Tree Genetics & Genomes
DOI 10.1007/s11295-008-0158-9

ORIGINAL PAPER

Genetic diversity within and among populations of shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.)

Shiqin Xu · C. G. Tauer · C. Dana Nelson

Received: 20 December 2007 / Revised: 2 April 2008 / Accepted: 14 April 2008

Abstract Shortleaf pine (n=93) and loblolly pine (n=112) trees representing 22 seed sources or 16 physiographic populations were sampled from Southwide Southern Pine Seed Source Study plantings located in Oklahoma, Arkansas, and Mississippi. The sampled trees were grown from shortleaf pine and loblolly pine seeds formed in 1951 and 1952, prior to the start of intensive forest management across their native ranges. Amplification fragment length polymorphism (AFLP) markers were developed and used to study genetic diversity and its structure in these pine species. After screening 48 primer pairs, 17 and 21 pairs were selected that produced 794 and 647 AFLPs in shortleaf pine and loblolly pine, respectively. High-AFLP-based genetic diversity exists within shortleaf pine and loblolly pine, and most (84.73% in shortleaf pine; 87.69% in loblolly pine) of this diversity is maintained within physiographic populations. The high value of unbiased measures of genetic identity and low value of genetic distance for all pairwise comparisons indicates that the populations have similar genetic structures. For shortleaf pine, there was no significant correlation between geographic distance and genetic distance (r=0.28), while for loblolly pine there was a weak but significant correlation (r=0.51).

Keyword Genetic diversity

Communicated by R. Sederoff

S. Xu · C. G. Tauer (✉)
Department of Natural Resource Ecology and Management,
Oklahoma State University,
Stillwater, OK 74078, USA
e-mail: chuck.tauer@okstate.edu

C. D. Nelson
Southern Institute of Forest Genetics, USDA Forest Service,
Saucier, MS 39574, USA

Introduction

Shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L) are important species for forest products in the southeast US. These species also provide habitat for wildlife and other environmental amenities, including soil stabilization, clean water and air, and carbon sequestration. Both species are widely distributed over most of the southeastern US (Fig. 1), suggesting that they possess a large amount of genetic variation due to adaptation to a variety of environments.

Early studies of natural variation in these two pine species necessarily relied on morphological traits. The Southwide Southern Pine Seed Source Study (SSPSSS), established in the 1950s, includes both species and has provided much of the early range-wide information, especially for shortleaf pine.

For loblolly pine, many additional studies of geographic variation have been reported. Schultz (1997) clearly summarized these, including the SSPSSS, stating that the species possesses considerable natural variation for many morphological traits. He reported that variation is generally clinal in nature, extending both north to south and east to west for many growth traits. For other traits, he reported ecotypic adaptation or the lack of significant variability. Some differences may exist between populations east and west of the Mississippi River for some traits, including growth rate, rust resistance, and drought tolerance.

The study of geographic variation in shortleaf pine is considerably less extensive and, except for the SSPSSS results, little has been reported. Through age 10, Wells and Wakley (1970) found no geographic pattern for survival, although northern sources survived best in northern plantings. They did report a clear relationship between growth and temperature at the seed source, with southern sources

