. The supernatant was poured off and the pellet resuspended in 5 ml of cold CTAB wash buffer (CTAB wash buffer includes 50 mM Tris, 25 mM EDTA, 0.35 M sorbital, and 0.1% β-mercaptoethanol; β-mercaptoethanol was added just before using. The pH of the CTAB wash buffer was 8.0 at 4^{0}C), brought to room temperature and transferred into a 50 ml

orkridge tube. About 1/5 volume of 5% sarkosyl was added into the tube. The tube was shaken gently by inversion and left at room temperature for 15 minutes. About 1/7 volume of 5 M NaCl was added and the tube was shaken gently by inversion. One tenth volume of 8.6% CTAB, 0.7 M NaCl solution was added and the tube was shaken gently by inversion. The tubes containing the mixture were incubated at 60^0C for 15 minutes. An equal volume of 24:1 chloroform/octanol was added and the tube was shaken gently by inversion until an emulsion was formed. The tube was centrifuged at 8000 RPM for 10 minutes at room temperature. The upper aqueous phase was transferred into a second 50 ml tube (if the aqueous layer was not clear, an equal volume of 24:1 chloroform/octanol was added to the second tube, shaken gently by inversion, and centrifuged at 8000 RPM for 10 minutes at room temperature again). A 2X volume of cold 95% ETOH was added to the second tube containing the clear aqueous layer and the tube was shaken gently by inversion to precipitate the DNA. 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 4^0C 76% ETOH, 10 mM NH_4Ac was added to the tube. The tube was left on the bench-top for 20 minutes. The ETOH, NH_4Ac was poured off and the DNA pellet dried. The DNA pellet was resuspend in about 150 ul TE buffer (the TE buffer includes 10 mM Tris with pH of 8.0 and 1 mM EDTA) .

AFLP markers were previously used by Remington et al. (1999) to construct genetic maps and by Remington and O'Malley (2000) to characterize embryonic stage inbreeding depression in loblolly pine. They used *Eco*RI and *Mse*I as the restriction digestion enzymes. From 48 primer pairs, Remington et al. (1999) found a large number of polymorphic fragments using 21 combinations of *Eco*RI (E) and *Mse*I (M) primers.

The selective nucleic acid sequences for *Eco*RI primers were 5'-ACA-3', 5'-ACC-3', 5'-ACG-3' and 5'-ACT-3'. The selective nucleic acid sequences for *Mse*I primers were 5'-CCAG-3', 5'-CCCG-3', 5'-CCGC-3', 5'-CCGG-3', 5'-CCTG-3', 5'-CCAA-3', 5'-CCAC-3', 5'-CCCA-3', 5'-CCGA-3', 5'-CCTA-3', 5'-CCTC-3' and 5'-CCTT-3'. The primers and the AFLP marker development protocols used by them were utilized in this study.

The protocols used by Remington et al. (1999) and Remington and O'Malley (2000) were modified as outlined below and used to screen shortleaf pine samples for AFLP markers:

1. DNA digestion: each reaction included 5 ul DNA (100 ng/ul), 0.25 ul rare cutter restriction endonuclase (RE) *Eco*RI (20 units/ul), 0.5 ul frequent cutter RE *Mse*I (10 units/ul), 5 ul 10X buffer for RE and 29.25 ul ddH_2O. The total volume was 40 ul. A master mix was used to ensure precision. Reactions were incubated for 2 hours at 37^0C. after which, the REs were inactivated at 70^0C for 15 minutes.

2. Ligation of adapter: each reaction included 1 ul *Eco*RI adaptor (5 pmol/ul), 2 ul *Mse*I adapter (25 pmol/ul), 1.5 ul 10X ligase buffer, 0.33 ul T_4 DNA ligase (3 unit/ul), 5.17ul ddH_2O and 40ul digestion mixture from step 1. The total volume was 50ul. A master mix was used to ensure precision. Reactions were incubated for 3 hours at 20^0C, or overnight. Then 10ul of the reaction mixture was loaded to a 1.5% agarose gel to check the digestion-ligation result. Another 10ul of reaction mixture was transferred into a new 200ul tube and 90ul H_2O was added and mixed well. The 1:10 diluted ligated mixture and undiluted portion were stored at -20^0C.

3. Pre-amplification: each reaction included 0.45ul *Eco*RI preamplification primer (100 ng/ul) and 0.45 ul *Mse*I preamplification primer (100 ng/ul), 0.6 ul 10 mM dNTPs, 3 ul 10X PCR-buffer, 1.8 ul 25mM M_gCl_2 (for buffer without M_gCl_2), 0.36 ul Taq polymerase (5unit/ul), 8.34 ul ddH_2O and 15 ul 1:10 diluted ligation mixture from step 2. The total volume was 30 ul. A master mix was used to ensure precision. The PCR program was 28 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute, then hold at 4°C. Then 10 ul of the PCR product was loaded to a 1.5% agarose gel to check the pre-amplification result. The pre-amplification PCR product was diluted 20 times (10 ul PCR products added to 190 ul water). All reaction mixtures (diluted or not) were stored at -20^{0}C.

4. Selective amplification: each reaction included two 0.40 ul *EcoR*I selective primers (1 pmol/ul) labeled with different dyes (one was IRDye 700 labeled and the other was IRDye 800 labeled), 1.50 ul unlabeled *Mse*I selective primer (10 ng/ul), 0.20 ul 10 mM dNTPs, 1 ul 10X PCR buffer, 0.60 ul 25 mM MgCl2, 0.12 ul Taq polymerase (5 unit/ul), 3.28 ul ddH2O and 2.50 ul 1:20 diluted pre-amplification PCR product from step 3. The total volume was 10 ul. A master mix was used to ensure precision. PCR was performed using a "touchdown" program: one cycle of 94°C for 10 seconds, 65°C for 30 seconds, and 72°C for 1 minute; twelve cycles of lowering the annealing temperature of 65°C by 0.7°C per cycle while keeping the 94°C for 10 seconds (denaturing) and the 72°C for 1 minute (extending); twenty-three cycles of increasing the extension time of 60 seconds by 1second/cycle while keeping 94°C for 10 seconds, 56°C for 30 seconds; hold at 4°C at completion. Finally 5.0µl of blue stop solution was added to each well, mixed

thoroughly, centrifuged briefly, denatured for 3 minutes at 94°C, and placed on ice immediately.

5. Gel analysis: LI-COR 25-cm plates, KBPLUS (6.5%) gel, 0.25-mM thickness spacers and rectangular 64-tooth combs were used. A 16-bit data collection system was used. The voltage was set to 1500 V, power to 40 W, current to 40 mA, temperature to 45°C, and scan speed to 4. The gel was focused and pre-run for 30 minutes. The wells were flushed completely with a 20 ml syringe to remove urea precipitate or pieces of gel before loading. About 0.5 µl each denatured sample and one lane of molecular size standard (50–700 bp) were loaded using an 8-channel Hamilton syringe. Each gel run took about 3 hours to visualize fragments up to 700 bp. The first bands (about 40 bp) normally appeared about 25 minutes after starting the run.

6. Image collection and analysis: real-time IRDye labeled AFLP band data (TIF images) were automatically collected and recorded during electrophoresis. Image data could be quickly viewed, printed, scored and analyzed.

For scoring at one specific locus, if there was one AFLP band in a sample lane, this band was marked as value "1", if there was no corresponding band for other samples, the value "0" was given. The "1" and "0" data were collected to evaluate genetic variation of shortleaf pine.

1.3.2 Data Analysis

Genetic variation was estimated at the level of species, population and region. Each population was represented by one seed source. The region west of the Mississippi River included 43 samples from seed sources 433, 481, 477, 475 and 423, and the region

east of the River had 50 samples from seed sources 401, 451, 487, 435, 461, 419 and 421(Figure 1.1).

The data of shortleaf pine Z15, loblolly pine 631 and two artificial hybrids were not included in data analysis.

Several different analyses using POPGENE version 1.31 (Yeh and Boyle, 1997) were used to examine genetic variation at all levels. First, AFLP marker diversity was calculated using the following estimates: percentage of polymorphic loci (p), observed number of alleles (na), effective number of alleles (ne) and average heterozygosity or gene diversity (h) (Nei, 1987). Also, the Ewens-Watterson test (Manly, 1985) was used to test polymorphic loci's selective advantage, disadvantage or neutrality and private alleles (Slatkin, 1985) were counted at the level of population and region.

Second, F-statistics were used to examine genetic variation among and within populations and regions. The gene diversity in the total population (Ht) is the sum of average gene diversity between subpopulations (Dst) and average gene diversity within subpopulations (Hs). The formula is Ht = Hs + Dst. The relative amount of gene differentiation among subpopulations was measured by the coefficient of gene differentiation (Gst). Gst = Dst/Ht. Estimated gene flow (Nm) was calculated by the formula Nm = 0.5 (1-Gst)/Gst (Mcdermott and McDonald, 1993).

Third, Nei's analysis of unbiased gene diversity in subdivided populations (Nei, 1987) was used to indicate genetic diversity at the level of populations in shortleaf pine. The Nei's unbiased genetic distance (1978) was used to generate a dendrogram based on the method of Unweighted Pair Group Method with Arithmatic Mean (UPGMA) to

demonstrate relationships among populations. Also, correlation analysis was used to find out the correlation relationship between genetic distances and geographic distances.

1.4 Results

1.4.1 Genetic Diversity

Eighteen primer pairs obtained by screening 48 primer pairs produced 794 loci, of which 523 were polymorphic (Table 1.2) in the 93 shortleaf pine samples.

Table 1.2 Primer Pairs Producing Polymorphic Loci in Shortleaf Pine

Primer Pair	# of Loci	# of Polymorphic Loci	% Polymorphic Loci
M-CCTGxZ-ACG	60	54	90.00
M-CCGAxE-ACG	41	35	85.37
M-CCAGxE-ACG	59	48	81.36
M-CCCGxE-ACA	67	54	80.60
M-CCCGxE-ACG	30	24	80.00
M-CCCGxE-ACG	15	12	80.00
M-CCTCxE-ACG	99	76	76.77
M-CCGAxZ-ACC	30	21	70.00
M-CCGAxE-ACT	30	20	66.67
M-CCCAxE-ACG	47	31	65.96
M-CCTTxE-ACG	49	32	65.31
M-CCTGxE-ACC	33	21	63.64
M-CCTAxE-ACG	63	38	60.32
M-CCGGxZ-ACT	36	21	58.33
M-CCGAxE-ACA	16	7	43.75
M-CCGCxE-ACT	31	11	35.48
M-CCTCxE-ACC	56	12	21.43
M-CCTTxE-ACC	32	6	18.75
Total	794	523	65.87

The first 8 primer pairs produced at least 70% polymorphic loci, so they provided the most information about shortleaf pine variation and they may be useful in studying shortleaf pine hybridization levels with other species. The details of the primer pairs and the markers are listed in the appendix.

Figure 1.2 is part of a typical AFLP gel picture produced by primer pair M-CCTCxE-ACG.

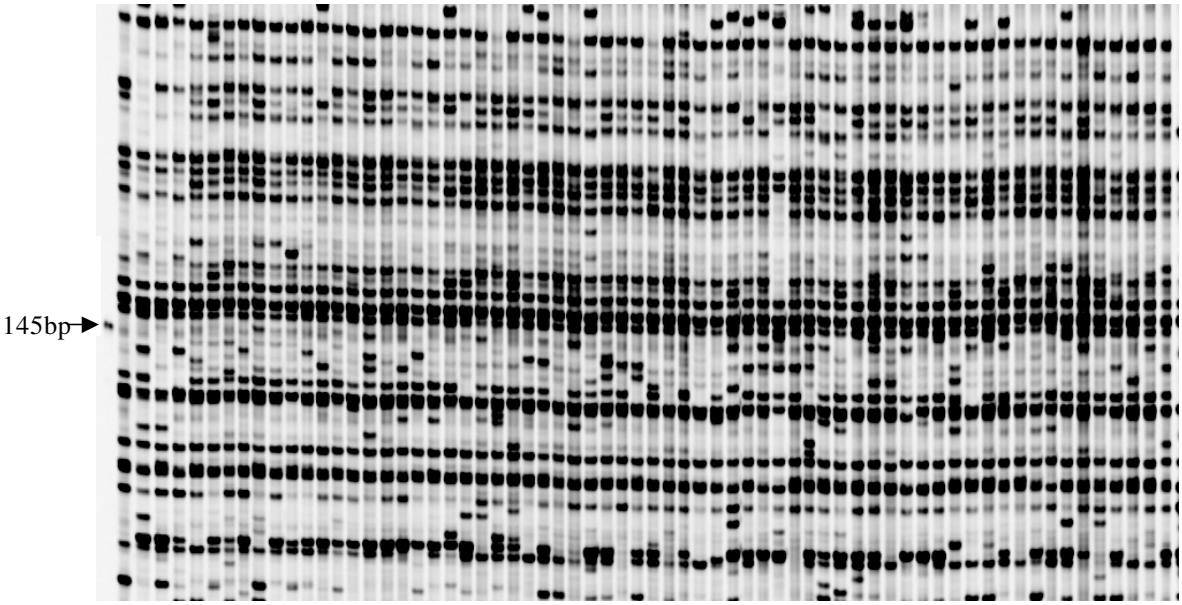

Figure 1.2 A part of the AFLP gel picture produced by primer pair M-CCTCxE-ACG
The 1st lane: a molecular standard, the 2nd lane: shortleaf pine Z15, the 3rd lane: loblolly pine 631, the 4th and 5th lanes: hybrids between Z15 and 631, the rest lanes: shortleaf pine samples from the SSPSSS planting.

The Ewens-Watterson test was used for testing loci neutrality at the level of 11 populations, and showed that 768 of the 794 loci were selectively neutral, 21 loci (loci ID: 92, 113, 141, 151, 180, 184, 276, 331, 538, 551, 619, A22, A27, A37, A39, A42, A45, A53, A58, A60 and A65) were selected against and 5 loci (loci ID: 608, 609, 613, 576 and 632) were favored by selection. The same test was applied to the region west (43 samples) and the region east (50 samples) of Mississippi River. At the regional level, 768 loci were selectively neutral, 19 loci (loci ID: 64, 92, 105, 180, 257, 260, 416, 419, 466, 520, 538, 566, S4, A22, A37, A39, A42, A53 and A58) were selected against, and 7 loci (loci ID: 86, 549, 576, 608, 609, 613 and 632) were selectively favored.

Ten AFLP bands were found only in one population and are called private alleles (Slatkin, 1985). Seven of the 10 private alleles were in populations in the east region and the other three were in populations in the west region (Table1.3). Of note, both 451(401) and 487 have three private alleles. At the regular level, the east region, including 50 samples, and west region, including 43 samples, have approximately similar numbers of private alleles; 12 in east and 10 in west. It is interesting to note that all alleles favored by selection were private alleles and these alleles are distributed fairly evenly in the populations and regions. The private alleles other than those selectively favored may be the results of an artifact of sampling, rare alleles in the species, or from out crossing with other pine species. However, these private alleles were not found in the loblolly pine sampled in this study.

Table 1.3 Private alleles in shortleaf pine populations by population and region

Population ID	Private allele ID
Eastern populations	
451&401	504, 576*, 642
487	144, 549, 613*
435	632*
Western populations	
433	609*
481	86
475	608*
Regions	
East region	252, 337, 470, 484, 504, 545, 549*, 576*, 613*, 620, 632*, 642
West region	86*, 122, 134, 167, 299, 409, 476, 491, 608*, 609*

* Alleles favored by selection

For shortleaf pine, the overall percentage of polymorphic loci was 65.87% (Table 1.4), the observed number of alleles was 1.66, the effective number of alleles was 1.24, and average heterozygosity was 0.15. Within populations, the mean percentage of

polymorphic loci (38.83%) was much lower than that within the species; the observed number of alleles (1.39) was a little lower than that within the species; the effective number of alleles (1.20) and average heterozygosity (0.12) were similar to the estimates within species.

Table 1.4 Summary of genetic diversity of shortleaf pine for all populations and regions based on 794 AFLP loci

Population ID	Percent Polymorphic Loci (P)	Observed # of Alleles (na)	Effective # of Alleles (ne)	Average heterozygosity (h)
East populations				
451 or 401	44.96	1.45	1.22	0.13
487	44.96	1.45	1.23	0.14
435	40.81	1.41	1.21	0.13
419	29.09	1.29	1.17	0.10
461	36.27	1.36	1.19	0.11
421	25.57	1.26	1.16	0.10
Mean	36.94	1.37	1.20	0.12
East Region	59.07	1.59	1.25	0.15
West populations				
433	39.04	1.39	1.20	0.12
477	39.55	1.40	1.21	0.13
481	52.14	1.52	1.28	0.17
475	40.43	1.40	1.20	0.12
423	34.26	1.34	1.18	0.11
Mean	41.08	1.41	1.21	0.13
West Region	63.48	1.63	1.28	0.17
Mean (within Populations)	38.83	1.39	1.20	0.12
Within Species	65.87	1.66	1.24	0.15

The genetic diversity measures in the east region were a little lower than for the west region (Table 1.4). The percentage of polymorphic loci was 59.07% in the east region and 63.48% in the west region; the east region had 1.59 observed alleles and 1.25 effective alleles while west region had 1.63 observed alleles and 1.28 effective alleles;

the average heterozygosity was 0.15 in the eastern region versus 0.17 in the western region.

1.4.2 Genetic Structure

Among populations, the values of G_{st} ranged from 0.0280 at locus L12 to 0.7482 at locus 5. The mean value of G_{st} was 0.1971, which means that 19.71% of the observed genetic diversity existed among the 11 subpopulations while 80.29% of the genetic diversity observed was within populations. The unbiased measures of genetic diversity were high and genetic distances were low for all pairwise comparisons, with the lowest genetic diversity (0.9481), and highest genetic distance (0.0533) between population 477 and 421, and highest genetic diversity (0.9867) and lowest genetic distance (0.0134) between population 487 and 435. The high value of genetic diversity and low value of genetic distance suggests that the genetic structure among subpopulations was very similar. Figure 1.3 is the phenogram got by UPGMA based on Nei's (1978) unbiased genetic distance.

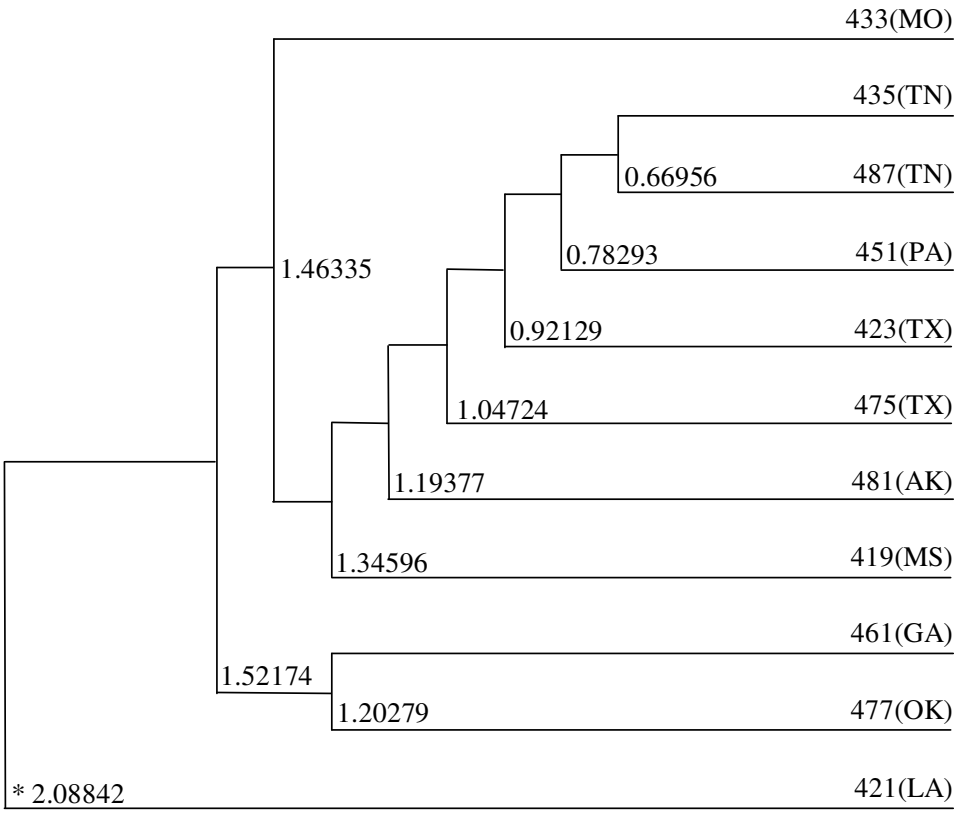

Figure 1.3 Phenogram of shortleaf pine populations based on Nei's (1978) unbiased genetic distance

* The genetic distances among groups

According to Figure 1.3, there appears to be a relationship between the genetic distance and geographic distance in some sub-regions. For example, two populations in TN (435 & 487) and two populations in TX (423 & 475) have relatively low genetic distances. However, across the entire region there is no apparent relationship between genetic distance and geographic distance. For example, the population from Morgan, TN(435) has a shorter genetic distance (0.921) between the Angelina TX(423) population than the distance (1.346) between the Lafayettle MS(419) population, but it is geographically more distant from 423 than 419.