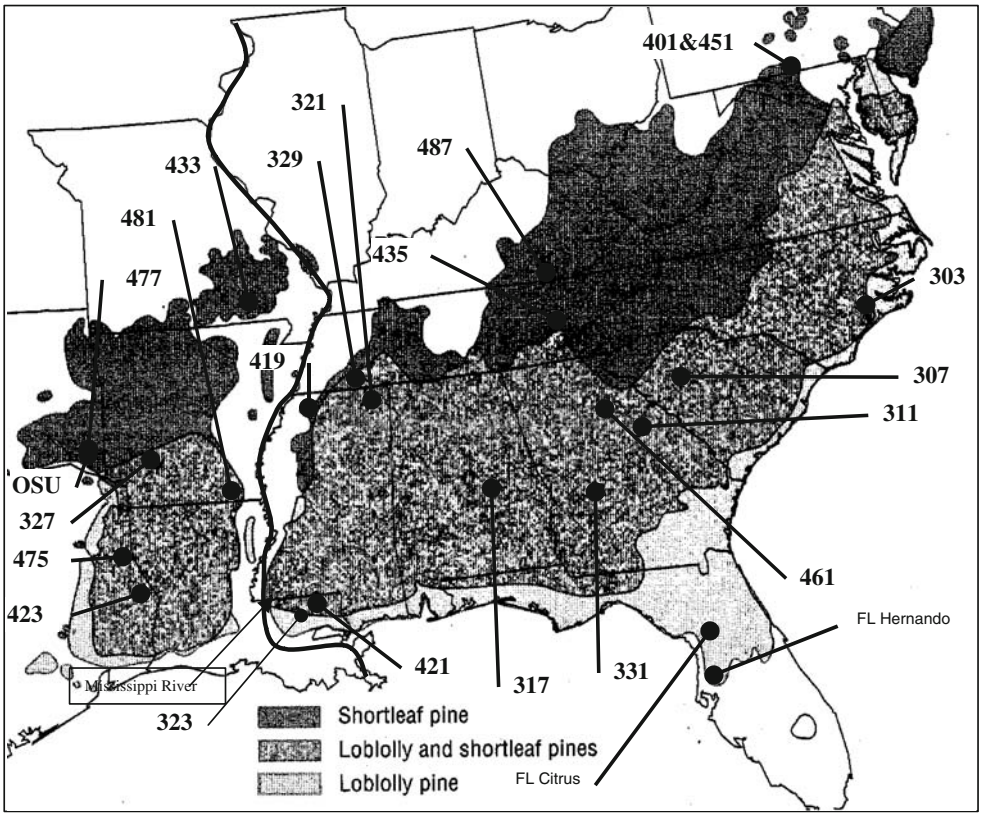

Fig. 1 The natural ranges of shortleaf pine and loblolly pine and approximate sample origins (300s are loblolly pine and 400s are shortleaf pine)

growing faster unless moved too far north. In a SSPSSS planting in Mississippi (Wells 1973), the only consistent genetic difference between eastern and western populations of shortleaf pine was in time of growth initiation, with sources west of the Mississippi River initiating growth earlier. Tauer (1980), reporting on 20-year results of two SSPSSS shortleaf pine plantings in Oklahoma, noted a north–south trend in growth but no discernable east–west trend. The relative performance of sources did not change after age 10.

Morphological studies were soon followed by biochemical (monoterpenes and isozymes) studies of genetic diversity. These studies generally confirmed the morphological data, showing north–south and east–west gradients. Some east to west of the Mississippi River differences in loblolly pine (Wells and Lambeth 1983; Wells and Wakley 1970; Florence and Rink 1979) were attributed to the presence of the river itself, while Schmidtling et al. (1999) and Wells et al. (1991) later attributed the differences to separate east and west glacial refugia during the Pleistocene. There were few differences noted for shortleaf pine when east and west of the river populations were compared (Edwards and Hamrick 1995; Raja et al. 1997), the notable exception being that the frequency of isocitrate dehydrogenase heterozygosity was higher west of the river. Heterozygosity at this locus was thought to indicate that the tree is a shortleaf pine × loblolly pine hybrid (Huneycutt and Askew 1989), but Xu et al. (2008) show that this marker is not reliable.

DNA marker technologies allow closer scrutiny of these differences. Amplification fragment length polymorphism (AFLP) markers are useful for studying population genetics in trees (Muluvi et al. 1999) because their use requires no previous sequence knowledge, has good repeatability, and can detect multiple loci. In this study, we used AFLPs to quantify and describe the genetic diversity found in natural shortleaf pine and loblolly pine populations sampled across their native ranges. This study also examines east–west genetic variation as defined by the Mississippi River.

Loblolly pine grows faster than shortleaf pine (except on the driest sites) and is generally preferred to shortleaf pine in artificial regeneration (Schultz 1997). More and more shortleaf pine is being replaced with improved loblolly pine on intensively managed forest lands (South and Buckner 2003). The US Forest Service (USFS) 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 USFS 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. The effect of such a high hybridization level on species integrity in the long term is unknown. The trees sampled for this study were derived from seeds collected in 1951 and 1952, when human influence due to management (specifically, replacing large tracts shortleaf pine with loblolly pine) was minimal. Thus, this study estimates genetic variation found in natural populations of shortleaf pine and loblolly pine approximately 50 years ago, that is, prior to intensive forest management. Our research should provide a reference or base-level dataset for addressing questions concerning diversity and hybridization changes between the 1950s, the present, and into the future.

Materials and methods

Needles and cones of shortleaf pine and loblolly pine were collected from 11 seed sources each (Fig. 1). The SSPSSS samples included 11 seed sources of shortleaf pine and nine seed sources of loblolly pine. The SSPSSS seed sources were created in 1951 and 1952 by collecting cones from 20 or more trees at each origin with the resulting seeds being mixed by source prior to planting. Two additional loblolly pine seed sources were collected; one was an Oklahoma State University (OSU) collection from Oklahoma source seed orchard selections made in the 1970s and 1980s with select tree ages ranging from 25 to 40 years old and the other a Florida collection made in 2005 from mature trees of a Florida source. The original locations of the seed sources and sample sizes obtained for this study are given in Table 1.

Collected needles were placed in plastic bags and kept on blue ice in coolers during overnight shipment. Upon arrival in the laboratory, the needles were frozen at −80°C for later use.

AFLP analysis

Total DNA was extracted from needles of the shortleaf pine samples using a modified CTAB protocol (Doyle and Doyle 1988) of C. G. Tauer's laboratory. A DNeasy Plant Mini Kit (QIAGEN, Inc.) was used for isolation of DNA from needle tissue of each loblolly pine sample. The primers and the AFLP protocols described by Remington et al. (1999) and Remington and O'Malley (2000) were utilized in this study, including the use of *Eco*RI and *Mse*I restriction enzymes and the same 48 primer pairs. The primer pairs include all combinations of the selective nucleotides *for Eco*RI primers (5'-ACA-3', 5'-ACC-3', 5'-ACG-3' and 5'-ACT-3') with the selective nucleotides for *Mse*I primers (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').

A LI-COR 4300 DNA Analyzer was used for AFLP fragment separation and detection. A single gel holds 64 samples; we therefore loaded the 205 samples per primer pair onto four gels. Samples from four nonstudy trees (one each shortleaf pine, Z15, and loblolly pine, SE631 (also designated GFC-631), and two interspecies F1 hybrids) were used as controls on each gel to ensure that the same loci were scored for all 205 samples loaded across the

Table 1 The origin and sample size of the shortleaf pine and loblolly pine sources in this study

Shortleaf pine[a]				Loblolly pine[a]			
Source ID	State	County	No of tress	Source ID	State	County	No of tress
401[b]	PA	Franklin	4	303	NC	Onslow	9
419	MS	Lafayette	5	307	SC	Newberry	10
421	LA	St. Helena	5	311	GA	Clarke	10
423	TX	Angelina	7	317	AL	Clay	11
433	MO	Dent	8	321	MS	Prentiss	10
435	TN	Morgan	9	323	LA	Livingston	10
451[b]	PA	Franklin	10	327	AR	Clark	11
461	GA	Clarke	8	329	TN	Hardeman	10
475	TX	Cherokee	10	331	GA	Spalding	10
477	OK	Pushmataha & McCurtain	8	OSU[c]	OK	McCurtain	11
481	AR	Ashley	10	FL[d]	FL	Hernando, Citrus	10
487	TN	Anderson	9				

[a] Ninety-three shortleaf pine and 92 loblolly pine samples were collected by Oklahoma State University Forest Resources Center personnel, Idabel, OK, USA. In addition, ten loblolly pine samples (of comparable age to the SSPSSS samples, i.e., average age=56, range 34–67) from an allopatric region of Florida were provided by Gregory Powell, University of Florida, Gainesville, FL, USA

[b] 401 belongs to the original collection made in 1951 and 451 to the collection made in 1955

[c] Not part of the SSPSSS, rather a local collection of equivalent age

[d] Present-day collection from allopatric region

multiple gels. At any specific locus, band presence was scored as "1" and band absence "0".