Figure 1.4 shows no correlation relationship between genetic distances and geographic distances (r=0.196)

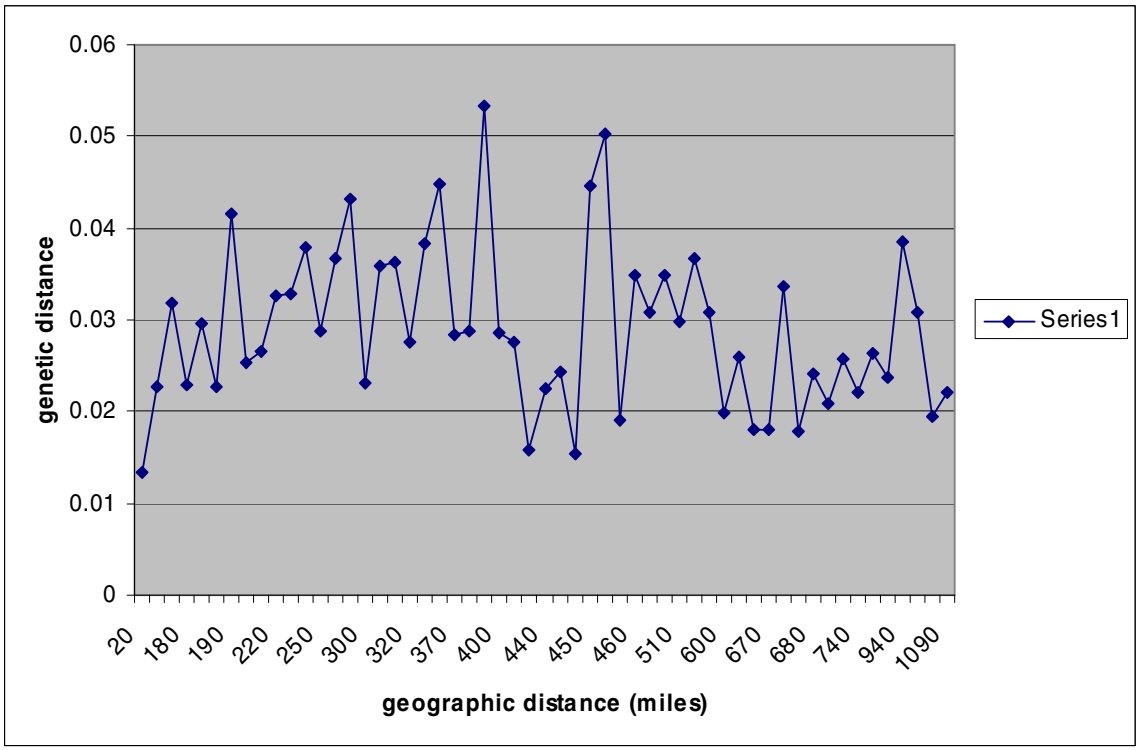

Figure 1.4 Correlations between shortleaf pine populations' genetic distances and geographic distances

Gene flow, Nm, was 2.0372 among populations, which means approximately two alleles migrate per generation. Wright (1931) noted that Nm of one or more would effectively annul any genetic difference between populations. Thus if Nm>1, it is assumed that there is a sufficient level of migration among populations to prevent differentiation. The relatively high rate of migrations (Nm=2.0372) among populations can explain the small genetic difference among populations (19.71%) in this study.

Between the two regions, the genetic diversity estimates (Gst) ranged from 0.000 at locus 48 to 0.267 at locus 160, with a mean of 0.0195. This Gst value suggests that

only 1.95% of the total genetic diversity found was between the two regions, therefore most of the genetic diversity (98.05%) occurs within both regions. The unbiased genetic diversity of the two regions is 0.9945 and the genetic distance is 0.0056. The high gene flow (Nm=25.1122) between the east and west regions has no doubt lead to the high similarity.

1.5 Discussion

Not many trees exist in SSPSSS plantings, so the sample sizes of some seed sources are not big. For example, the seed sources 419 and 421 only have 5 samples. The small sample sizes may lead to a little askew results.

To our knowledge, this study is the first to use AFLPs to explore genetic diversity in shortleaf pine. When compared with previous studies based on isoenzyme markers, our study differs as follows:

First, AFLPs revealed a lower overall percentage of polymorphic loci (65.87%) than Raja et al. (1997) (87.2%) and Edwards and Hamrick (1995) (91%). Sun et al. (1999) found similar differences when they compared the genetic diversity obtained by isozyme, RAPD and microsatellite markes in E*lymus caninus*. RAPD revealed 58% polymorphic loci while isozyme showed 73% polymorphic loci in their study. Though they used RAPDs and we used AFLPs, the nature of RAPDs and AFLPs is similar. Both marker types are dominant and they reflect random diversity of coding and non-coding regions across the whole genome, while isozyme markers reflect diversity of coding regions only. The AFLPs were used in this study because AFLPs have better repeatability than RAPDs.

Second, this study revealed higher (Gst=0.1971) genetic diversity among populations than Raja et al. (1997) (0.089) or Edwards and Hamrick (1995) (0.026). The difference may be caused by the marker loci sampled in the different studies. Raja et al. (1997) and Edwards and Hamrick (1995) used isoenzyme loci, and as most of the isoenzymes reflect essential biological functions in *Pinus,* strong selection on these isoenzyme loci would prevent the accumulation of much variation by mutation (most

mutations being unfavorable) during evolution. Accordingly, genetic variation estimates based on isoenzymes would be low among populations. However, non-coding regions can accumulate change in a neutral manner. In this study, the majority (97%) of the 794 AFLP loci were selectively neutral, as shown by Ewens-Watterson neutrality test. Mutations of selectively neutral loci are not harmful or probably do not change the phenotypes of the individuals, so the neutral mutated loci have no selection pressure. In the long evolution process without selection pressure, one certain locus may accumulate several different kinds of neutral mutations in subpopulations. As a result, these neutral mutations would result in increased genetic variation among subpopulations when using AFLPs. Thus the level of variation at selected loci may differ from that of neutral loci (Nei, 1987).

Third, more markers were used in our AFLP study than in the isoenzyme studies. This study was based on the data of 794 AFLP markers, while only 39 markers were studied by Raja et al. (1997) and 22 by Edwards and Hamrick (1995). The number of markers used in different methods can affect genetic diversity results (Messmer et al., 1991; Smith et al., 1992). Generally, the more markers used, the more precise are the results obtained (Moser and Lee, 1994). Results based on more loci in this study may better represent the genetic diversity across shortleaf pine's genome while limited isozyme loci may only represent genetic diversity in limited coding regions of the genome.

Isoenzyme markers represent the variation of a highly restricted number of enzyme related genes (less than 3% of the genome codes for all proteins in the human genome and less than 30% in *Arabidopsis thanliana* (Arabidopsis Genome Initiative

2000). Thus, only a very small fraction of variation in a species is observed by isozyme studies. AFLPs or RAPDs reflect variation of both coding and non-coding regions, including the nuclear, mitochondrial and chloroplast genome. Therefore, AFLPs (or RAPDs) and isoenzyme markers may reflect genetic diversity of different genome regions. To date researchers have reported low correlations between results based on isozyme markers and RAPDs in various organisms (r=0.204, Sun et al. (1999); r=0.38, Lanner-Herrera et al. (1996); r=0.36, Heun et al. (1994)). Since AFLPs are similar in nature to RAPDs, the correlation between the results from AFLPs and isoenzymes may also be expected to low as we found.

Though AFLPs and isozyme markers may mirror different kinds of genetic diversity, it is interesting to note that our study based on AFLPs and previous studies based on isoenzyme markers draw some similar conclusions in genetic diversity estmates.

As seen in Table 1.4, for this study, genetic diversity measures within populations were lower than within species. Raja et al. (1997), and Edwards and Hamrick (1995) reported similar estimates. The ten private alleles (seven private alleles in the Raja et al. (1997) study and three in Edwards and Hamrick (1995)) may in part result in the lower value of genetic diversity observed within populations than within species.

In this study, all the genetic diversity measures in the western region were slightly higher than those in the eastern region. This same trend was observed by Raja et al. (1997). However, Edwards and Hamrick's (1995) results were different. In their study, all the genetic diversity measures within the eastern region, except expected heterozygosity (He), were slightly higher than those in the western region. Since the differences between east and west regions are small, Edwards and Hamrick (1995)

conclusion that the east and west regions have similar level of genetic diversity seems reasonable.

In summary, all the studies, both those isoenzyme and the AFLP markers revealed that: 1) high genetic diversity existed in shortleaf pine and most of the genetic diversity was within subpopulations; 2) gene flow was high among subpopulations; 3) there was no obvious relationship between population genetic distances and geographic distances; and 4) east and west regions had similar genetic diversity.

Since AFLPs and isoenzyme markers reflect variation of different parts of the genome, it may be best to combine them to get a comprehensive estimate of the genetic diversity for any organism.

Appendix: Primer Pairs and Markers IDs in Shortleaf Pine

(Similar analysis was done for loblolly pine samples from SSPSSS planting in chapter 2 and the same AFLP naming system was used. Marker names beginning with L, S or A were polymorphic among shortleaf pine and loblolly pine; markers with L: high frequency in loblolly pine and low frequency in shortleaf pine; markers with S: high frequency in shortleaf pine and low frequency in loblolly pine; markers with A: similar frequency in shortleaf pine and loblolly pine; other markers: only found in shortleaf pine).

1. **Primer Pair:** M-CCAG X E-ACG

Markers	Band size (bp)
1	55
2	59
3	62
4	60
5	61
6	64
7	63
8	66
9	65
10	67
11	68
12	69
13	81
14	82
15	86
16	85
17	103
18	106
19	130
20	132

Markers	Band size (bp)
21	149
22	155
23	156
24	150
25	159
26	170
27	171
28	161
29	162
30	203
31	210
32	215
33	220
34	222
35	229
36	241
37	242
S1	80
L1	95
A1	100

Markers	Band size (bp)
L2	120
L3	125
S2	145
L4	204
L5	230
L6	270
L7	275
A2	856
A3	99
A4	102
A5	104
A6	105
A7	110
A8	135
A9	140
A10	141
A11	148
A12	160
A13	240

2. **Primer Pair:** M-CCCG X E-ACG

Markers	Band size (bp)
38	50
39	70
40	76
41	77
42	75
43	78
44	100
45	108
46	118
47	95
48	105
49	107
50	122
51	125
52	128

Markers	Band size (bp)
53	130
54	140
55	134
56	135
57	138
67	240
68	241
69	230
70	235
71	254
72	255
73	290
74	299
75	257
76	301

Markers	Band size (bp)
77	305
78	310
79	315
80	345
81	346
A14	120
A15	215
L8	256
S3	270
A16	271
A17	946
A18	124
A19	200
A20	208
A21	255

3. **Primer Pair:** M-CCCG X E-ACA

Markers	Band size (bp)
58	142
59	144
60	150
61	151
62	195
63	152
64	154
65	209
66	220
82	59
83	61
84	62
85	60
86	63
87	65
88	66
89	71
90	77
91	70
92	75
93	76
94	79
95	91

Markers	Band size (bp)
96	78
97	90
98	92
99	99
100	100
101	103
102	110
103	112
104	117
105	104
106	111
107	118
108	125
109	119
110	134
111	137
112	138
113	135
114	136
115	147
116	149
117	148

Markers	Band size (bp)
118	150
119	151
120	160
121	181
122	155
123	180
124	182
125	195
126	201
127	204
128	210
129	215
130	190
131	196
132	205
133	220
134	225
135	230
136	231
137	235
A22	120
A23	133

4. **Primer Pair:** M-CCTG X E-ACG

Markers	Band size (bp)
138	60
139	61
140	62
141	63
142	64
143	65
144	85
145	90
146	95
147	82
148	100
149	105
150	110
151	115
152	125
153	130
154	135
155	140
156	146
157	147

Markers	Band size (bp)
158	160
159	162
160	156
161	157
162	158
163	164
164	166
165	206
166	210
167	211
168	240
169	242
170	245
171	250
172	253
173	260
174	262
175	252
176	263
177	280

Markers	Band size (bp)
178	360
179	290
180	362
181	363
182	370
183	364
184	375
S4	70
A24	80
A25	155
L9	204
L10	320
A27	78
A28	81
A29	101
A30	102
A31	120
A32	145
A33	254

5. **Primer Pair:** M-CCTG X E-ACC

Markers	Band size (bp)
185	55
186	56
187	60
188	66
189	68
190	69
191	65
192	67
193	68
194	99
195	101

Markers	Band size (bp)
196	90
197	95
198	106
199	130
200	135
201	140
202	110
203	120
204	125
205	150
206	146

Markers	Band size (bp)
207	149
208	151
209	153
210	154
211	155
212	152
213	153
S5	105
L11	225
A26	275

6. **Primer Pair:** M-CCGA X E-ACT

Markers	Band size (bp)
214	61
215	62
216	77
217	78
218	79
219	60
220	70
221	76
222	81
223	82

Markers	Band size (bp)
224	80
225	90
226	100
227	105
228	108
229	140
230	142
231	148
232	190
233	203

Markers	Band size (bp)
234	99
235	110
236	195
237	245
238	250
239	205
240	230
L12	165
A35	202

7. **Primer Pair:** M-CCGA X E-ACC

Markers	Band size (bp)
241	54
242	60
243	61
244	75
245	55
246	79
247	81
248	95
249	110
250	85

Markers	Band size (bp)
251	101
252	105
253	121
254	130
255	119
256	120
257	135
258	141
259	140
260	142

Markers	Band size (bp)
261	144
262	148
263	152
264	146
L13	70
L14	100
A36	80
A37	90
A38	125
A39	150

8. **Primer Pair:** M-CCGA X E-ACG

Markers	Band size (bp)
265	60
266	64
267	62
268	66
269	75
270	128
271	118
272	125
273	130
274	140
275	142
276	141
277	150
278	151

Markers	Band size (bp)
279	144
280	155
281	204
282	210
283	215
284	220
285	230
286	232
287	250
288	257
289	270
290	275
291	285
292	295

Markers	Band size (bp)
293	305
L15	76
L16	90
A40	256
A41	300
A42	55
A43	98
A44	100
A45	105
A46	110
A47	120
A48	280
A49	290

9. **Primer Pair:** M-CCGA X E-ACA

Markers	Band size (bp)
294	60
295	66
296	70
297	75
298	65
299	100

Markers	Band size (bp)
300	101
301	110
302	120
303	130
304	149

Markers	Band size (bp)
305	105
306	121
307	150
308	160
309	156

10. **Primer Pair:** M-CCTT X E-ACG

Markers	Band size (bp)
310	55
311	65
312	66
313	67
314	70
315	80
316	85
317	76
318	79
319	90
320	105
321	125
322	126
323	144
324	120
325	121
326	130

Markers	Band size (bp)
327	147
328	148
329	152
330	145
331	146
332	155
333	171
334	166
335	170
336	169
337	172
338	173
339	174
340	202
341	204
342	210

Markers	Band size (bp)
343	175
344	200
345	215
346	225
347	220
348	230
349	251
A50	60
A51	75
L17	78
S7	80
A52	101
A53	142
A54	250
A55	68
A56	150

11. **Primer Pair:** M-CCTT X E-ACC

Markers	Band size (bp)
350	55
351	56
352	57
353	60
354	66
355	71
356	73
357	80
358	65
359	69
360	70

Markers	Band size (bp)
361	89
362	90
363	91
364	105
365	106
366	110
367	120
368	121
369	122
370	130
371	140

Markers	Band size (bp)
372	145
373	146
374	150
375	151
376	152
377	155
378	85
379	115
380	143
S8	90

12. **Primer Pair:** M-CCGC X E-ACT

Markers	Band size (bp)
381	65
382	66
383	67
384	69
385	70
386	72
387	85
388	120
389	125
390	130
391	135

Markers	Band size (bp)
392	140
393	68
394	71
395	80
396	225
397	240
398	149
399	230
400	245
401	246

Markers	Band size (bp)
402	250
403	260
404	266
405	265
406	270
407	271
408	280
409	272
410	275
A57	150

13. **Primer Pair:** M-CCGG X Z-ACT

Markers	Band size (bp)
411	51
412	52
413	53
414	64
415	60
416	62
417	63
418	65
419	66
420	67
421	76
422	80
423	81

Markers	Band size (bp)
424	70
425	75
426	89
427	90
428	91
429	92
430	96
431	100
432	120
433	130
434	140
435	93

Markers	Band size (bp)
436	95
437	146
438	151
439	153
440	155
441	150
442	152
443	154
A58	145
S9	254
A59	55
A60	145

14. **Primer Pair:** M-CCCA X E-ACG

Markers	Band size (bp)
444	66
445	65
446	66
447	68
448	69
449	70
450	72
451	71
452	77
453	75
454	76
455	81
456	82
457	83
458	90
459	110

Markers	Band size (bp)
460	115
461	91
462	105
463	106
464	120
465	116
466	130
467	131
468	132
469	133
470	134
471	150
472	160
473	165
474	170
475	175

Markers	Band size (bp)
476	140
477	141
478	181
479	185
480	190
481	195
482	197
483	180
484	196
485	198
486	231
487	240
L18	230
A61	80
A62	125

15. **Primer Pair:** M-CCTA X E-ACG

Markers	Band size (bp)
488	60
489	64
490	65
491	62
492	63
493	75
494	80
495	85
496	95
497	96
498	97
499	99
500	101
501	110
502	100
503	101
504	103
505	111
506	112
507	113
508	125

Markers	Band size (bp)
509	136
510	140
511	130
512	135
513	143
514	150
515	152
516	153
517	154
518	155
519	156
520	145
521	121
522	160
523	161
524	162
525	163
526	164
527	204
528	210
529	200

Markers	Band size (bp)
530	201
531	215
532	230
533	240
534	220
535	235
536	245
537	250
538	251
539	252
540	253
541	254
542	255
543	256
544	257
545	260
546	265
547	270
A63	90
S10	120
A64	142

16. **Primer Pair:** M-CCTC X E-ACG

Markers	Band size (bp)
548	55
549	56
550	57
551	58
552	60
553	62
554	59
555	61
556	63
557	70
558	71
559	72
560	81
561	73
562	80
563	82
564	91
565	83
566	90
567	92
568	95
569	99
570	100
571	101
572	102
573	105
574	106
575	107
576	108
577	109
578	110
579	112
580	120
581	121

Markers	Band size (bp)
582	122
583	115
584	116
585	123
586	140
587	130
588	131
589	141
590	142
591	143
592	144
593	150
594	151
595	152
596	154
597	146
598	153
599	155
600	165
601	166
602	167
603	168
A65	111
A66	180
L19	345
L20	160
604	170
605	156
606	160
607	171
608	185
609	189
610	190

Markers	Band size(bp)
611	191
612	192
613	193
614	201
615	200
616	205
617	206
618	207
619	208
620	209
621	210
622	212
623	214
624	220
625	216
626	217
627	225
628	240
629	230
630	245
631	260
632	265
633	270
634	280
635	285
636	282
637	283
638	284
639	305
640	310
641	290
642	295
643	300

17. **Primer Pair:** M-CCTC X E-ACC

Markers	Band size (bp)
644	55
645	61
646	62
647	70
648	60
649	63
650	71
651	72
652	73
653	80
654	96
655	97
656	100
657	101
658	102
659	75
660	95
661	103
662	120

Markers	Band size (bp)
663	125
664	126
665	127
666	128
667	129
668	140
669	142
670	146
671	147
672	150
673	155
674	165
675	170
676	171
677	175
678	104
679	121
680	172

Markers	Band size (bp)
681	180
682	185
683	200
684	201
685	205
686	206
687	207
688	230
689	231
690	232
691	240
692	245
693	255
694	260
695	270
696	280
697	198
698	200

References

Arabidopsis Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, Nature. 408: 796–815.