Data analysis

Not many SSPSSS plantings remain (all data are available from C. D. Nelson by request), and of those that do, some seed sources have as few as four trees remaining. For example, in this study, we were able to sample only five trees each from sources 419 and 421. To properly represent populations with adequate sample sizes, the seed sources within the same physiographic regions were grouped as follows: 419 and 421 (northwest Mississippi and southeast Louisiana), 423 and 475 (both southeast Texas), 435 and 487 (both east Tennessee), 311 and 307 (piedmont, east Georgia and South Carolina), 331 and 317 (piedmont west Georgia and Alabama), and 329 and 321 (west Tennessee and northeast Mississippi). In total, the 11 seed sources were pooled such that eight physiographic populations were formed within each species.

Genetic variation was estimated at the level of species, regions (defined as east or west of Mississippi River) and physiographic population. The region west of the Mississippi River included 43 shortleaf pine samples (from seed sources 433, 481, 477, 475, and 423) and 22 loblolly pine samples (from sources OSU and 327), and the region east of the river included 50 shortleaf pine samples (from seed sources 401, 451, 487, 435, 461, 419, and 421) and 80 loblolly pine samples (from sources 329, 321, 317, 331, 311, 307, 303, and 323; Fig. 1).

Several different analyses using POPGENE version 1.31 (Yeh and Boyle 1997) were used to examine genetic variation at each of the three levels: species, regions, and populations. The POPGENE software recognizes marker type and the estimates are appropriate to the dominant nature of AFLPs. First, AFLP marker diversity was calculated using the following estimates: percentage of polymorphic loci, which includes all polymorphic loci regardless of allele frequencies (p), observed number of alleles (na), effective number of alleles (the reciprocal of homozygosity, Hartl and Clark 1989; ne) and average heterozygosity (h) estimated with the assumption that a large number of loci are used and the average heterozygosity is low (Nei 1978). Also, the Ewens–Watterson test (Manly 1985) was used to test polymorphic loci's selective advantage, disadvantage, or neutrality.

Second, F statistics were used to examine genetic variation among and within populations and regions. Gene diversity in the species (Ht) is the sum of average gene diversity among populations (Dst) and average gene diversity within populations (Hs), where $Ht = Hs + Dst$. The

Table 2 AFLP primer pairs producing polymorphic loci in shortleaf pine and loblolly pine

Primer Pair	Shortleaf pine			Loblolly pine		
	# of loci	# of polymorphic loci	% Polymorphic loci	# of loci	# of polymorphic loci	% Polymorphic loci
M-CCTGxE-ACG	60	54	90.00	63	60	95.24
M-CCGAxE-ACG	41	35	85.37	26	19	73.08
M-CCAGxE-ACG	59	48	81.36	55	47	85.45
M-CCCGxE-ACA	67	54	80.60	31	20	64.52
M-CCCGxE-ACG	45	36	80.00	17	11	64.70
M-CCGAxE-ACC	30	21	70.00	15	9	60.00
M-CCTTxE-ACG	49	32	65.31	45	32	71.11
M-CCTAxE-ACG	63	38	60.32	27	17	62.96
M-CCGAxE-ACT	30	20	66.67	24	11	45.83
M-CCCAxE-ACG	47	31	65.96	22	12	54.54
M-CCTGxE-ACC	33	21	63.64	38	17	44.74
M-CCGGxE-ACT	36	21	58.33	18	6	33.33
M-CCGAxE-ACA	16	7	43.75	11	2	18.18
M-CCGCxE-ACT	31	11	35.48	16	4	25.00
M-CCTCxE-ACC	56	12	21.43	39	6	15.38
M-CCTTxE-ACC	32	6	18.75	25	2	8.00
M-CCTCxE-ACG	99	76	76.77			
M-CCTCxZ-ACG				21	12	57.14
M-CCAGxE-ACA				37	9	24.32
M-CCTCxE-ACT				48	5	10.42
M-CCCAxE-ACT				29	1	3.45
M-CCCAxE-ACC				40	1	2.50
Total	794	523	65.87	647	303	46.68

relative amount of gene differentiation among populations was measured by the coefficient of gene differentiation, Gst, where Gst=Dst/Ht.