Chen, J. W., Tauer, C. G., Bai, G., Huang, Y., Payton, M. E., and Holley, A. G. 2004. Bidirectional introgression between Pinus taeda and Pinus echinata: Evidence from morphological and molecular data. Can. J. For. Res. 34: 2508-2516.

Delocourt, P. A, Delcourt, H. R., and Davidson, J. L. 1983. Mapping and calibration of modern pollen-vegetation relationships in the southeastern United States. Rev. Palaeobot. PalynoL. 39:1-45.

Dorman, W. K., 1976, The genetics and breeding of southern pine. Agriculture handbook No. 471. U.S. Department of Agriculture Forest Service, Washington, D.C.

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

Edwards, M. A., and Hamrick, J. L. 1995. Genetic variation in shortleaf pine, Pinus echinata Mill. (Pinaceae). For. Genet. 2(1): 21-28.

GemMill, C. E. C., Ranker, T. A., Ragone, D., Perlman, S. P., and Wood, K. R. 1998. Conservation genetics of the endangered endemic Hawaiian genus (Brighamia Campanulaceae). Am. J. Bot. 85(4):528–539.

Heun, M., Murphy, J. P., and Phillips, T. D. 1994. A comparison of RAPD and isozyme analyses for determining the genetic relationships among Avena sterilis L. accessions. Theor. Appl. Genet. 87: 689–696.

Ivey, C. T., and Richards, J. H. 2001. Genetic diversity of everglades sawgrass, Cladium Jamaicense (Cyperaceae). Int. J. Plant Sci. 162:817–825.

Lanner-Herrera, C., Gustafsson, M., Fält, A.-S., and Bryngelsson,T. 1996. Diversity in natural populations of wild Brassica oleracea as estimated by isozyme and RAPD analysis. Genet. Res. Crop Evol. 43: 13–23.

Manly, B. F. J. 1985. The statistics of natural selection. Chapman and Hall. London. New York. P272-282.

McDermott, J. M., and McDonald, B. A. 1993. Gene flow in plant pathosystems. Annu. Rev. Phytopathol. 31:353-373.

Messmer, M. M., Melchinger, A. E., Woodman, W. L., Lee, E. A., and Lamkey, K. R. 1991. Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize population: Comparison of allozyme and RFLP data. Theor. Appl. Genet. 83: 97–107.

Moser, H., and Lee, M. 1994. RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variation in oats. Theor. Appl. Genet. 87: 947–956.

Muluvi, G. M., Sprent, J. I., Soranzo, N., Provan, J., Odee, D., Folkard, G., McNicol, J. W., and Powell, W. 1999. Amplified fragment length polymorphism (AFLP) analysis of genetic variation in Moringa oleifera Lam. Mol. Ecol. 8: 463-470.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 89: 583-590.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University press, New York, p187-192.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. H. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (Pinus echinata). Can. J. For. Res. 27: 740-749.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. 1998. Regeneration methods affect genetic variation and structure in shortleaf pine (Pinus echinata Mill.) For. Genet. 5: 171-178.

Remington, D. L., and O'Malley, D. M. 2000. Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. Genetics. 155: 337-348.

Remington, D. L., Whetten, R. W., Liu, B. H., and O'Malley, D. M. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in P*inus taeda*. Theor. Appl. Genet. 98: 1279-1292.

Schmidtling, R. C., Myszewski, J. H., and McDaniel, C. E. 2005. Geographic variation in shortleaf pine (*Pinus echinata Mill*.) - Cortical monoterpenes. Southern Forest Tree Improvement Conference 28: 161-167. Jun. 20-23, Raleigh, North Carolina.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution. 39: 53-65.

Smith, O. S., Smith, J. S. C., Bowen, S. L., and Tenborg, R. A. 1992. Numbers of RFLP probes necessary to show associations between lines. Maize Genet. Coop. Newslett. 66: 66.

Sun, G. L., Díaz, O., Salomon, B., and Bothmer, R. 1999. Genetic diversity in Elymus caninus as revealed by isozyme, RAPD, and microsatellite markers. Genome. 42: 420–431.

Tarayre, M., and Thompson, J. D. 1997. Population genetic structure of the gynodioecious Thymus vulgaris L. (Laboratoryiatae) in southern France. J. Evol. Biol. 10:157–174.

Tauer, G. C, and McNew, W. R. 1985. Inheritance and correlation of growth of shortleaf pine in two environments. Silvae Genet. 34: 5-11.

Wright, S. 1931. Evolution in Mendelian populations. Genetics. 16: 97-159.

Yeh, F. C., and Boyle, T. J. B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg. J. Bot. 129: 157.

II. GENETIC DIVERSITY AND STRUCTURE IN NATURAL STANDS OF LOBLOLLY PINE (*PINUS TAEDA* L.)

2.1 Abstract

One hundred and twelve loblolly pine trees from 11 seed sources were sampled from Southwide Southern Pine Seed Source Study (SSPSSS) plantings in Mississippi. These samples represent loblolly pine trees from seed produced prior to extensive forest management throughout its geographic range. Eighteen primer pairs obtained by screening 48 primer pairs produced AFLP markers at 647 loci in the samples. The AFLP markers were used to estimate genetic diversity and structure of loblolly pine populations. Throughout the species, loblolly pine was polymorphic at 46.68% (p) of the 647 loci, had 1.47 observed alleles (na) and 1.19 effective alleles (ne) per polymorphic locus. The average heterozygosity (h) was 0.12. Western populations were slightly less diverse than eastern ones. Western populations had lower p, h, na and ne than eastern populations. Genetic structure analysis showed 15.92% of the genetic variation existed among the 11 subpopulations and 84.08% of the genetic variation was within populations. The high values of unbiased measures of genetic identity and low values of genetic distance for all pairwise comparisons indicted that the subpopulations have similar genetic structures. The high inter-population gene flow (Nm=2.64) may explain the high genetic similarity among subpopulations. High gene flow (Nm=22.81) occurred between eastern and western populations. No apparent relationship exists between loblolly pine geographic distance and genetic distance.

2.2 Introduction

Loblolly pine (*Pinus taeda* L.) is perhaps the most important timber species in the United States. Loblolly pine is used for construction lumber, plywood, posts, poles, paper and many other products. Since loblolly pine grows faster than shortleaf pine for at least the first 30 years following planting, more and more native shortleaf pine is being replaced with improved loblolly pine. As a result, more and more improved loblolly pine seedlings are needed for regeneration. A number of programs with the objective of improvement of loblolly pine have been established. For example, the Western Gulf Forest Improvement Program (WGFIP) was founded in 1969, with the objective of providing the Western Gulf Region of the United States with the best genetic quality loblolly pine seed for use in forest regeneration programs. However, how these improvement practices will affect loblolly pine genetic diversity in the long term is unknown.

Genetic diversity provides the initial raw material needed for adaptation and evolution of populations and species (Ledig, 1988; Namkoong, 1991). Tree populations with sustained losses in genetic diversity may become less resistant to biotic or abiotic stress, and have reduced productivity, fitness and health (Bergmann and Scholz, 1987; Bergmann et al., 1990; Raddi et al., 1994). Thus genetic diversity is an essential factor affecting sustainability of forest resources. Moreover, the successes of breeding and genetic improvement programs partly depend on the richness of genetic diversity in desirable traits. However, breeding and genetic improvement practices often reduce genetic diversity (Rajora et al., 2000). To estimate the effect of breeding and genetic improvement on loblolly pine biodiversity in the long term, a base line population

estimate is needed. The loblolly pines sampled in this study were collected from trees in remaining Southwide Southern Pine Seed Source Study (SSPSSS) planting in Mississippi. These trees were raised from seed collected in 1951 and 1952. These seeds were formed at a time when man's influence on forest species diversity due to management was presumed minimal. Thus, the data collected in this study will provide a reference or base level for estimating the effect of the current improvement programs and other activities on loblolly pine genetic diversity.

Prior to the advent of molecular methods, morphological traits such as growth rate (Wells and Wakeley, 1966), wood specific gravity (Byram and Lowe, 1988) and drought resistance (van Buijtenen, 1966) were used to study genetic diversity in loblolly pine. Later, the allozyme electrophoresis technique was used (Roberds and Conkle, 1984). However, the use of morphological characters and allozyme electrophoresis techniques has serious limits. For example, morphological characters of trees are easily affected by environmental factors and the allozyme electrophoresis technique is time-consuming, labor-intensive, expensive, and only a limited number of loci can be studied. Since DNA based markers may distinguish hybrids that can not be discriminated by their morphology and allozyme markers, the use of DNA markers to identify hybrids and study genetic structure has rapidly developed. This study used AFLPs to estimate genetic variation in loblolly pine because this technique requires no previous sequence knowledge, has good repeatability and can detect multiple loci.

The genetic variation of adaptive characters such as growth, disease resistance and survival of loblolly pine populations east of the Mississippi River are reported to be different from that west of the river (Wells and Wakeley, 1970). There were two

hypothesis developed to explain the cause of the east-west differences for loblolly pine. One proposed by Wells et al. (1991) suggests the genetic differentiation is ancient and caused by separation during or preceding the Pleistocene. Florence and Rink (1979) developed the other hypothesis, which states that the pineless landform of the Mississippi River Valley restricted gene flow between loblolly pine in the east and west regions and this has lead to the east-west divergence. This study explored the east-west genetic variation in addition to species diversity at the DNA markers level.

2.3 Materials and Methods

Loblolly pine samples were collected throughout its range as shown in Figure 2.1. The samples of loblolly were from 9 seed sources of a SSPSSS planting in OK and AR, one seed source (OSU) from seed orchard selections made in the 1970's and 1980's with ages around 25 to 40 years old, and one seed source (FL) from a 2005 collection and these trees represent loblolly pine from an allopatric region. For the SSPSSS planting, cones were collected from 20 or more trees in each area and the resulting seeds were mixed to establish the seed resource.

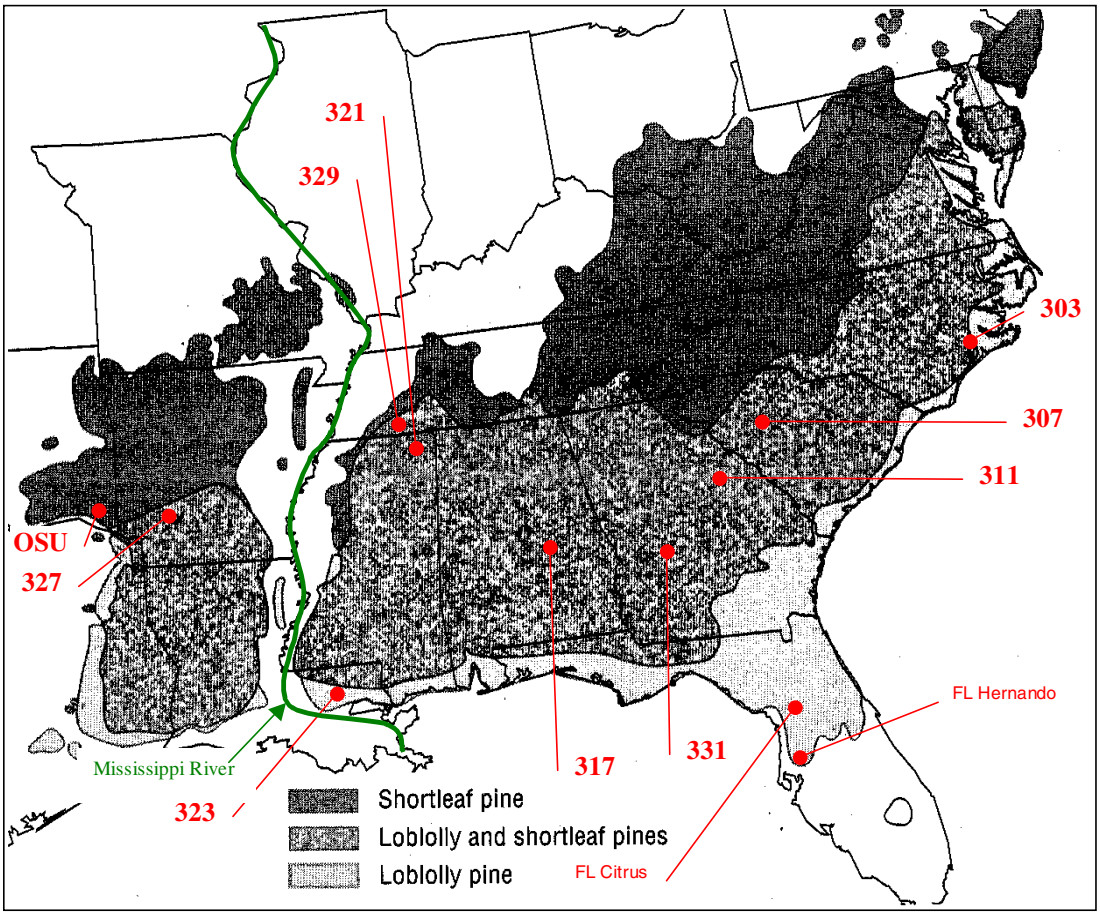

Figure 2.1 The origin of seed sources sampled and the natural range of loblolly pine
The numbers are seed source IDs of samples.

The seed sources sampled for this research were from the origins shown in Figure 2.1 and Table 2.1.

Table 2.1 The origin of the loblolly pine sources sampled in this study

Source ID	State	County	No of tress
303	NC	Onslow	9
307	SC	Newberry	10
311	GA	Clarke	10
317	AL	Clay	11
321	MS	Prentiss	10
323	LA	Livingston	10
327	AR	Clark	11
329	TN	Hardeman	10
331	GA	Spalding	10
OSU*	OK	McCurtain	11
FL[&]	FL	Hernando, Citrus	10

* Not part of the SSPSSS, rather a local collection of equivalent age;
& present day collection from allopatric region

In total, needles from 112 loblolly pine trees were sampled. One hundred and two loblolly pine samples from SSPSSS were collected by Oklahoma State University Forest Resources Center personnel, Idabel, OK, USA. Ten Florida loblolly pine samples were provided by Gregory Powell, University of Florida, Gainesville, FL, USA.

When using the 4300 DNA Analyzer from LI-COR for AFLP analysis, only 64 samples can be loaded in one gel. Consequently, the remaining 29 samples had to be loaded in a second gel. To ensure the same locus was scored for all 93 samples, loblolly pine 631, shortleaf pine Z15, and two hybrids between them were used as standards or check lanes. The shortleaf pine parent Z15, was provided by Bruce Bongarten, Warnell School of Forest Resource, University of Georgia. Z15 is from North Carolina. The loblolly pine parent 631, and the artificial hybrids (F1) between Z15 x 631 were supplied

by Dana Nelson, USDA Forest Service, Southern Institute of Forest Genetics, Saucier, MS, USA. Loblolly pine 631 is from the west central piedmont of Georgia County, GA.

Needles were placed in plastic bags and kept cool with blue ice in a cooler during overnight shipment to the laboratory. Upon arrival, the needles were frozen at -80^{0}C for later use.

2.3.1 AFLP Analysis

A DNeasy Plant Mini kit for isolation of DNA from Qiagen was used to extract DNA from the needle tissue of each loblolly sample.

The primers and the AFLP marker development protocols used by Remington et al (1999) to construct genetic maps and by Remington and O'Malley (2000) to characterize embryonic stage inbreeding depression in loblolly pine were utilized in this study. They used *Eco*RI and *Mse*I as the restriction digestion enzymes. From 48 primer pairs, Remington et al (1999) found a large number of polymorphic fragments using 21 primer combinations of *Eco*RI (E) and *Mse*I (M) primers. The selective nucleic acid sequences for *Eco*RI primers were 5'-ACA-3', 5'-ACC-3', 5'-ACG-3' and 5'-ACT-3'. The selective nucleic acid sequences for *Mse*I primers were 5'-CCAG-3', 5'-CCCG-3', 5'-CCGC-3', 5'-CCGG-3', 5'-CCTG-3', 5'-CCAA-3', 5'-CCAC-3', 5'-CCCA-3', 5'-CCGA-3', 5'-CCTA-3', 5'-CCTC-3' and 5'-CCTT-3'.

The protocols used by Remington et al. (1999) and Remington and O'Malley (2000) were modified as outlined below and used to screen loblolly pine samples for AFLP markers:

1. DNA digestion: each reaction included 5ul DNA (100 ng/ul), 0.25 ul rare cutter restriction endonuclase (RE) *Eco*RI (20 units/ul), 0.5 ul frequent cutter RE *Mse*I (10

units/ul), 5 ul 10X buffer for RE and 29.25 ul ddH_2O. The total volume was 40 ul. A master mix was used to ensure precision. Reactions were incubated for 2 hours at 37^0C. After which, the REs were inactivated at 70^0C for 15 minutes.

2. Ligation of adapter: each reaction included 1 ul *Eco*RI adaptor (5pmol/ul), 2 ul *Mse*I adapter (25 pmol/ul), 1.5 ul 10X ligase buffer, 0.33 ul T_4 DNA ligase (3 unit/ul), 5.17 ul ddH_2O and 40 ul digestion mixture from step 1. The total volume was 50 ul. A master mix was used to ensure precision. Reactions were incubated for 3 hours at 20^0C, or overnight. Then 10 ul of the reaction mixture was loaded to a 1.5% agarose gel to check the digestion-ligation result. Another 10 ul of reaction mixture was transferred into a new 200 ul tube and 90 ul H_2O added, and mixed well. The 1:10 diluted ligated mixture and undiluted portion were stored at -20^0C.

3. Pre-amplification: each reaction included 0.45 ul *Eco*RI preamplification primer (100 ng/ul) and 0.45 ul *Mse*I preamplification primer (100 ng/ul), 0.6 ul 10 mM dNTPs, 3 ul 10X PCR-buffer, 1.8 ul 25 mM M_gCl_2 (for buffer without M_gCl_2), 0.36 ul Taq polymerase (5 unit/ul), 8.34 ul ddH_2O and 15 ul 1:10 diluted ligation mixture from step 2. The total volume was 30 ul. A master mix was used to ensure precision. The PCR program was 28 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute, then hold at 4°C. Following PCR, 10 ul of the PCR product was loaded to a 1.5% agarose gel to check the pre-amplification result. The pre-amplification PCR product was diluted 20 times (10 ul PCR product added to190 ul water). All reaction mixtures (diluted or not) were stored at -20^0C.