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

Results

The 17 selected AFLP primer pairs for shortleaf pine produced 794 loci, of which 523 were polymorphic across the 93 trees sampled. For loblolly pine, the 21 selected primer pairs produced 647 loci, of which 303 were polymorphic across the 112 trees sampled (Table 2).

The top eight primer pairs produced at least 60% polymorphic loci in both species, providing a relatively high efficiency rate for producing AFLP markers. The details of the primer pair results and the specific markers are available from C. G. Tauer upon request. Figure 2 is a typical AFLPs gel image of 64 samples produced by primer pair M-CCAGxE-ACG.

The Ewens–Watterson test was used to test locus neutrality at the level of eight populations in both species. In shortleaf pine, 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 not favored by selection and five loci (loci ID: 608, 609, 613, 576, and 632) were favored by selection. In loblolly pine, 633 of the 647 loci tested were selectively neutral; ten loci (loci ID: 85, 87, 88, 192, 290, 485, 513, L6, A62, and A66) were not favored by selection and four loci (loci ID: 5, 11, 123, and 132) were favored by selection. The same test was applied to regions west (43 samples in shortleaf pine and 22 samples in loblolly pine) and east (50 samples in shortleaf pine and 80 samples in loblolly pine) of the Mississippi River. At the regional level in shortleaf pine, 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 not favored by selection and seven loci (loci ID: 86, 549, 576, 608, 609, 613, and 632) were selectively favored. In

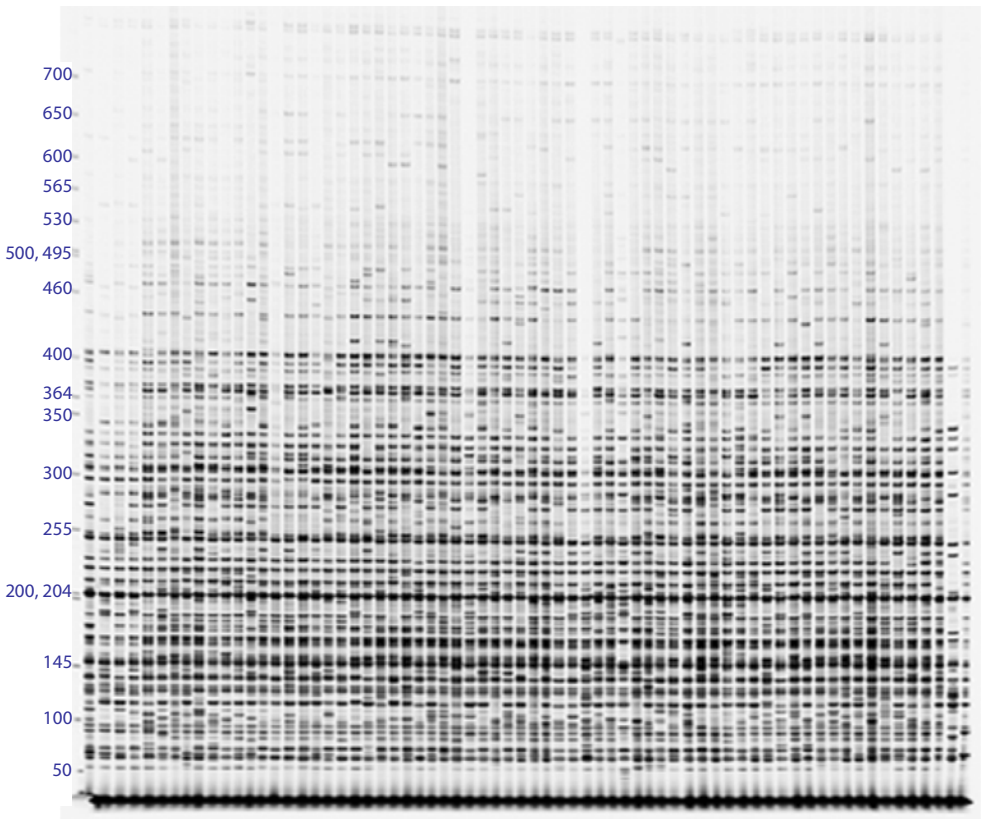

Fig. 2 An AFLP gel image of shortleaf pine samples produced by primer pair M-CCAGxE-ACG

Lane 1: a molecular standard; lane 2: shortleaf pine Z15; lane 3: loblolly pine 631; lanes 4 & 5: hybrids between Z15 and 631; the remaining lanes: loblolly pine samples from the SSPSSS planting

loblolly pine, 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 not favored by selection and 4 loci (loci ID: 5, 11, 123, and 132) were favored by selection.