4. Selective amplification: each reaction included two 0.4 ul *Eco*RI selective primers (1 pmol/ul) labeled with different dyes (one was IRDye 700 labeled and the other

was IRDye 800 labeled), 1.5ul unlabeled *Mse*I selective primer (10 ng/ul), 0.2ul 10 mM dNTPs, 1ul 10X PCR buffer, 0.6ul 25 mM MgCl2, 0.12ul Taq polymerase (5 unit/ul), 3.28 ul ddH2O and 2.5 ul 1:20 diluted pre-amplification PCR product from step 3. The total volume was 10 ul. A master mix was used to ensure precision. PCR was performed using a "touchdown" program: one cycle of 94°C for 10 seconds, 65°C for 30 seconds, and 72°C for 1 minute; twelve cycles of lowering the annealing temperature of 65°C by 0.7°C per cycle while keeping the 94°C for 10 seconds (denature step) and the 72°C for 1 minute (extension step); twenty-three cycles of increasing the extension time of 60 seconds by 1second/cycle while keeping 94°C for 10 seconds, 56°C for 30 seconds; hold at 4°C at completion. Following PCR 5.0 µl of blue stop solution was added to each well, mixed thoroughly, centrifuged briefly, denatured for 3 minutes at 94°C, and placed on ice immediately.

5. Gel analysis: LI-COR 25-cm plates, KBPLUS (6.5%) gel, 0.25-mM thickness spacers and rectangular 64-tooth combs were used. A 16-bit data collection system was used. The voltage was set to 1500 V, power to 40 W, current to 40 mA, temperature to 45°C, and scan speed to 4. The gel was focused and pre-run for 30 minutes. The /'p/wells were flushed completely with a 20 ml syringe to remove urea precipitate or pieces of gel before loading. About 0.5 µl of each denatured sample and the molecular sizing standard (50–700 bp) were loaded using an 8-channel Hamilton syringe. The run took about 3 hours to visualize fragments up to 700 bp. The first bands (about 40 bp) normally appeared about 25 minutes after starting the run.

6. Image collection and analysis: real-time IRDye laboratoryeled AFLP band data (TIF images) were automatically collected and recorded during electrophoresis. Image data could be quickly viewed, printed, scored and analyzed.

For scoring at one specific locus, if there was one AFLP band in a sample, this band was marked as value "1", if there was no corresponding band in the other samples, the value "0" was given. The "1" and "0" data were collected to evaluate genetic variation of loblolly pine.

2.3.2 Data Analysis

Genetic variation was estimated at the level of species, population and region. Each population was represented by one seed resource of 9 or 10 trees. The region west of the Mississippi River included 22 sample trees from sources OSU and 327, and the region east of the River had 80 sample trees from sources 329, 321, 317, 331, 311, 307, 303 and 323 (Figure 2.1).

The data of shortleaf pine Z15, loblolly pine 631 and two artificial hybrids were not included in data analysis.

Several different analyses using POPGENE version 1.31 (Yeh and Boyle, 1997) were used to examine genetic variation at all levels. First, AFLP marker diversity was calculated using the following estimates: percentage of polymorphic loci (p), observed number of alleles (na), effective number of alleles (ne) and average heterozygosity or gene diversity (h) (Nei, 1987). Also, the Ewens-Watterson test (Manly, 1985) was used to test the polymorphic loci's selective advantage, disadvantage or neutrality, and private alleles were counted at the level of populations and regions.

Second, F-statistics were used to examine genetic variation among and within populations and regions. Gene diversity in the total population (Ht) is the sum of average gene diversity between subpopulations (Dst) and average gene diversity within subpopulations (Hs). The formula is Ht = Hs + Dst. The relative amount of gene differentiation among subpopulations was measured by the coefficient of gene differentiation (Gst). Gst = Dst/Ht. Estimated gene flow (Nm) was calculated by the formula Nm = 0.5 (1-Gst)/Gst (McDermott and McDonald, 1993).

Third, Nei's analysis of unbiased gene diversity in subdivided populations (Nei, 1987) was used to indicate genetic diversity at the level of populations in loblolly pine. The Nei's unbiased genetic distance (1978) was used to generate a dendrogram based on the method of Unweighted Pair Group Method with Arithmatic Mean (UPGMA) to demonstrate relationships among populations. Also, correlation analysis was used to find out the correlation relationship between genetic distances and geographic distances.

2.4 Results

2.4.1 Genetic Diversity

Twenty-one primer pairs obtained by screening 48 primer pairs produced 647 loci, of which 303 were polymorphic (Table 2.2) in the 112 loblolly pine samples.

Table 2.2 Primer Pairs Producing Polymorphic Loci in Loblolly Pine

Primer Pair	# of Loci	# of Polymorphic Loci	% Polymorphic Loci
M-CCTGxE-ACG	63	60	95.24
M-CCAGxE-ACG	55	47	85.45
M-CCGAxE-ACG	26	19	73.08
M-CCTTxE-ACG	45	32	71.11
M-CCCGxE-ACA	17	11	64.70
M-CCCGxE-ACA	31	20	64.52
M-CCTAxE-ACG	27	17	62.96
M-CCGAxE-ACC	15	9	60.00
M-CCTCxZ-ACG	21	12	57.14
M-CCCAxE-ACG	22	12	54.54
M-CCGAxE-ACT	24	11	45.83
M-CCTGxE-ACC	38	17	44.74
M-CCGGxE-ACT	18	6	33.33
M-CCGCxE-ACT	16	4	25.00
M-CCAGxE-ACA	37	9	24.32
M-CCGAxE-ACA	11	2	18.18
M-CCTCxE-ACC	39	6	15.38
M-CCTCxE-ACT	48	5	10.42
M-CCTTxE-ACC	25	2	8.00
M-CCCAxE-ACT	29	1	3.45
M-CCCAxE-ACC	40	1	2.50
Total	647	303	46.68

The first 8 primer pairs produced at least 60% polymorphic loci, so they provide the most information about loblolly pine variation and they would prove most useful in

studying loblolly pine hybridization levels with other species. The details of the primer pairs and the markers are listed in the appendix.

Figure 2 is a typical AFLPs gel picture produced by primer pair M-CCAGxE-ACG.

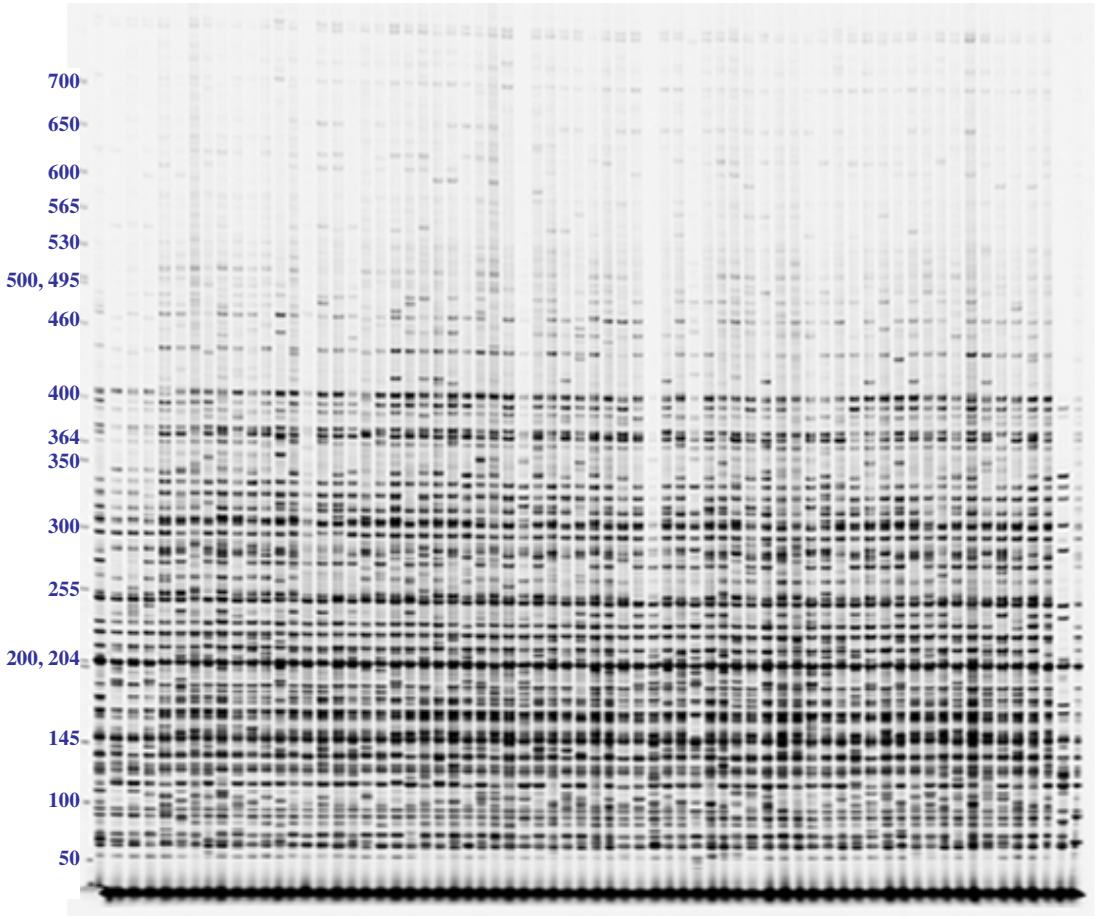

Figure 2.2 A typical AFLP gel picture produced by primer pair M-CCAGxE-ACG

The 1st lane: a molecular standard, the 2nd lane: shortleaf pine Z15, the 3rd lane: loblolly pine 631, the 4th and 5th lanes: hybrids between Z15 and 631, the rest lanes: loblolly pine samples from the SSPSSS planting.

The Ewens-Watterson test for neutrality at the level of the 11 populations showed that 633 of the 647 loci tested were selectively neutral, 10 loci (loci ID: 85, 87, 88, 192, 290, 485, 513, L6, A62 and A66) were selected against and 4 loci (loci ID: 5, 11, 123 and

132) were favored by selection. The same test was applied for the east and west regions. At the regional level, 629 of the 647 loci were selectively neutral, 14 loci (loci ID: 8, 85, 87, 192, 407, 410, 485, 513, 518, L1, L6, A6, A45 and A62) were selected against and 4 loci (loci ID: 5, 11, 123 and 132) were favored by selection.

Six AFLP bands were found in only one population (Table 2.3) and are called private alleles (Slatkin, 1985). Five of the six private alleles were in the eastern populations and the other one was in the western populations. At the regional level, the east had 23 private alleles while the west had only one private allele. It is interesting to note that all selection favored alleles were private alleles. Besides these selectively favored alleles, the other private alleles may be the results of an artifact of sampling, simply rare alleles, or from crosses with other pine species. For example, the locus S5 in the east region was found at high frequency in all shortleaf pine populations sampled in this study. A2 and A23 were evenly distributed in shortleaf pine populations, but they have the frequency of 21.5% and 5.4% in shortleaf pine respectively.

Table 2.3 Private alleles in loblolly pine populations by population and region

Population ID	Private allele ID
East populations	
311	13
317	123*
303	132*
fl	191
323	5*
West populations	
OSU	11*
Regions	
East region	5*, 12, 13, 25, 27, 31, 83, 89, 108, 111, 123*, 132*, 135, 142, 143, 145, 191, 360, 502, 516, S5, A21, A23
West region	11*

* The alleles favored by selection

For loblolly pine sampled from the Mississippi SSPSSS planting, the overall percentage of polymorphic loci was 46.68% (Table 2.4), the observed number of alleles was 1.47, the effective number of alleles was 1.19 and the average heterozygosity was 0.12. The trees (FL) sampled from the allopatric region of recent origin had a lower number of polymorphic loci (29.37%), a lower number of observed alleles (1.31) and effective alleles (1.17), and lower average heterozygosity (0.10) when compared to the trees from the SSPSSS. All samples from the SSPSSS were in the sympatric region with shortleaf pine (Figure 2.1).

Within populations, the mean percentage of polymorphic loci (30.54%) was much lower than that within the species; all other measures including the observed number of alleles (1.31), the effective number of alleles (1.17) and average heterozygosity (0.10) were slightly lower than within species estimates.

Genetic diversity measures in the east region were higher than those for the west region (Table 2.4). The percentage of polymorphic loci was 46.06% in the east region and 35.09% in the west region; the east region had 1.46 observed alleles, 1.21 effective alleles and the west region had 1.35 observed alleles and 1.18 effective alleles; the average heterozygosity was 0.13 in the east region versus 0.11 in the west region.

Table 2.4 Summary of genetic diversity estimates for loblolly pine for all populations and regions based on 647 loci

Population ID	Percent Polymorphic Loci (P)	Observed # of Alleles (na)	Effective # of Alleles (ne)	Average heterozygosity (h)
East populations				
303	31.07	1.31	1.18	0.11
329	31.07	1.31	1.16	0.10
321	28.13	1.28	1.16	0.09
307	34.93	1.35	1.19	0.11
311	26.74	1.27	1.15	0.09
317	35.55	1.36	1.20	0.12
331	29.37	1.29	1.17	0.10
323	31.68	1.32	1.17	0.10
Mean	31.07	1.31	1.17	0.10
East Region	46.06	1.46	1.21	0.13
West populations				
OSU	24.27	1.24	1.13	0.08
327	32.61	1.33	1.18	0.11
Mean	28.44	1.29	1.16	0.10
West Region	35.09	1.35	1.18	0.11
Mean (Within Populations)	30.54	1.31	1.17	0.10
Within Species	46.68	1.47	1.19	0.12
FL	29.37	1.30	1.17	0.10

2.4.2 Genetic Structure

Among populations, the values of G_{st} ranged from 0.0155 at locus 485 to 0.4238 at locus L12. The mean G_{st} was 0.1592, which means that 15.92% of the observed genetic diversity exists among the 11 subpopulations and 84.08% of that genetic diversity observed is within populations. The unbiased measure of genetic identity was high and genetic distance was low for all pairwise comparisons, with the lowest genetic identity

(0.9752) and highest genetic distance (0.0251) between population 311 and 307, and highest genetic identity (0.9908) and lowest genetic distance (0.0093) between population FL and 303. In general, the high value of genetic identity and low value of genetic distance indicts that the genetic structure among populations is very similar. Figure 2.3 is the phenogram resulting from UPGMA based on Nei's (1978) unbiased genetic distance.

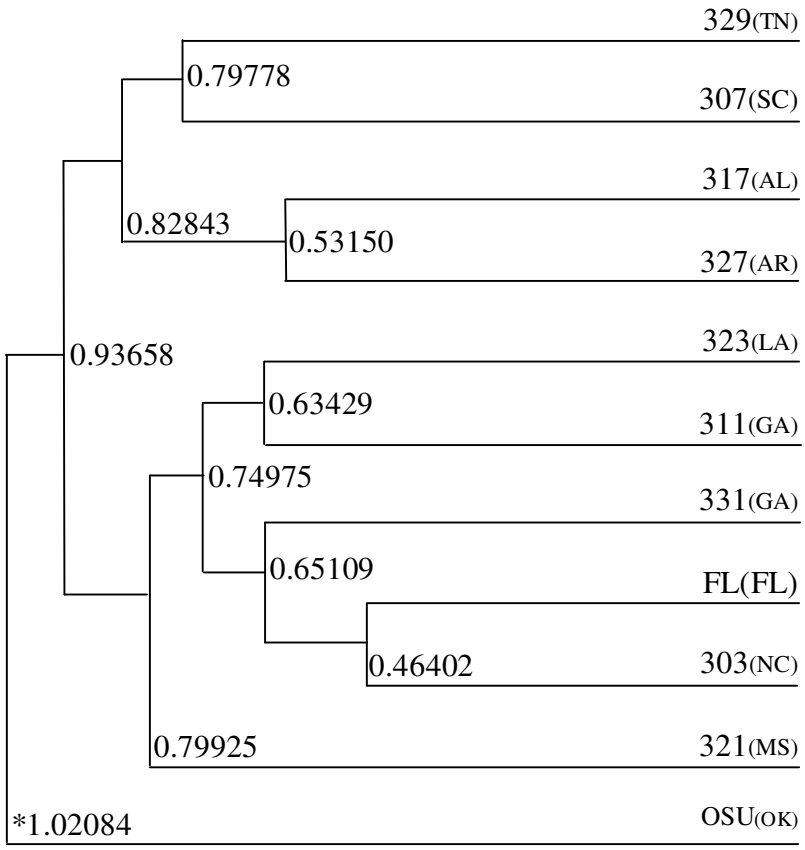

Figure 2.3 Phenogram of loblolly populations using Nei's (1978) unbiased genetic distance

* the genetic distances among groups

In viewing Figure 2.3, it seems there is no apparent relationship between genetic distance and geographic distance. For example, populations 307 and 311 are in close proximity geographically, but the highest genetic distance existed between them.

Likewise, populations FL and 303 have the lowest genetic distance, but they are far away from each other geographically.

Figure 2.4 shows no correlation relationship between genetic distances and geographic distances (r=0.222).

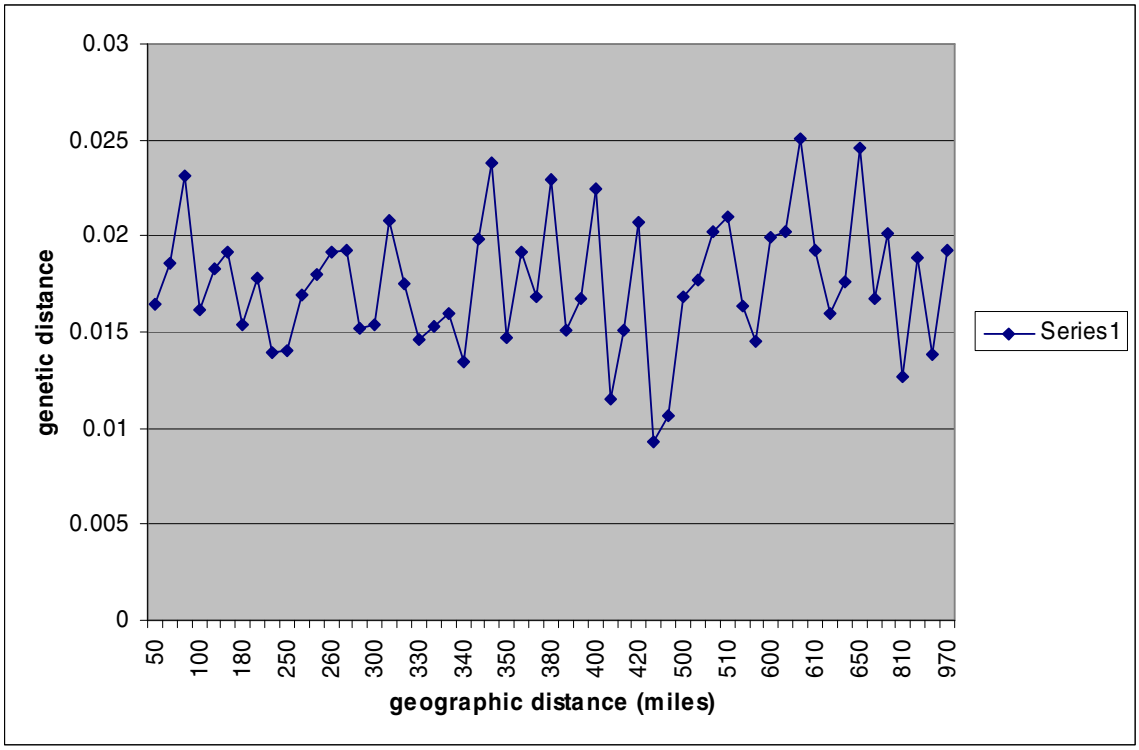

Figure 2.4 Correlations between loblolly pine populations' genetic distances and geographic distances

Gene flow, Nm, was 2.64 among populations, which means approximately three alleles migrate among populations per generation. Wright (1931) noted that a single allele migration every two generations (Nm=0.5) can effectively annul any genetic difference caused by drift. Thus if Nm>1, it is assumed that there is a sufficient level of migration among populations to prevent differentiation. The relatively high rate of

migrations (Nm=2.64) among populations can explain the relatively small genetic differences found among populations (15.92%) in this study.