For shortleaf pine, the overall percentage of polymorphic loci was 65.87% (Table 3), 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 (43.71%) was much lower than that within the species; the observed number of alleles (1.44) was a little lower than that within the species; the effective number of alleles (1.22) and average heterozygosity (0.14) were essentially similar to the estimates within species. The genetic diversity measures in the east region were a little lower than the west region. 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 east region versus 0.17 in the west region.

For loblolly pine sampled from the SSPSSS, the overall percentage of polymorphic loci was 46.68% (Table 3); the observed number of alleles was 1.47; the effective number of alleles was 1.19 and average heterozygosity was 0.12. The trees of Florida origin (source FL), sampled apart from the SSPSSS to represent the allopatric region, had a lower number of polymorphic loci (29.37%), a lower number of observed alleles (1.30) and effective alleles (1.17), and

Table 3 Genetic diversity of shortleaf pine and loblolly pine for all populations and regions based on 794 (shortleaf pine) and 647 (loblolly pine) AFLP loci

Population ID	P	na	ne	h
Shortleaf pine				
East				
451 and 401	44.96	1.45	1.22	0.13
435 and 487	52.14	1.52	1.24	0.15
419 and 421	37.91	1.38	1.20	0.12
461	36.27	1.36	1.19	0.11
Mean	42.82	1.43	1.21	0.13
East region	59.07	1.59	1.25	0.15
West				
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
423 and 475	47.61	1.48	1.22	0.13
Mean	44.59	1.45	1.23	0.14
West region	63.48	1.63	1.28	0.17
Mean (within populations)	43.71	1.44	1.22	0.14
Species	65.87	1.66	1.24	0.15
Loblolly pine				
East				
303	31.07	1.31	1.18	0.11
329 and 321	38.02	1.38	1.19	0.11
311 and 307	39.57	1.40	1.20	0.12
317 and 331	40.03	1.40	1.21	0.12
323	31.68	1.32	1.17	0.10
Mean	36.07	1.36	1.19	0.11
East region	46.06	1.46	1.21	0.13
West				
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)	33.82	1.34	1.18	0.11
Species	46.68	1.47	1.19	0.12
FL	29.37	1.30	1.17	0.10

P Percent polymorphic loci, *na* observed number of alleles, *ne* effective number of alleles, *h* average heterozygosity

lower average heterozygosity (0.10) when compared to the trees from the SSPSSS, but the Florida source was similar to the other samples at the physiographic population level. Within physiographic populations, the mean percentage of polymorphic loci (33.82%) was much lower than that within the species; all other measures including the observed number of alleles (1.34), the effective number of alleles (1.18), and average heterozygosity (0.11) were slightly lower than within species estimates. Genetic diversity measures in the east region were higher than those for the west region. 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.

The mean value of G_{st} of all loci among shortleaf pine populations was 0.1527. The mean G_{st} of all loci among loblolly pine populations was 0.1231. These G_{st} values suggest that in both species, a relatively small portion (15.27% in shortleaf pine; 12.31% in loblolly pine) of the observed genetic diversity exists among populations while a majority (84.73% in shortleaf pine; 87.69% in loblolly pine) of the genetic diversity observed was within populations. For both species, the unbiased measures of genetic identity were high and genetic distances were low for all pairwise comparisons. In shortleaf pine, the lowest genetic identity (0.9644) and highest genetic distance (0.0362) was between populations 433 and 481, and highest genetic identity (0.9893) and lowest genetic distance (0.0107) was between populations 435/487 and 475/423. In loblolly pine, the lowest genetic identity (0.9771) and highest genetic distance (0.0231) was between populations 303 and 327, and highest genetic identity (0.9935) and lowest genetic distance (0.0065) was between populations 331/317 and 311/307. In general, the high value of genetic identity and low value of genetic distance suggests that the genetic structure among populations was very similar.