Between the two regions, the Gst values range from 0.000 at locus 289 to 0.2645 at locus 414, with a mean of 0.0214. This Gst value suggests that only 2.14% of the total genetic diversity found is between the two regions and most of the genetic diversity (97.86%) occurs within regions. The unbiased genetic diversity of the two regions is 0.9954 and the genetic distance is 0.0046. The high gene flow (Nm=22.81) has no doubt led to the high similarity between the two regions.

2.5 Discussion

It is interesting to note that all selection favored alleles were private alleles and the same selection favored alleles were found at both the population level and the region level in this study. These alleles may be involved in very important functions for loblolly pine to survive in certain locations and they were maintained by selection.

To our knowledge, this study is the first to use AFLPs to explore genetic diversity in loblolly pine. Compared with previous studies based on isoenzyme markers and microsatellite markers, our study differs in the following ways:

First, AFLPs revealed lower mean percentage of polymorphic loci within populations (30.54%) than that (64.9%) reported by Schmidtling et al. (1999), whose isoenzyme study also used samples from a SSPSSS planting. Sun et al. (1999) reported similar differences in results when they compared genetic diversity measured using isozyme, RAPD and microsatellite marker in E*lymus caninus*. RAPDs revealed 58% polymorphic loci while isoenzymes found 73% polymorphic loci in their study. Though they used RAPDs and we used AFLPs, the molecular nature of RAPDs and AFLPs is similar. Both are dominant markers and they reflect random diversity of coding and non-coding regions across the whole genome, while isozyme markers reflect diversity only in coding regions.

Second, this study revealed higher (Dst=0.1592) genetic diversity among populations than that of Schmidtling et al. (Dst=1999) (0.066). Schmidtling et al. (1999) used isoenzyme loci and most isoenzymes reflect essential biological functions, so strong selection on these isoenzyme loci prevents the accumulation of much variation by mutation during evolution. Accordingly, genetic variation estimates based on

isoenzymes would be low among populations. However, non-coding regions can accumulate change in a neutral manner. In this study, the majority (97.84%) of the 647 AFLP loci were selectively neutral, as shown by the Ewens-Watterson neutrality test. The mutations of selectively neutral loci are presumablly not harmful and do not change the phenotypes of individuals, so the mutated neutral loci have no selection pressure. In the evolutionary process, in the absence of selection pressure, any locus may accumulate several different kinds of neutral mutations in subpopulations. As a result, these selection neutral mutations would result in increased genetic variation among subpopulations when measured using AFLPs. Thus the variation at selected loci may differ from those of neutral loci (Nei, 1987) and we revealed higher genetic variation at neutral loci than at selected loci studied by Schmidtling et al. (1999).

Third, our study did not find a clear east-west difference in genetic diversity measures. In contrast, Al-Rabab'ah and Williams (2002) reported that there exists clear east-west genetic differentiation based on microsatellite markers in terms of three factors (chord distance, allelic diversity and diagnostic alleles) examined by principal components analysis. In our study, though there is a big difference between east and west in number of private alleles, the differences in percentage of polymorphic loci, observed number of alleles, effective number of alleles, Nei's gene diversity are small. Schmidtling et al. (1999) reported only a subtle east-west difference in allozymic frequencies but a large difference in number of rare alleles (20 in the east region versus 2 in the west) in agreement with our results.

Fourth, more markers were used in this AFLP study than in the isoenzyme and microsatellite studies. This study was based on the data of 647 AFLP markers, while only

18 isoenzyme loci were studied by Schmidtling et al. (1999) and 18 microsatellite loci in the study of Al-Rabab'ah and Williams (2002). The number of markers used in different methods can affect genetic diversity results (Messmer et al. 1991; Smith et al. 1992). Generally, the more markers used, the more precise are the results obtained (Moser and Lee, 1994). Therefore, the results of this study using many loci may better represent the genetic diversity of loblolly pine than the limited isozyme loci or microsatellite loci studies.

Isoenzyme markers represent the variation of a highly restricted number of enzyme related genes (less than 3% of the genome codes for all proteins in the human genome and less than 30% *in Arabidopsis thanliana* (Arabidopsis Genome Initiative 2000). Thus, only a very small fraction of variation in a species is observed by isozyme studies. AFLPs or RAPDs reflect variation of both coding and non-coding regions, including the nuclear, mitochondrial and chloroplast genome. Microsatellite markers are located in non-coding repetitive regions and they reflect variation of the non-coding region only. Therefore, AFLPs (or RAPDs), isoenzyme markers and macrosatellite markers may reflect genetic diversity of different genome regions. Since coding sequences are under higher selection pressure to maintain functions and non-coding regions have low or no selection pressure, the coding and non-coding sequences undergo different evolutionary processes. For example, repetitive sequences change by amplification and transposition more rapidly than single copy sequences (Sun et al., 1999). So far, researchers have found low correlations between results based on isozyme markers and RAPD markers (r=0.204, Sun et al., 1999; r=0.38, Lanner-Herrera et al., 1996; r=0.36, Heun et al., 1994), and between RAPD and microsatlellite markers

(r=0.235, Russell et al., 1997; r=0.267, Sun et al., 1999) in different organisms. Since AFLPs are similar in nature to RAPDs, the correlation between the results based on AFLPs and isoenzyme or AFLPs and microsatellite may also be low.

Although AFLPs, isozyme and microsatellite markers may mirror different types and levels of genetic diversity, it is interesting to note that our study based on AFLPs, and previous studies based on isoenzyme and microsatellite markers draw some similar conclusions. For example, gene flow (Nm) between the east and west regions in all the studies was high enough to minimize any east-west genetic differentiation (Nm ranged from 1.87 to 6.71, from 3.54 to 9.37, and 22.81 in Al-Rababah and Williams's data (2002), Schmidtling et al.'s (1999) data and our study, respectively). The high gene flow between the two regions does not agree with Florence and Rink's (1979) hypothesis that restricted gene flow between the two regions caused the east-west divergence, but do support the hypothesis of Wells et al. (1991) that the genetic differentiation is ancient and caused by separation during or preceding the Pleistocene.

All the studies, whether based on AFLPs, isoenzyme or microsatellite markers, revealed some common results concerning the genetics of loblolly pine. These are: 1) high genetic diversity exists in loblolly pine and most of the genetic diversity is within subpopulations; 2) gene flow is high among subpopulations; 3) there is no obvious relationship between population genetic distances and geographic distances; and 4) genetic differences between the east and the west regions are probably minimal (although the macrosattelite study did not agree on this point).

Since AFLPs, isoenzyme and microsatellite markers reflect variation of different parts of the genome, it may be best to consider results obtained by all the marker types to get the most comprehensive estimate of genetic diversity for any organism.

Appendix: Primer Pairs and Locus IDs in Loblolly Pine

(Similar analysis was done for shortleaf pine sample from SSPSSS planting in chapter 1 and the same AFLP naming system was used. Marker names beginning with L, S or A were polymorphic among shortleaf pine and loblolly pine; markers with L: high frequency in loblolly pine and low frequency in shortleaf pine; markers with S: high frequency in shortleaf pine and low frequency in loblolly pine; markers with A: similar frequency in shortleaf pine and loblolly pine; other markers: only found in loblolly pine)

1. **Primer Pair:** M-CCAG X E-ACG

Markers	Band size (bp)
1	60
2	65
3	66
4	67
5	68
6	69
7	70
8	71
9	75
10	82
11	90
12	94
13	121
14	130
15	134
16	149
17	151
18	152
19	158

Markers	Band size (bp)
20	161
21	153
22	154
23	155
24	165
25	170
26	200
27	206
28	210
29	280
30	215
31	241
32	310
33	305
S1	80
L1	95
A1	100
L2	120

Markers	Band size (bp)
L3	125
S2	145
L4	204
L5	230
L6	270
L7	275
A2	856
A3	99
A4	102
A5	104
A6	105
A7	110
A8	135
A9	140
A10	141
A11	148
A12	160
A13	240

2. **Primer Pair:** M-CCAG X E-ACA

Markers	Band size (bp)
34	55
35	57
36	65
37	66
38	58
39	59
40	80
41	81
42	90
43	91
44	92
45	93
46	105

Markers	Band size (bp)
47	106
48	108
49	109
50	67
51	100
52	115
53	120
54	121
55	122
56	123
57	124
58	110

Markers	Band size (bp)
59	150
60	152
61	153
62	154
63	155
64	170
65	175
66	200
67	204
68	206
69	208
70	151

3. **Primer Pair:** M-CCCG X E-ACG

Markers	Band size (bp)
71	70
72	54
73	55
74	190
75	195
76	200

Markers	Band size (bp)
77	75
A14	120
A15	215
L8	256
S3	270
A16	271

Markers	Band size (bp)
A17	946
A18	124
A19	200
A20	208
A21	255

4. **Primer Pair:** M-CCCG X E-ACA

Markers	Band size (bp)
78	55
79	56
80	58
81	65
82	57
83	60
84	70
85	75
86	106
87	76
88	105

Markers	Band size (bp)
89	107
90	110
91	115
92	116
93	117
94	121
95	140
96	141
97	125
98	132

Markers	Band size (bp)
99	145
100	146
101	155
102	170
103	190
104	210
105	150
106	160
A22	120
A23	133

5. **Primer Pair:** M-CCTG X E-ACG

Markers	Band size (bp)
107	51
108	52
109	53
110	54
111	55
112	56
113	57
114	58
115	82
116	59
117	95
118	100
119	96
120	110
121	111
122	112
123	113
124	115
125	116
126	125
127	130

Markers	Band size (bp)
128	135
129	139
130	140
131	142
132	143
133	144
134	159
135	161
136	180
137	181
138	190
139	191
140	192
141	200
142	206
143	210
144	215
145	220
146	225
147	230
148	235

Markers	Band size (bp)
149	240
150	250
151	252
152	260
153	270
154	310
155	315
156	355
157	375
S4	70
A24	80
A25	155
L9	204
L10	320
A27	78
A28	81
A29	101
A30	102
A31	120
A32	145
A33	254

6. **Primer Pair:** M-CCTG X E-ACC

Markers	Band size (bp)
158	56
159	70
160	71
161	55
162	72
163	75
164	76
165	73
166	77
167	80
168	90
169	99
170	101

Markers	Band size (bp)
171	102
172	91
173	92
174	120
175	125
176	140
177	135
178	150
179	110
180	115
181	160
182	170
183	152

Markers	Band size (bp)
184	153
185	180
186	190
187	195
188	201
189	210
190	215
191	175
192	199
S5	105
L11	225
A26	275

7. **Primer Pair:** M-CCCA X E-ACT

Markers	Band size (bp)
193	60
194	65
195	66
196	67
197	68
198	69
199	70
200	71
201	72
202	73

Markers	Band size (bp)
203	105
204	110
205	115
206	120
207	125
208	130
209	135
210	150
211	160
212	190

Markers	Band size (bp)
213	195
214	201
215	204
216	210
217	215
218	230
219	235
220	340
S6	155

8. **Primer Pair:** M-CCCA X E-ACC

Markers	Band size (bp)
221	55
222	60
223	65
224	70
225	75
226	80
227	81
228	90
229	95
230	96
231	98
232	101
233	102
234	103

Markers	Band size (bp)
235	104
236	110
237	115
238	120
239	125
240	130
241	135
242	140
243	145
244	150
245	155
246	180
247	185

Markers	Band size (bp)
248	190
249	195
250	200
251	204
252	210
253	215
254	220
255	225
256	230
257	290
258	300
259	350
A34	144

9. **Primer Pair:** M-CCCA X E-ACT

Markers	Band size (bp)
260	55
261	56
262	57
263	60
264	75
265	80
266	90
267	77

Markers	Band size (bp)
268	75
269	105
270	110
271	100
272	115
273	125
274	130
275	135

Markers	Band size (bp)
276	140
277	145
278	180
279	120
280	160
281	204
L12	165
A35	202

10. **Primer Pair:** M-CCGA X E-ACC

Markers	Band size (bp)
282	65
283	66
284	95
285	110
286	120

Markers	Band size (bp)
287	145
288	105
289	119
290	151
L13	70

Markers	Band size (bp)
L14	100
A36	80
A37	90
A38	125
A39	150

11. **Primer Pair:** M-CCGA X E-ACG

Markers	Band size (bp)
291	56
292	60
293	57
294	58
295	125
296	130
297	135
298	145
299	75

Markers	Band size (bp)
300	150
301	190
302	204
303	210
304	230
L15	76
L16	90
A40	256
A41	300

Markers	Band size (bp)
A42	55
A43	98
A44	100
A45	105
A46	110
A47	120
A48	280
A49	290

12. **Primer Pair:** M-CCGA X E-ACA

Markers	Band size (bp)
305	60
306	61
307	65
308	70

Markers	Band size (bp)
309	75
310	80
311	101
312	105

Markers	Band size (bp)
313	120
314	65
315	102

13. **Primer Pair:** M-CCTC X E-ACT

Markers	Band size (bp)
316	60
317	65
318	70
319	73
320	75
321	80
322	96
323	92
324	95
325	96
326	97
327	101
328	102
329	104
330	105
331	106

Markers	Band size (bp)
332	107
333	120
334	125
335	130
336	135
337	140
338	145
339	103
340	110
341	160
342	165
343	170
344	175
345	180
346	190
347	195

Markers	Band size (bp)
348	196
349	197
350	206
351	207
352	208
353	209
354	210
355	211
356	220
357	221
358	230
359	235
360	150
361	155
362	240
363	236

14. **Primer Pair:** M-CCTT X E-ACG

Markers	Band size (bp)
364	60
365	65
366	66
367	70
368	71
369	72
370	77
371	80
372	85
373	86
374	99
375	100
376	101

Markers	Band size (bp)
377	102
378	103
379	120
380	131
381	132
382	133
383	134
384	135
385	140
386	150
387	130
388	141
389	154

Markers	Band size (bp)
390	160
391	161
392	162
393	163
394	185
395	190
396	201
397	202
398	205
399	206
400	208
401	155
402	180

15. **Primer Pair:** M-CCTC X E-ACC

Markers	Band size (bp)
403	55
404	66
405	61
406	65
407	69
408	70
409	79
410	81
411	90
412	95
413	99
414	100
415	102
416	103
417	104

Markers	Band size (bp)
418	110
419	115
420	120
421	125
422	130
423	140
424	144
425	146
426	182
427	203
428	210
429	215
430	220
431	230
432	135

Markers	Band size (bp)
433	137
434	232
435	235
436	290
437	248
438	282
A50	60
A51	75
L17	78
S7	80
A52	101
A53	142
A54	250
A55	68
A56	150

16. **Primer Pair:** M-CCTT X E-ACC

Markers	Band size (bp)
439	58
440	60
441	65
442	70
443	75
444	80
445	81
446	85
447	89

Markers	Band size (bp)
448	95
449	100
450	105
451	110
452	120
453	135
454	138
455	140

Markers	Band size (bp)
456	145
457	150
458	152
459	155
460	160
461	55
462	115
S8	90

17. **Primer Pair:** M-CCGC X E-ACT

Markers	Band size (bp)
463	60
464	61
465	62
466	65
467	70

Markers	Band size (bp)
468	75
469	80
470	100
471	130
472	135

Markers	Band size (bp)
473	146
474	147
475	85
476	132
A57	150

18. **Primer Pair:** M-CCGG X E-ACT

Markers	Band size (bp)
477	52
478	53
479	60
480	65
481	70
482	115
483	120

Markers	Band size (bp)
484	130
485	99
486	140
487	155
488	160
489	210

Markers	Band size (bp)
490	215
491	150
A58	145
S9	254
A59	55
A60	145

19. **Primer Pair:** M-CCCA X E-ACG

Markers	Band size (bp)
492	61
493	70
494	60
495	62
496	99
497	120
498	75
499	90

Markers	Band size (bp)
500	124
501	150
502	122
503	151
504	122
505	155
506	160

Markers	Band size (bp)
507	170
508	180
509	240
510	156
L18	230
A61	80
A62	125

20. **Primer Pair:** M-CCTC X E-ACG

Markers	Band size (bp)
511	150
512	155
513	135

Markers	Band size (bp)
514	140
A65	111
A66	180

Markers	Band size (bp)
L19	345
L20	160

21. **Primer Pair:** M-CCTA X E-ACG

Markers	Band size (bp)
515	55
516	56
517	60
518	57
519	61
520	85
521	94
522	95
523	70

Markers	Band size (bp)
524	80
525	110
526	115
527	135
528	140
529	145
530	150
531	99
532	101

Markers	Band size (bp)
533	160
534	152
535	165
536	175
537	170
538	172
A63	90
S10	120
A64	142

References

Al-Rabab'ah, M. A. and Williams, C. G. 2002. Population dynamics of Pinus taeda L. based on nuclear microsatellites. For. Ecol. Manage. 163: 263–271.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 408: 796–815.

Bergmann, F., and Scholz, F. 1987. The impact of air pollution on the genetic structure of Norway spruce. Silvae genet. 36: 80-83.

Bergmann, F., Gregorius, H-R., and Larsen, J. B. 1990. Levels of genetic variation in European silver fir (Abies alba) — are they related to the species' decline? Genetica. 82: 1-10.

Byram, T. D., and Lowe, W. J. 1988. Specific gravity variation in a loblolly pine seed source study in the Western Gulf Region. Forest. Sci. 34:798-803.

Edwards, M. A., and Hamrick, J. L. 1995. Genetic variation in shortleaf pine, Pinus echinata Mill. (Pinaceae). For. Genet. 2(1): 21-28.

Florence, Z., and Rink, G. 1979. Geographic patterns of allozymic variation in loblolly pine. In: Proceeding of the 15th Southern Forest Tree Improvement Conference, Starkville, MS, 19-21 June 1979. Mississippi State University, MS, pp. 33-41.

Heun, M., Murphy, J. P., and Phillips, T. D. 1994. A comparison of RAPD and isozyme analyses for determining the genetic relationships among Avena sterilis L. accessions. Theor. Appl. Genet. 87: 689–696.

Lanner-Herrera, C., Gustafsson, M., Fält, A.-S., and Bryngelsson, T. 1996. Diversity in natural populations of wild Brassica oleracea as estimated by isozyme and RAPD analysis. Genet. Res. Crop Evol. 43: 13–23.

Ledig, F. T. 1988. The conservation of diversity in forest trees. Bioscience. 38: 471-479.

Manly, B. F. J. 1985, The statistics of natural selection. Chapman and Hall. London. New York. Pp. 272-282.

McDermott, J. M., and McDonald, B. A. 1993. Gene flow in plant pathosystems. Annu. Rev. Phytopathol. 31:353-373.

Messmer, M. M., Melchinger, A. E., Woodman, W. L., Lee, E. A., and Lamkey, K. R. 1991. Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize population: Comparison of allozyme and RFLP data. Theor. Appl. Genet. 83: 97–107.

Moser, H., and Lee, M. 1994. RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variation in oats. Theor. Appl. Genet. 87: 947–956.

Namkoong, G., 1991. Biodiversity — issues in genetics, forestry and ethics. Forestry Chronicle. 68: 438 443.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, pp.187-192.

Raddi, S., Stefanini, F. M., Camussi, A., and Giannini, R. 1994. Forest decline index and genetic variability in Picea abies (L) Karst . For. Genet. 1: 33-40.

Rajora, O. P., Rahman, M. H., Buchert, G. P., and Dancik, B. P. 2000. Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (P*inus strobus*) in Ontario, Can. Mol. Eco. 9: 339-348.

Remington, D. L., and O'Malley, D. M. 2000. Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. Genetics. 155: 337-348.

Remington, D. L., Whetten, R. W., Liu, B. H., and O'Malley, D. M. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in Pinus taeda. Theor. Appl. Genet. 98: 1279-1292.

Roberds, J. H., and Conkle, M. T. 1984. Genetic structure in loblolly pine stands: allozyme variation in parents and progeny. Forest. Sci. 30: 319-329.

Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W., and Waugh, R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95: 714—722.

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

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution. 39: 53-65.

Smith, O. S., Smith, J. S. C., Bowen, S. L., and Tenborg, R. A. 1992. Numbers of RFLP probes necessary to show associations between lines. Maize Genet. Coop. Newslett. 66: 66.

Sun, G. L., Díaz, O., Salomon, B., and Bothmer, R. 1999. Genetic diversity in Elymus caninus as revealed by isozyme, RAPD, and microsatellite markers. Genome. 42: 420–431.

van Buijtenen, J. P. 1966. Testing loblolly pines for drought resistance. Texas Forest Service Technical Report 13. pp.15.

Wells, O. O., and Wakeley, P. C. 1970. Variation in longleaf pine from several geographic sources. Forest. Sci. 16: 28-45.

Wells, O. O., and Wakeley, P. C. 1966. Geographic variation in survival, growth and fusiform-rust infection of planted loblolly pine. For. Sci. Monog. 11: 40.

Wells, O. O., Switzer, G. L., and Schmidtling, R. C. 1991. Geographic variation in Mississippi loblolly pine and sweetgum. Silvae genet. 40: 105-119.

Wright, S. 1931. Evolution in Mendelian populations. Genetics. 16: 97-159.

Yeh, F. C., and Boyle, T. J. B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg. J. Bot. 129: 157.

III. HYBRIDIZATION BETWEEN NATURAL POPULATIONS OF SHORTLEAF PINE (*PINUS ECHINATA* MILL.) AND LOBLOLLY PINE (*PINUS TAEDA* L.)

3.1 Abstract

Two hundred and five shortleaf pine and loblolly pine samples from 22 seed sources were sampled from Southwide Southern Pine Seed Source Study (SSPSSS) plantings or equivalent origin. These samples represent shortleaf and loblolly pine formed prior to intensive forest management throughout their geographic range. Ninety-six AFLPs were produced by 17 primer pairs after screening 48 primer pairs in these samples. Two hybrids in the loblolly pine samples and two hybrids in the shortleaf pine samples were found using the *IDH* (Isocitrate dehydrogenase) marker. Two more hybrids in the shortleaf pine samples were found combing the 96 AFLPs with *IDH* markers using software NewHybrids version 1.1 beta. This study suggested that later generation hybrids can be found using molecular markers and confirmed that *IDH* is a powerful marker to detect hybrids between the two species. To more efficiently detect hybrids codominant markers are needed because codominant markers can provide more genetic information than dominant markers. Hybridization frequency varied geographically, ranging from 25% in MO to 0% in other places in this study. Also, the hybridization level was higher in populations west of Mississippi River than east of the river (9.3% west vs. 0% east in shortleaf pine populations, 4.5% west vs. 1.1% east in loblolly populations and 7.7% west vs. 0.71% east in all populations). The results suggest that the potential for the existence of hybrids or creation of hybrids should be considered in forest management decisions.

3.2 Introduction

Loblolly pine (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) are both of considerable economic importance in the southeast United States. Both species can be used for construction lumber, plywood posts, poles, paper, and other products. They have broad geographic ranges and a large sympatric region (Figure 3.1).

Research has shown that shortleaf pine and loblolly pine probably have the most similar karyotypes among the southern pine (Saylor, 1972), so they are expected to cross easily with each other. As early as 1933, artificial hybrids between them were created by the institute of Forest Genetics Research in California and reported by Schreiner (1937). In nature, however, there are other conditions such as flowering time which affects possible hybridization. Loblolly pine has mature male and female strobili from the end of February to the middle of March, and shortleaf pine has mature strobili about 2-3 weeks later. The general lack of overlap of strobili development generally results in no or low levels of hybridization between the two species. But strobili and maturity time may vary by as much as 3 weeks among trees in the same stand. Strobili maturity is also affected by seasonal climatic fluctuations, which may lead to overlapping times. Thus, hybridization between the two species may occur in sympatric populations in some years (Dorman and Barber, 1956).

As early as 1953, researchers reported trees with morphologies intermediate between loblolly pine and shortleaf pine, suggesting hybrids do occur naturally in the sympatric region (Hare and Switzer, 1969; Zobel, 1953). Some trees in loblolly pine populations in sympatric regions were found to have resistance to fusiform rust (Henry and Bercaw, 1956), to which loblolly pine is generally susceptible but shortleaf pine is

resistant. Likewise, some trees in shortleaf stands have resistance to littleleaf disease, to which shortleaf pine is susceptible and loblolly pine is resistant. Recent studies have revealed a relatively high level of hybridization among trees in a shortleaf pine and loblolly pine population in west-central Arkansas (Raja et al., 1998; Chen et al., 2004) and somewhat lower levels in Georgia (Edwards et al., 1997).

Prior to the advent of molecular tools, morphological characters were used to study hybrids between loblolly pine and shortleaf pine (Mergen et al., 1965; Cotton et al., 1975). Later, isoenzymes, in particular the *IDH* (Isocitrate dehydrogenase) isoenzyme, were used to identify hybrids (Huneycutt and Askew, 1989; Edwards and Hamrick, 1997; Chen et al., 2004). But these morphological and isoenzyme markers are of limited in utility. For example, morphological characters of trees are easily affected by environmental factors. Also, it proved difficult to choose the suitable set of morphological traits to efficiently distinguish hybrids (Hicks, 1973). The isoenzyme *IDH* is a good marker to find first generation hybrids but it can only detect some of the later generation hybrids. More markers are needed to reliablly detect later hybrid generations. Since DNA based markers may distinguish species that can not be discriminated by their morphology, phenology or isoenzyme markers, many such DNA markers have been developed and used to identify hybrids of shortleaf pine and loblolly pine (Chen et al., 2004; Edwards and Hamrick., 1997). This study explored the use of AFLP markers combined with the *IDH* marker to find hybrids between shortleaf pine and loblolly pine. Marker data were analyzed using the software NewHybrids version 1.1 beta (Anderson and Thompson, 2002; Anderson, 2003).

Edwards and Hamrick (1995) found a higher level (4.6%) of hybridization between shortleaf pine and loblolly pine in shortleaf pine populations located west of the Mississippi River than (1.1%) east of the river. This study used AFLPs and the *IDH* marker to further examine possible differences between western and eastern populations. Previous research (Chen et al., 2004; Raja et al., 1997; Edwards and Hamrick, 1997) found relatively high hybridization levels between these two species in some regions, but the hybridization level throughout most of their ranges is unknown. This study sampled shortleaf pine and loblolly pine from allopatric and sympatric populations to study the hybridization level throughout their natural ranges.

Since loblolly pine grows faster than shortleaf pine at least for the first 30 years, more and more shortleaf pine has been replaced with improved loblolly pine. The US Forest Service is one of only a few organizations which regenerates shortleaf pine, usually relying on natural regeneration. As a result, the shortleaf pine stands naturally regenerated by the Forest Service are becoming surrounded by more and more loblolly pine. Thus, it is reasonable to ask if the hybridization level is increasing in naturally regenerated shortleaf pine in areas surrounded by expanding loblolly pine plantings. The hybridization level may play a very important role in shortleaf pine or loblolly pine genetic integrity in the future. If we can estimate how intensive forest management affects hybridization levels, we can deduce how intensive loblolly pine management may affect both shortleaf pine and loblolly pine genetic integrity in a long term. Thus, the samples collected for this study were from Southwide Southern Pine Seed Source Study (SSPSSS) plantings. These plantings contain trees grown from seed collected in 1951 and 1952, when man's influence due to management was minimal. The hybridization

level of these samples will provide a reference or base level to evaluate the effects of currently intensive forest management. This information can serve to develop guidance for shortleaf pine and loblolly pine genetic conservation.

3.3 Materials and Methods

Loblolly pine and shortleaf pine have extensive geographic ranges and a large sympatric region. Samples from the allopatric and sympatric populations were collected as shown in Figure 1:

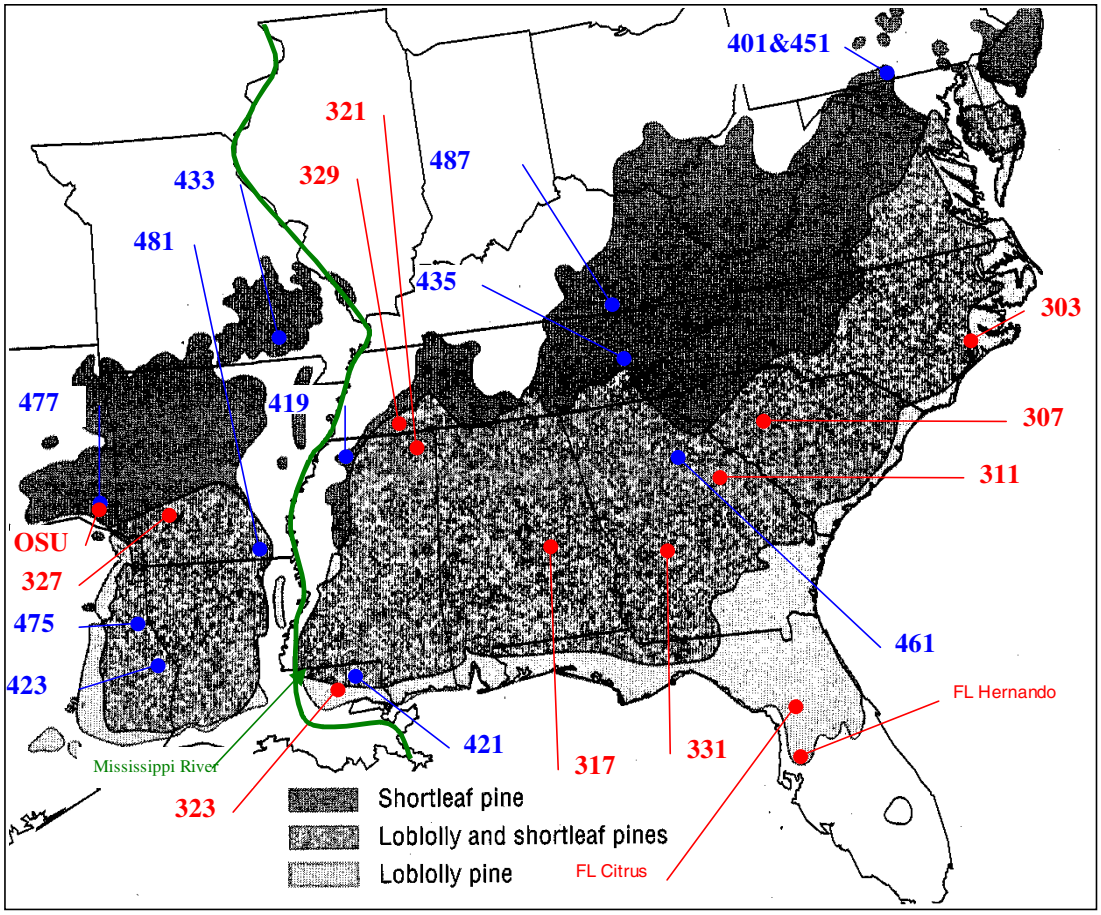

Figure 3.1 The origin of the shortleaf pine and loblolly pine samples, and the species natural ranges.

300's are loblolly pine and 400's are shortleaf pine

Needles and cones of shortleaf pine and loblolly pine were collected from 22 seed sources each (Figure 3.1). The seed sources were created by collecting cones from 20 or more trees at each origin and the resulting seeds were mixed. Trees grown from these

seeds were grown and planted into the SSPSSS planting, which we subsequently sampled. The locations of the seed sources sampled in this research are given in Table 3.1 and Table 3.2.

Table 3.1 The origin of the shortleaf pine sources sampled in this study

Source ID	State	County	No of tress
401*	PA	Franklin	4
419	MS	Lafayette	5
421	LA	St. Helena	5
423	TX	Angelina	7
433	MO	Dent	8
435	TN	Morgan	9
451*	PA	Franklin	10
461	GA	Clarke	8
475	TX	Cherokee	10
477	OK	Pushmataha & McCurtain	8
481	Ark	Ashley	10
487	TN	Anderson	9

(*401 belongs to the original collection made in 1951 and 451 to the collection made in 1955)

Table 3.2 The origin of the loblolly pine sources sampled in this study

Source ID	State	County	No of tress
303	NC	Onslow	9
307	SC	Newberry	10
311	GA	Clarke	10
317	AL	Clay	11
321	MS	Prentiss	10
323	LA	Livingston	10
327	AR	Clark	11
329	TN	Hardeman	10
331	GA	Spalding	10
OSU*	OK	McCurtain	11
FL[&]	FL	Hernando, Citrus	10

*Not part of the SSPSSS, rather a local collection of equivalent age;

&present day collection from allopatric region

The 93 shortleaf pine and the 102 loblolly pine samples (except the Florida collection) were collected by Oklahoma State University Forest Resources Center personnel, Idabel, OK, USA. Ten loblolly pine samples from Florida were provided by Gregory Powell, University of Florida, Gainesville, FL, USA.

When using the 4300 DNA Analyzer from LI-COR for AFLP analysis, only 64 samples can be loaded in one gel. Consequently, the remaining 141 samples had to be loaded in a second gel. To ensure the same locus was scored for all 205 samples, loblolly pine 631, shortleaf pine Z15, and two hybrids between them were used as standards or check lanes. The shortleaf pine parent Z15, was provided by Bruce Bongarten, Warnell School of Forest Resources, University of Georgia. Z15 is from North Carolina. Loblolly pine parent 631, and the artifical hybrids (F1) between Z15 x 631 were supplied by Dana Nelson, USDA Forest Service, Southern Institute of Forest Genetics, Saucier, MS, USA. Loblolly pine 631 is from the west central piedmont of Georgia County, GA.

Collected needles and cones were placed in plastic bags and kept cool with blue ice in coolers during overnight shipment. Upon arrival in the laboratory, the needles were frozen at -80^0C for later use. Cones were placed on laboratory benches to air dry. When the cones opened, the seeds were collected. The seeds were stored frozen at -20^0C for later use.

3.3.1 AFLPs Analysis

Total DNA was extracted from needles of shortleaf pine using a modified CTAB protocol (Doyle and Doyle, 1988) used by our laboratory. A DNeasy Plant Mini kit for isolation of DNA from Qiagen was used to extract DNA from the needle tissue of each loblolly pine sample.

The primers and the AFLP marker development protocols used by Remington et al (1999) to construct genetic maps and by Remington and O'Malley (2000) to characterize embryonic stage inbreeding depression in loblolly pine were utilized in this study. They used *Eco*RI and *Mse*I as the restriction digestion enzymes. From 48 primer pairs, Remington et al. (1999) found a large number of polymorphic fragments using 21 primer combinations of *Eco*RI (E) and *Mse*I (M) primers. The selective nucleic acid sequences for *Eco*RI primers were 5'-ACA-3', 5'-ACC-3', 5'-ACG-3' and 5'-ACT-3'. The selective nucleic acid sequences for *Mse*I primers were 5'-CCAG-3', 5'-CCCG-3', 5'-CCGC-3', 5'-CCGG-3', 5'-CCTG-3', 5'-CCAA-3', 5'-CCAC-3', 5'-CCCA-3', 5'-CCGA-3', 5'-CCTA-3', 5'-CCTC-3' and 5'-CCTT-3'.

The protocols used by Remington et al. (1999), and Remington and O'Malley (2000) were modified as outlined below and used to screen shortleaf pine and loblolly pine samples for AFLP markers:

1. DNA digestion: each reaction included 5 ul DNA (100 ng/ul), 0.25 ul rare cutter restriction endonuclase (RE) *Eco*RI (20 units/ul), 0.5 ul frequent cutter RE *Mse*I (10 units/ul), 5 ul 10X buffer for RE and 29.25 ul ddH_2O. The total volume was 40 ul. A master mix was used to ensure precision. Reactions were incubated for 2 hours at 37^0C. After which the REs were inactivated at 70^0C for 15 minutes.

2. Ligation of adapter: each reaction included 1 ul *Eco*RI adaptor (5 pmol/ul), 2 ul *Mse*I adapter (25 pmol/ul), 1.5 ul 10X ligase buffer, 0.33 ul T_4 DNA ligase (3 unit/ul), 5.17 ul ddH_2O and 40 ul digestion mixture from step 1. The total volume was 50 ul. A master mix was used to ensure precision. Reactions were incubated for 3 hours at 20^0C, or overnight. An additional 10 ul of the reaction mixture was loaded to a 1.5% agarose

gel to check the digestion-ligation result. The 10 ul of reaction mixture was transferred into a new 200 ul tube and 90 ul H_2O added, and mixed well. The 1:10 diluted ligated mixture and undiluted portion were stored at -20^0C.

3. Pre-amplification: each reaction included 0.45 ul *Eco*RI preamplification primer (100ng/ul) and 0.45 ul *Mse*I preamplification primer (100ng/ul), 0.6 ul 10 mM dNTPs, 3 ul 10X PCR-buffer, 1.8 ul 25 mM M_gCl_2 (for buffer without M_gCl_2), 0.36 ul Taq polymerase (5unit/ul), 8.34 ul ddH_2O and 15 ul 1:10 diluted ligation mixture from step 2. The total volume was 30 ul. A master mix was used to ensure precision. The PCR program was 28 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute, then hold at 4°C. Following PCR 10 ul of the PCR product was loaded to a 1.5% agarose gel to check the pre-amplification result. The pre-amplification PCR product was diluted 20 times (10 ul PCR product plus 190 ul water). All reaction mixtures (diluted or not) were stored at -20^0C.

4. Selective amplification: each reaction included two 0.4 ul *EcoR*I selective primers (1 pmol/ul) labeled with different dyes (one was IRDye 700 labeled and the other was IRDye 800 labeled), 1.5 ul unlabeled *Mse*I selective primer (10 ng/ul), 0.2 ul 10 mM dNTPs, 1 ul 10X PCR buffer, 0.6 ul 25 mM $MgCl_2$, 0.12 ul Taq polymerase (5 unit/ul), 3.28 ul ddH_2O and 2.5 ul 1:20 diluted pre-amplification PCR product from step 3. The total volume was 10 ul. A master mix was used to ensure precision. PCR was performed using a "touchdown" program: one cycle of 94°C for 10 seconds, 65°C for 30 seconds, and 72°C for 1 minute; twelve cycles of lowering the annealing temperature of 65°C by 0.7°C per cycle while keeping the 94°C for 10 seconds (denature step) and the 72°C for 1 minute (extension step); twenty-three cycles of increasing the extension time of 60

seconds by 1second/cycle while keeping 94°C for 10 seconds, 56°C for 30 seconds; hold at 4°C at completion. Following PCR 5.0µl of blue stop solution was added to each well, mixed thoroughly, centrifuged briefly, denatured for 3 minutes at 94°C, and placed on ice imMediately.

5. Gel analysis: LI-COR 25-cm plates, KBPLUS (6.5%) gel, 0.25-mM thickness spacers and rectangular 64-tooth combs were used. A 16-bit data collection system was used. The voltage was set to 1500 V, power to 40 W, current to 40 mA, temperature to 45°C, and scan speed to 4. The gel was focused and pre-run for 30 minutes. The wells were flushed completely with a 20 ml syringe to remove urea precipitate or pieces of gel before loading. About 0.5µl each denatured sample and the molecular size standard (50–700 bp) were loaded using an 8-channel Hamilton syringe. The run took about 3 hours to visulize fragments up to 700 bp. The first bands around 40bps normally appeared about 25 minutes after starting the run.

6. Image collection and analysis: real-time IRDye laboratoryeled AFLP band data (TIF images) were automatically collected and recorded during electrophoresis. Image data could be quickly viewed, printed, scored and analyzed.

For scoring at one specific locus, if there was one AFLP band in a sample, this band was marked "+", if there was no corresponding band in another sample, the value "-" was given. The "+" and "-" data were collected for analysis.

3.3.2 *IDH* Analysis

IDH is a co-dominant marker. For conifers, the haploid (n) megagametophyte tissue of the germinating seed is from the mother trees and is preferred for *IDH* analysis. Needles may also be used, but this is more difficult. In this study, for trees for which

seeds were availaboratoryle, ten germinated seeds of each tree were used to obtain megagametophyte tissue. The maternal genotype can be effectively estimated by this sample size (Yeh and Layton, 1979). Huneycutt and Askew (1989) demonstrated that it was possible to pool all 10 megagametophytes and to use this pooled material as a single tissue sample to genotype the tree. If the pooled material results in a single band, the tree is homozygous for either parent-band; if the pooled material results in a double band, this tree is a hybrid. In this study, 10 seeds of each of 110 samples were obtained and tested with the *IDH* marker. For the remaining 95 samples needles were used for the *IDH* analysis and about 0.05 g needle tissue each tree was used.

The protocol used in our laboratory for the *IDH* analysis followed that of Raja et al. (1997) slightly modified as follows:

1. Sample preparation: when haploid megagametophyte tissue was used, seeds in the cones of the sample trees were extracted, dried to approximately 6% moisture content, and frozen at -20°C for later use. Before use, the seeds were thawed to room temperature for 1 h and then immersed in water overnight prior to stratification. Water was drained. Immature and empty seeds were thrown away and the good seeds were stratified moist at 4°C for 60 days to break dormancy before they were germinated on moist filter paper in Petri dishes at room temperature. When the radicle of a seed was about 2 to 5 mM long, the seed was placed in a second Petri dish on moist filter paper at 4^0C until 10 seeds of one tree were obtained. Megagametophytes from 10 seed of each tree were isolated, maintained on ice, and ground in 0.14 M Wendel and Parks (1982) extraction buffer. The extraction buffer included 0.04 M Na-phosphate, 0.20 M sucrose, 0.001 M EDTA, 0.003 M DTT, 0.00 3M ascorbic acid, 0.003 M sodium bisulfite, 0.006

M dietlyldithiocarbamate, 5% PVP-40 and 0.1% β-mercaptoethanol (pH is adjusted to 7.3). When needles were used, about 0.05 g of needle tissue was ground into fine powder in liquid nitrogen using mortar and pestle. Then 0.14 M extraction buffer was added into the tubes containing the fine powder of the samples. The sample and the extraction buffer were mixed well. A paper wick (12x3.5 mm, Whatman chromatography paper, no.3 MM) was inserted into the extraction buffer to collect the sample.

2. Gel preparation: An 11% starch gel was used. The electrode buffer was diluted 20 times and used as gel buffer. The electrode buffer included 0.04 M citric acid and the pH was adjusted to 8.1 with N-(3-aminopropyl) morpholine A. The protocol of Conkle et al. (1982) for gel preparation and loading was utilized with the following modifications: 40 s heating in a microwave oven after the boiling water was added to the starch suspension to avoid premature solidification and to strengthen the gels; heating the vacuum flask on a hot plate while degassing; and using a spatula imMediately after pouring to remove air bubbles. If the gels were prepared the day before they were used, they were stored at room temperature covered with plastic film. The gels were kept at 4^0C for 1 hour before use.

3. Loading samples: The gel was cut into one small piece and one big piece. The smaller one was moved toward the edge of the gel glass until an opening at the origin was about 1 cm. The wicks were placed on the fresh-cut gel surface of the larger piece. The bottom of the wicks touched the gel glass. When all the wicks were in place, the smaller one was pushed back against the wicks on the larger one. The gel was covered with plastic wrap. The gel was connected to a power supply at 4^0C in a refrigerator to run. After the current had been on for 15 minutes, the power was turned off and the wicks

were removed from the gel. The gel without wicks was put back in refrigerator to complete the run.

4. Electrophoresis: The gels were run for 4 hours 30 minutes at a current of 60 to 65 mA.

5. Staining: When the gel run was finished, the gel was sliced into two pieces. The bottom piece was put in a staining buffer at 37^0C for at least 30 minutes in the dark when seeds were used. If needles were used, the gel was kept in the staining buffer in the dark overnight at room temperature following 37^0C for 30 minutes. The staining buffer includs 25 ml 0.2 M tris-HCL (pH is 8.0), 200 mg DL-isocitric acid, 2 ml 1% $MgCL_2$, 2 ml 10 mg/ml NADP, 2 ml 10 mg/ml NBT and 2 ml 1 mg/ml PMS.

3.3.3 Hybrid Analysis

The software NewHybrids version 1.1 beta (Anderson and Thompson, 2002; Anderson, 2003) was used to analyze AFLP and *IDH* data looking for hybrids. The software provides six genotype categories: pure species 1, pure species 2, F1, F2 hybrids of them, the first backcross generation to pure species 1 and the first backcross generation to pure species 2. The results show the estimated probability that each individual belongs to each different genotype category.

All the AFLP markers that produced polymorphic bands across the two different species were scored and used in the analysis. According to personal communication with Dr. Anderson, the author of software NewHybrids (2006), it is not necessary to select species-specific markers. He recommends use of all AFLPs that were polymorphic across the two species. The theory underlying NewHybrids allows the analysis of markers that are not necessarily perfect diagnostic.

3.4 Results

3.4.1 AFLP Markers

Ninety-six AFLP markers found to be polymorphic in both loblolly pine and shortleaf pine were produced using 17 primer pairs after 48 primer pairs were screened (Table 3.3). These primers were used for the analysis of all 93 shortleaf and 112 loblolly pine samples.

Table 3.3 The 96 AFLPs that are polymorphic in both loblolly pine and shortleaf pine

Primer pairs	Number of markers	markers
M-CCAGXE-ACG	22	S1[1] (80[2]), L1(95), A1(100), L2(120), L3(125), S2(145), L4(204), L5(230), L6(270), L7(275), A2(856), A3(99), A4(102), A5(104), A6(105), A7(110), A8(135), A9(140), A10(141), A11(148), A12(160), A13(240)
M-CCTGXE-ACG	12	S4(70), A24(80), A25(155), L9(204), L10(320), A27(78), A28(81), A29(101), A30(102), A31(120), A32(145), A33(254)
M-CCGAXE-ACG	12	L15(76), L16(90), A40(256), A41(300), A42(55), A43(98), A44(100), A45(105), A46(110), A47(120), A48(280), A49(290)
M-CCCGXE-ACG	10	A14(120), A15(215), L8(256), S3(270), A16(271), A17(946), A18(124), A19(200), A20(208), A21(255)
M-CCTCXE-ACC	9	A50(60), A51(75), L17(78), S7(80), A52(101), A53(142), A54(250), A55(68), A56(150)
M-CCGAXE-ACC	6	L13(70), L14(100), A36(80), A37(90), A38(125), A39(150)
M-CCGGXE-ACT	4	A58(145), S9(254), A59(55), A60(145)
M-CCTCXE-ACG	4	A65(111), A66(180), L19(345), L20(160)
M-CCTGXE-ACC	3	S5(105), L11(225), A26(275)
M-CCCAXE-ACG	3	L18(230), A61(80), A62(125)
M-CCTAXE-ACG	3	A63(90), S10(120), A64(142)
M-CCCGXE-ACA	2	A22(120), A23(133)
M-CCCAXE-ACT	2	L12(165), A35(202)
M-CCCAXE-ACT	1	S6(155)
M-CCCAXE-ACC	1	A34(144)
M-CCTTXE-ACC	1	S8(90)
M-CCGCXE-ACT	1	A57(150)

1: name of the marker; 2: estimated size of the marker

The first 5 primer pairs produced 9 or more AFLP markers and provided a majority of the information about the hybridization level between shortleaf and loblolly pine. They are very informative and would be useful in any further study of hybridization between these two species.

Figure 3.2 is a part of a typical AFLP gel picture produced by primer pair M-CCTCxE-ACG.

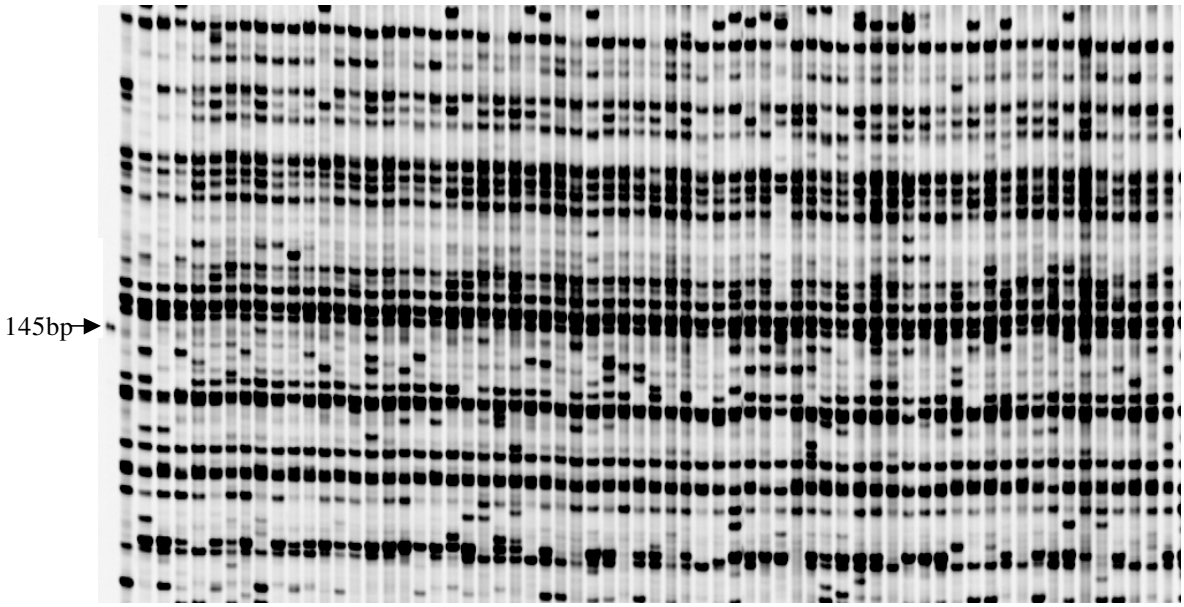

Figure 3.2 A part of the AFLP gel picture produced by primer pair M-CCTCxE-ACG
The 1st lane: a molecular standard, the 2nd lane: shortleaf pine Z15, the 3rd lane: loblolly pine 631, the 4th and 5th lanes: hybrids between Z15 and 631, the rest lanes: samples from the SSPSSS.

3.4.2 *IDH* Marker

Figure 3.3 is a picture of a stained *IDH* starch gel.

Figure 3.3 Picture of the *IDH* stained starch gel

Lane 1: a natural hybrid, lane 2, 3, 4, 5, 6, 7, 8, 12, 13, 14, 15, 16, 17: loblolly pine samples, line 9, 10, 11: shortleaf pine samples, arrow: indicting migration direction of *IDH*.

In Figure 3.3, The *IDH* band in loblolly pine samples migrates faster than that in shortleaf pine samples. The hybrid has an *IDH* band from both shortleaf pine and loblolly pine.

Two hybrids (327-2 and 321-4) were found in loblolly pine samples and two hybrids (433-1 and 433-2) were found in shortleaf pine samples by the *IDH* marker. Tree 433-1, 321-4 and 327-2 were found to be hybrids using seeds and 433-2 was detected using needles.

3.4.3 Hybrid Analysis

The 96 AFLPs and the *IDH* data were analyzed by NewHybrids version 1.1 beta (Anderson and Thompson, 2002; Anderson, 2003). Result is shown in appendix. According to the results, 433-2 has 96% probability, 481-7 has 92% probability and 481-9 has 72% probability of being a backcross to shortleaf. That is to say, two extra hybrids (481-7 and 481-9), which could not be detected by *IDH* alone, were found by combining the 96 AFLPs and the *IDH* data. Tree 433-2 can be found as hybrid by *IDH* alone or by combining the 96 AFLP and the *IDH* data.

In all six hybrids found in this study: 433-1, 433-2, 481-7 and 481-9 were from shortleaf pine samples, and 321-4 and 327-2 were from loblolly pine samples. Trees 433-1 and 433-2 came from seed source 433 (Dent, MO), so the estimated hybridization rate for this source is 25% (2/8). Trees 481-7and 481-9 were from source 481 (Ashley, AR) thus the hybridization rate is 20% (2/10) in this seed source. Tree 321-4 was from seed source 321 (Prentiss, MS), so the hybridization rate of this source is 10% (1/10). Tree 327-2 was from source 327 (Clark, AR), giving a source hybridization rate of 10% (1/10).

According to Figure 1, shortleaf pine seed source 433, 477, 481, 475 and 423 are located west of Mississippi River. All the hybrids (433-1, 433-2, 481-7 and 481-9) found in shortleaf pine samples in this study were from west of the Mississippi River. Accordingly, the shortleaf pine hybridization rate is 9.3% (4/43) west of the river and 0% east of the River. Loblolly seed source OSU and 327 are from west of the Mississippi River and one hybrid (327-2) was found in this area. The other hybrid in the loblolly pine samples (321-4) was from east of the Mississippi River. The hybridization rate in loblolly pine population is 4.5% (1/22) in the west and 1.1% (1/90) in the east. In all the samples, the hybridization rate is 7.7% (5/65) in western populations and 0.71% (1/140) in eastern populations.

The hybridization rate of the 93 shortleaf pine sampled throughout its range in this study is 4.3% (4/93), and the hybridization rate of the 112 loblolly pine samples is 1.79% throughout its range. In total, 2.9% (6/205) of the 205 samples were hybrids.

3.5 Discussion

The *IDH* isoenzyme is a codominant marker useful in detecting hybrids between shortleaf pine and loblolly pine (Ernest et al., 1990; Huneycutt and Askew, 1989; Edwards et al., 1997; Chen et al., 2004). To date, it is the only reliable locus thought to be fixed for different alleles between these two species. However, *IDH* will reliable detect only F1 hybrids. According to Mendelin genetics, 50% of the F2 hybrids will be homozygous at the *IDH* locus, as well as in the first backcross (BC1) generation. In other words, more than 50% later generation hybrids will be homozygous at *IDH* locus. Therefore, more markers are needed to identify later generation hybrids. This study demonstrated that AFLPs in conjunction with *IDH* with the help of special software can identify later generation hybrids that can not be detected by *IDH* alone.

However, AFLPs are dominant markers and they are not as informative as codominant markers. If the 4 *IDH* hybrids detected in this study were F2 or BC1 generation, on average, 4 more hybrids would be expected according to Mendelin genetics. If the 4 hybrids detected were later generation, beyond F2 or BC1, more than 4 additional hybrids would probably be in the populations. But, the large number of AFLPs (96 in this study) combined with the codominate *IDH* marker data only detected 2 more hybrids. In addition, the two hybrids could not be found if only AFLPs data were used and the AFLPs had to combine with the *IDH* data to detect hybrids using NewHybrids. Thus, more codominant markers, such as SSRs, may be needed to reliablly identify most or all later generation hybrids. In theory, only a few codominat markers should identify more hybrids than a large number of dominant markers.

Three of the four hybrids detected by *IDH* were not found by the AFLPs maybe because the dominant AFLPs are not as informative as codominant *IDH*. These three hybrids are probably later generation hybrids, whose quick recombination AFLP loci mask the hybrid nature while their *IDH* loci maintained a heterozygous status. Chen et al. (2004) and Edwards et al. (1997) also presented evidence that some hybrids found in their studies might be later generation hybrids.

According to the NewHybrid analysis data presented in the appendix, three of the *IDH* hybrids (433-1, 327-2 and 321-4) show low possibilities (1.3%, 0.9% and 0.5% respectively) to be hybrids. In Comparison, five trees, 477-8, 487-6, 451-7, 481-5 and 307-4, have an average 35.66% probability of being hybrids (backcross or F2). Moreover, more than 4 additional hybrids would be expected if 4 *IDH* hybrids were later generation, beyond F2 or BC1. We only found 2 additional hybrids. Thus, 477-8, 487-6, 451-7, 481-5 and 307-4 may be also later generation hybrids. Since their probability to be hybrids are lower than 50%, and they were not detected by *IDH* marker, more markers or software reliable beyond the F_2 and first backcross generation may be needed to identify their hybrid nature. These trees were not included in the results and discussion as hybrids.

This study found a relatively high level (15%) of hybridization between shortleaf pine and loblolly pine in Arkansas. This result is consistent with previous studies (15% by Raja et al., 1997; 14% by Chen et al., 2004). Of note, even though Raja et al. (1997) and this study used samples from SSPSSS plantings, the hybridization levels of some seed sources were surprisingly high, up to 25% in MO in this study and 34% in southeastern Arkansas in Raja et al.'s study (unpublished data), because these trees were

originally selected to represent the species in the SSPSSS tests. One possible reason is that the trees were originally selected based on their morphological traits. Since later generation hybrids, in particular backcross, often have a morphology similar to the backcross parents (Edwards et al., 1997; Chen et al., 2004), it is conceivable such hybrids were selected as representative of shortleaf pine or loblolly pine.

This study agrees with the other studies that the hybridization level was higher in populations west of Mississippi River than eastern populations. In our study, the hybridization rate was 9.3% west vs. 0% east in shortleaf populations, 4.5% west vs. 1.1% east in loblolly populations and 7.7% west vs. 0.7% east in all populations. Edwards and Hamrick (1995) found hybridization level at 4.6% west vs. 1.1% east in shortleaf pine. The different percentages of hybridization levels reported in different studies may be due to number, location, time of samples. Edwards and Hamrick (1995) pointed out that the day length and warm climate found in the western and lower latitudes of the shortleaf pine population might lead to more strobili maturation overlap between the two species and result in more opportunities to hybridize.

The hybrids found in this study were from Missouri, Mississippi and Arkansas. Edwards and Hamrick (1995) also detected hybrids from these three states. This study, Raja et al. (1997, unpublished data) and Edwards and Hamrick (1995) found hybrids in shortleaf pine populations far north of any natural loblolly pine populations. Schmidtling et al. (2005) pointed to two possible reasons leading to hybrids in allopatraic shortleaf pine populations. The first reason was that gene flow could be due to long-distance pollen transport. The second one was that the loblolly pine ranged farther north 5,000 to 7,000 years ago because the climate was warmer during the Hypsithermal geological

period. Possibly, the apparently later generation hybrids found in the allopatric region today result from F1s formed during the Hypsithermal geological period.

The hybridization frequency between the two species varied among populations from different places in all the studies. The hybridization level was relatively high in some locations (eg. Arkansas), which may have forest management implications.

Appendix: Probabilities of Each Sample Belonging to Different Genotypes

tree ID	shortleaf pine	loblolly pine	F1	F2	Backcross-shortleaf	Backcross-loblolly
433-1	0.98248	0	0	0.00442	0.01307	0.00002
433-2	0.01235	0	0.00006	0.02485	0.96268	0.00006
433-3	0.98945	0	0	0.00416	0.00635	0.00004
433-4	0.98552	0	0	0.00479	0.00961	0.00008
433-5	0.91027	0	0	0.00457	0.08516	0.00001
433-6	0.98795	0	0	0.00472	0.00664	0.00069
433-7	0.9897	0	0	0.00497	0.00524	0.0001
433-8	0.99101	0	0	0.00443	0.00455	0.00002
461-1	0.98477	0	0	0.00531	0.00909	0.00083
461-2	0.99042	0	0	0.00476	0.00481	0.00001
461-3	0.98671	0	0	0.00562	0.00742	0.00025
461-4	0.98707	0	0	0.00411	0.0088	0.00002
461-5	0.99172	0	0	0.00546	0.00277	0.00005
461-6	0.98719	0	0	0.00409	0.00867	0.00005
461-7	0.99196	0	0	0.00429	0.00375	0
461-8	0.98564	0	0	0.00402	0.01033	0
477-1	0.93854	0	0	0.00386	0.0576	0.00001
477-2	0.98538	0	0	0.00433	0.01027	0.00002
477-3	0.98507	0	0	0.00551	0.00942	0.00001
477-4	0.95646	0	0	0.00384	0.0397	0
477-5	0.9847	0	0	0.00566	0.00954	0.00011
477-6	0.98198	0	0	0.00401	0.014	0.00001
477-7	0.98513	0	0	0.00451	0.01032	0.00004
477-8	0.60006	0	0	0.00633	0.3936	0.00002
435-1	0.98646	0	0	0.0044	0.00909	0.00005
435-2	0.98957	0	0	0.0036	0.00683	0
435-3	0.98709	0	0	0.00437	0.0085	0.00003
435-4	0.98665	0	0	0.00578	0.0075	0.00007
435-5	0.98783	0	0	0.00457	0.00731	0.00029
435-6	0.99103	0	0	0.00501	0.00396	0
435-7	0.98745	0	0	0.00401	0.00854	0
435-8	0.99142	0	0	0.00418	0.00439	0.00002
435-12	0.9822	0	0	0.00415	0.01364	0.00001
487-1	0.98793	0	0	0.00554	0.00613	0.0004
487-2	0.97493	0	0	0.00404	0.02103	0
487-3	0.98526	0	0	0.00416	0.01057	0.00001
487-4	0.98775	0	0	0.00383	0.00841	0
487-5	0.98188	0	0	0.00411	0.01399	0.00002

tree ID	shortleaf pine	loblolly pine	F1	F2	Backcross-shortleaf	Backcross-loblolly
487-6	0.58599	0	0	0.00485	0.40916	0
487-7	0.99031	0	0	0.00527	0.0044	0.00001
487-8	0.98807	0	0	0.00329	0.00864	0
487-9	0.98526	0	0	0.00489	0.00983	0.00002
451-1	0.99059	0	0	0.00485	0.00455	0.00001
451-2	0.99219	0	0	0.00303	0.00478	0
451-3	0.99204	0	0	0.00542	0.00246	0.00008
451-4	0.99169	0	0	0.00226	0.00605	0
451-5	0.9927	0	0	0.00241	0.00489	0
451-6	0.8858	0	0	0.00228	0.11192	0
451-7	0.6838	0	0	0.00251	0.31369	0
451-8	0.98851	0	0	0.00397	0.00752	0
451-9	0.99013	0	0	0.00553	0.00434	0
451-10	0.98767	0	0	0.0038	0.00853	0
475-1	0.98285	0	0	0.00575	0.01124	0.00017
475-2	0.98744	0	0	0.00499	0.00667	0.00091
475-3	0.98892	0	0	0.00425	0.00682	0.00001
475-4	0.98918	0	0	0.00412	0.0067	0
475-5	0.99002	0	0	0.00536	0.00399	0.00063
475-6	0.98783	0	0	0.00455	0.00755	0.00006
475-7	0.98963	0	0	0.00565	0.00454	0.00019
475-8	0.98647	0	0	0.00424	0.00893	0.00036
475-9	0.98864	0	0	0.00289	0.00847	0
475-10	0.98909	0	0	0.00457	0.00619	0.00014
481-1	0.98107	0	0	0.00461	0.0128	0.00152
481-2	0.98843	0	0	0.0034	0.00817	0
481-3	0.99135	0	0	0.00501	0.00361	0.00002
481-4	0.99021	0	0	0.00417	0.00562	0
481-5	0.60723	0	0.00001	0.00479	0.38795	0.00002
481-6	0.98822	0	0	0.00466	0.00707	0.00004
481-7	0.03353	0	0	0.04629	0.9197	0.00048
481-8	0.99133	0	0	0.00362	0.00504	0.00001
481-9	0.25951	0	0	0.01277	0.72771	0
481-10	0.99022	0	0	0.00536	0.00442	0.00001
423-1	0.83009	0	0	0.00417	0.16572	0.00001
423-2	0.9909	0	0	0.00384	0.00524	0.00001
423-3	0.9622	0	0	0.00581	0.0318	0.00019
423-4	0.99087	0	0	0.00418	0.0049	0.00005
423-5	0.97382	0	0	0.00283	0.02335	0
423-6	0.9793	0	0	0.00529	0.0154	0.00001
423-7	0.98966	0	0	0.00531	0.00434	0.00068
401-1	0.98275	0	0	0.00516	0.01204	0.00004
401-2	0.98978	0	0	0.00417	0.00602	0.00003

tree ID	shortleaf pine	loblolly pine	F1	F2	Backcross-shortleaf	Backcross-loblolly
401-3	0.9864	0	0	0.00416	0.00943	0.00001
401-4	0.98769	0	0	0.00342	0.00888	0
419-1	0.99165	0	0	0.0038	0.00455	0.00001
419-2	0.99136	0	0	0.00567	0.00294	0.00004
419-3	0.98846	0	0	0.00327	0.00827	0
419-4	0.97498	0	0	0.00572	0.01927	0.00003
419-5	0.99019	0	0	0.00565	0.00399	0.00017
421-1	0.98407	0	0	0.00354	0.01239	0
421-2	0.98201	0	0.00001	0.00421	0.01374	0.00003
421-3	0.98784	0	0	0.00466	0.00743	0.00006
421-4	0.98803	0	0	0.00536	0.0062	0.00041
421-5	0.98399	0	0.00001	0.00375	0.01225	0
329-1	0	0.99207	0	0.00098	0	0.00695
329-2	0	0.99539	0	0.00019	0	0.00442
329-3	0	0.99346	0	0.00087	0	0.00567
329-4	0	0.99524	0	0.00001	0	0.00475
329-5	0	0.98905	0	0.00061	0	0.01034
329-6	0	0.99092	0	0.0007	0	0.00838
329-7	0	0.99238	0	0.00136	0	0.00626
329-8	0	0.99325	0	0.00076	0	0.006
329-9	0	0.99495	0	0.00056	0	0.00449
329-10	0	0.99419	0	0.00005	0	0.00576
323-1	0	0.99151	0	0.00497	0	0.00353
323-2	0	0.99209	0	0.00395	0	0.00396
323-3	0	0.98612	0	0.00197	0	0.01191
323-4	0	0.99306	0	0.00018	0	0.00676
323-5	0	0.99509	0	0.00033	0	0.00458
323-6	0	0.99517	0	0.00001	0	0.00482
323-7	0	0.9929	0	0.00025	0	0.00685
323-8	0	0.9935	0	0.0032	0	0.00331
323-9	0	0.99224	0	0.00037	0	0.0074
323-10	0	0.99246	0	0.00287	0	0.00467
331-1	0	0.99313	0	0.00277	0	0.0041
331-2	0	0.99173	0	0.00006	0	0.00821
331-3	0	0.99155	0	0.00054	0	0.00791
331-4	0	0.99035	0	0.00173	0	0.00791
331-5	0	0.99202	0	0.00038	0	0.0076
331-6	0	0.99047	0	0.00157	0	0.00795
331-7	0	0.99237	0	0.00255	0	0.00508
331-8	0	0.99354	0	0.00001	0	0.00645
331-9	0	0.99131	0.00001	0.00269	0	0.00599
331-10	0	0.99373	0	0.00041	0	0.00585
311-1	0	0.991	0	0.00283	0	0.00617

tree ID	shortleaf pine	loblolly pine	F1	F2	Backcross-shortleaf	Backcross-loblolly
311-2	0	0.98884	0	0.00307	0	0.00808
311-3	0	0.99111	0	0.00369	0	0.00519
311-4	0	0.99242	0	0.0017	0	0.00588
311-5	0	0.99243	0	0.00096	0	0.00661
311-6	0	0.99021	0	0.00164	0	0.00815
311-7	0	0.99084	0	0.00328	0	0.00589
311-8	0	0.99452	0	0.00018	0	0.0053
311-9	0	0.9938	0	0.0006	0	0.0056
311-10	0	0.99105	0	0.00127	0	0.00768
FL-1	0	0.99046	0	0.00008	0	0.00946
FL-2	0	0.99163	0	0.00092	0	0.00746
FL-3	0	0.99054	0	0.00062	0	0.00883
FL-4	0	0.99355	0	0.0006	0	0.00585
FL-5	0	0.99004	0	0.00395	0	0.006
FL-6	0	0.98932	0	0.00092	0	0.00976
FL-7	0	0.98789	0	0.00163	0	0.01047
FL-8	0	0.98951	0	0.0021	0	0.00839
FL-9	0	0.98975	0	0.00021	0	0.01005
FL-10	0	0.98598	0	0.00281	0	0.0112
303-1	0	0.99335	0	0.00003	0	0.00661
303-2	0	0.99221	0	0.00037	0	0.00742
303-3	0	0.94483	0	0.00234	0	0.05283
303-4	0	0.99281	0	0.00113	0	0.00606
303-5	0	0.9921	0	0.00149	0	0.00641
303-6	0	0.99348	0	0.00008	0	0.00644
303-7	0	0.99458	0	0.00009	0	0.00533
303-8	0	0.99129	0	0.00321	0	0.0055
303-9	0	0.99322	0	0.00134	0	0.00544
307-1	0	0.99238	0	0.00177	0	0.00585
307-2	0	0.99303	0	0.0008	0	0.00618
307-3	0	0.9928	0	0.00096	0	0.00623
307-4	0	0.49694	0.00003	0.27884	0.00001	0.22418
307-5	0	0.98326	0	0.00146	0	0.01528
307-6	0	0.98542	0	0.00115	0	0.01343
307-7	0	0.99133	0	0.00033	0	0.00834
307-8	0	0.98966	0.00002	0.003	0	0.00732
307-9	0	0.99213	0	0.00078	0	0.00709
307-10	0	0.99002	0	0.00092	0	0.00906
321-1	0	0.9914	0	0.00477	0	0.00384
321-2	0	0.99443	0	0.00019	0	0.00539
321-3	0	0.99262	0	0.00056	0	0.00682
321-4	0	0.99377	0	0.00078	0	0.00545
321-5	0	0.99244	0	0.0032	0	0.00436

tree ID	shortleaf pine	loblolly pine	F1	F2	Backcross-shortleaf	Backcross-loblolly
321-6	0	0.99364	0	0.00085	0	0.00551
321-7	0	0.99205	0	0.00194	0	0.00601
321-8	0	0.99347	0	0.00102	0	0.00551
321-9	0	0.94008	0.00003	0.00278	0	0.05711
321-10	0	0.991	0	0.00253	0	0.00647
317-1	0	0.97924	0	0.00462	0	0.01613
317-2	0	0.98774	0	0.00072	0	0.01154
317-3	0	0.99162	0	0.00261	0	0.00577
317-4	0	0.99243	0	0.00197	0	0.0056
317-5	0	0.99227	0	0.00052	0	0.00721
317-6	0	0.99148	0	0.00141	0	0.00711
317-7	0	0.98968	0	0.00384	0	0.00648
317-8	0	0.99436	0	0.00078	0	0.00486
317-9	0	0.99462	0	0.00086	0	0.00452
317-10	0	0.98988	0	0.00303	0	0.00709
317-11	0	0.98805	0	0.00258	0	0.00937
327-1	0	0.99384	0	0.00089	0	0.00527
327-2	0	0.98983	0	0.00127	0	0.0089
327-3	0	0.9905	0	0.00129	0	0.00821
327-4	0	0.99266	0	0.00245	0	0.00489
327-5	0	0.99497	0	0.00075	0	0.00428
327-6	0	0.99149	0	0.00181	0	0.0067
327-7	0	0.99254	0	0.00136	0	0.0061
327-8	0	0.99351	0	0.00273	0	0.00377
327-9	0	0.99364	0	0.00042	0	0.00595
327-10	0	0.91641	0	0.04005	0	0.04354
327-11	0	0.99237	0	0.00097	0	0.00666
OSU-1	0	0.99461	0	0.0008	0	0.00459
OSU-2	0	0.99212	0	0.00357	0	0.00431
OSU-3	0	0.99107	0	0.00045	0	0.00847
OSU-4	0	0.99299	0	0.00007	0	0.00695
OSU-5	0	0.9945	0	0.00003	0	0.00546
OSU-6	0	0.99262	0	0.00058	0	0.00681
OSU-7	0	0.99457	0	0.00001	0	0.00542
OSU-8	0	0.99151	0	0.00214	0	0.00635
OSU-9	0	0.99299	0	0.00009	0	0.00692
OSU-10	0	0.99214	0	0.00123	0	0.00663
OSU-11	0	0.9927	0	0.00135	0	0.00595

References

Anderson, E. C., and Thompson, E. A. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics. 160: 1217-1229.

Anderson, E. C. 2003. User's guide to the program newhybrids version 1.1 beta. http://ib.berkeley.edu/laboratorys/slatkin/eriq/software/software.htm

Chen, J. W., Tauer, C. G., Bai, G., Huang, Y., Payton, M. E., and Holley, A. G. 2004. Bidirectional introgression between Pinus taeda and Pinus echinata: Evidence from morphological and molecular data. Can. J. For. Res. 34: 2508-2516.

Conkle, M. T., Hodgskiss, P. D., Nunnally, L. B., and Hunter, S. C. 1982. Starch gel electrophoresis of conifer seeds: a laboratory manual. USDA For. Serv. Tech. Rep. PSW-64. Washington, DC: USDA Forest Service.

Cotton, M. H., Hicks, R. R., and Flake, R.H. 1975. Morphological variability among loblolly and shortleaf pines of east Texas with reference to natural hybridization. Castanea. 40:309-319.

Dorman, K. W., and Barber, J. C., 1956. Time of flowering and seed ripening in southern pines. USDA Forest Service, Southeastern Forest Experiment Station, Station Paper 72.

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

Edwards, M. A., and Hamrick, J. L. 1995. Genetic variation in shortleaf pine, Pinus echinata Mill. (Pinaceae). For. Genet. 2: 21-28.

Edwards, M. A., Hamrick, J. L., and Price, R. A., 1997. Frequency and direction of hybridization in sympatric populations of Pinus taeda and P. echinata (Pinaceae). Am. J. Bot. 84: 879–886.

Hare, R. C., and Switzer, G. L. 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.

Henry, B. W., and Bercaw, T. E. 1956. Shortleaf-loblolly hybrid pines free of fusiform rust after 5 years' exposure. J. For. 54: 779.

Hicks, R. R. 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., and Askew, G. R. 1989. Electrophoretic identification of loblolly pine shortleaf pine hybrids. Silvae Genet. 38: 95-96.

Little, E. L., Jr., and Righter, F. I. 1965. Botanical descriptions of forty artificial pine hybrids. USDA For. Ser. Tech. Bull. 1345: pp47.

Mergen, F., Stairs, G. R., and Snyder, E. B. 1965. Natural and controlled loblolly x shortleaf pine hybids in Mississippi. Forest Sci. 11: 306-314.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. H. 1997. Isoenzyme variation and genetic structure in natural populations of shortleaf pine (Pinus echinata). Can. J. For. Res. 27: 740-749.

Raja, R. G., Tauer, C. G., Wittwer, R. F., and Huang, Y. 1998. Regeneration methods affect genetic variation and structure in Shortleaf Pine (Pinus echinata Mill.) For. Genet. 5: 171-178.

Remington, D. L., and O'Malley, D. M. 2000. Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. Genetics. 155: 337-348.

Remington, D. L., Whetten, R. W., Liu, B. H., and O'Malley, D. M. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in P*inus taeda*. Theor. Appl. Genet. 98: 1279-1292.

Saylor, L. C. 1972. Karyotype analysis of the genus Pinus-subgenus Pinus. Silvae Genet. 21: 155-163.

Schmidtling, R. C., Myszewski, J. H., and McDaniel, C. E. 2005. Geographic variation in shortleaf pine (*Pinus echinata Mill*.) - Cortical monoterpenes. Proceedings to Southern Forest Tree Improvement Conference 28: 161-167. Jun. 20-23, Raleigh, North Carolina.

Schreiner, E. J. 1937. Improvement of forest trees. 1937 Yearbook of Agriculture. USDA. pp1242-1279.

Watts, W. A. 1983. A vegetational history of the eastern United States 25,000 to 10,000 years ago. In: S.C. Porter (ed) The late Pleistocene. Late-Quaternary Environments of the United States, pp294-321. University of Minnesota Press, Minneapolis.

Wendel, J. F., and Parks, C. R. 1982. Genetic control of isozyme variation in Camalia japonica L. J. Hered. 73: 197-204.

Yeh, F. C., and Layton, C. 1979. The organization of genetic variability in central and marginal populations of lodgepole pine Pinus contorta spp. latifolia. Can. J. Genet. Cytol. 21: 487-503.

Zobel, B. J. 1953. Are there natural loblolly-shortleaf pine hybrids? J. For. 51: 494–495.

VITA

Shiqin Xu

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STUDY OF GENETIC DIVERSITY AND HYBRIDIZATION IN NATURAL STANDS OF SHORTLEAF (*PINUS ECHINATA* M.) AND LOBLOLLY (*PINUS TAEDA* L.) PINE POPULATIONS

Major Field: Plant Science

Biographical:

Person Data: Born in Shiyan City, Hubei Province, the People's Republic of China on June 27, 1972, the daughter of Minggao Xu and Yunfeng San.

Education: Graduated from No.1 Yunyang High school, Shiyan City, Hubei province in July 1990; received Bachelor of Science degree in Vegetable Science from Huazhong Agricultural University, Wuhan, Hubei province in July 1994; obtained Master of Science degree in Vegetable Breeding from Huazhong Agricultural University, Wuhan, Hubei province in July 1997; completed the requirements for the Doctor of Philosophy degree with a major in Plant Science in Oklahoma State University in December, 2006.

Experience: January 2003 to December 2006: Graduate Research Assistant, Department of Forestry, Oklahoma State University, Stillwater, Oklahoma, USA; July 1997 to April 2001: Agronomist, Shanghai Municipal Seed Company, Shanghai, the People's Republic of China.

Name: Shiqin Xu Date of Degree: December, 2006

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: GENETIC DIVERSITY AND HYBRIDIZATION IN NATURAL STANDS OF SHORTLEAF (*PINUS ECHINATA* M.) PINE AND LOBLOLLY (*PINUS TAEDA* L.) PINE

Pages in Study: 111 Candidate for the Degree of Doctor of Philosophy

Major Field: Plant Science

Scope and Method of Study: molecular markers (AFLPs and *IDH*), population genetics, conversation, forest genetics

Findings and Conclusions: Genetic diversity in natural stands of shortleaf pine and loblolly pine is high. The majority (over 80%) of this genetic diversity is found within subpopulations and less than 20% is found among subpopulations. The subpopulations in both shortleaf pine and loblolly pine have similar level of genetic diversity. The populations located east of the Mississippi River and those west of the River have similar level of genetic diversity. Gene flow is high among subpopulations and between populations east and west of the Mississippi River. No apparent relationship exists between populations' geographic distances and genetic distances. The hybridization rate between the two species varies from place to place, ranging from 25% in MO to 0 in other populations in this study. The hybridization frequency is higher in populations west of Mississippi Rive than in those east of the River. *IDH* is a useful marker to detect hybrids of shortleaf and loblolly pine. AFLPs are good markers for genetic diversity and structure study, but not efficient in finding later generation hybrids. More codominant markers are needed to detect later generation hybrids. The results of this study may have important forest management implications.

ADVISER'S APPROVAL: Charles G. Tauer