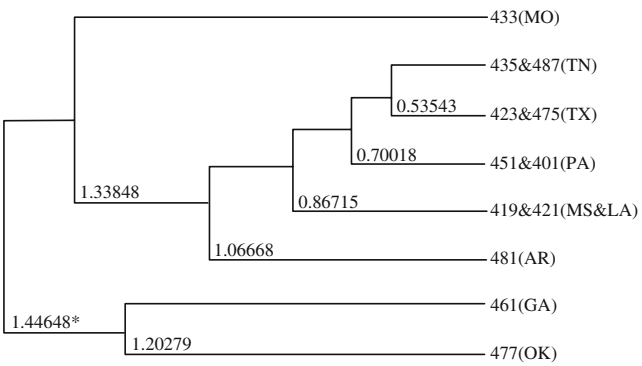

Fig. 3 Phenogram of shortleaf pine populations based on Nei's (1978) unbiased genetic distance (*asterisk*, the genetic distances among groups)

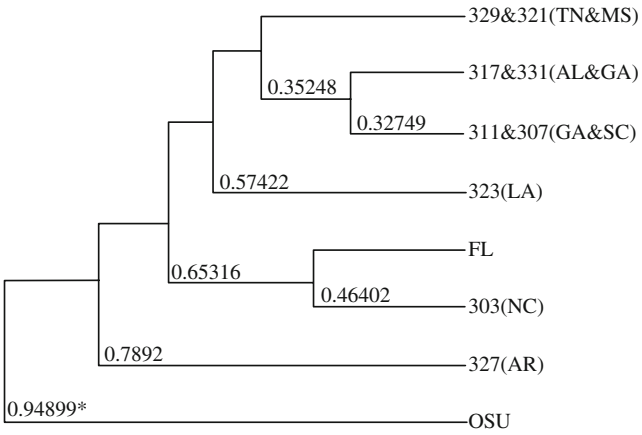

Fig. 4 Phenogram of loblolly pine populations based on Nei's (1978) unbiased genetic distance (*asterisk*, the genetic distances among groups)

Figures 3 and 4 are the phenograms of shortleaf pine and loblolly pine populations obtained by UPGMA based on Nei's (1978) unbiased genetic distance. As shown in Figs. 3 and 4, there is no obvious relationship between genetic distance and geographic distance for either species. The estimated correlation between genetic distance and geographic distance for shortleaf pine was not significant (r=0.28). There was a weak but significant correlation between genetic distance and geographic distance for loblolly pine (r=0.51).

Between the east and west regions, the mean genetic diversity estimate (Gst) for both shortleaf pine and loblolly pine was 0.02. This mean value suggests that only a small portion (2%) of the total genetic diversity found was between the two regions for both species; therefore, most of the genetic diversity (98%) occurs within both regions. The unbiased genetic identity of the two regions is 0.9945 in shortleaf pine and 0.9954 in loblolly pine, and the genetic distance between two regions is 0.0056 in shortleaf pine and 0.0046 in loblolly pine.

Note that only a few SSPSSS plantings remain, and of those that do, some seed sources have as few as four surviving trees. The possible affects of the small population sample sizes in this study are further addressed in the "Discussion," but in general, the small sample sizes for some sources, even after pooling into physiographic populations would suggest that some caution be given these results.

Discussion

To our knowledge, this study is the first to use AFLPs to explore genetic diversity in shortleaf pine and loblolly pine. When compared with previous studies based on

isoenzyme markers and microsatellite markers, our study results differ in four main ways as described in the following paragraphs.

First, AFLPs revealed a lower overall percentage of polymorphic loci. We found 65.87% polymorphic loci in shortleaf pine while Raja et al. (1997) found 87.2% and Edwards and Hamrick (1995) found 91% using isoenzyme markers. In loblolly pine, 46.68% polymorphic loci were found using AFLPs while Schmidtling et al. (1999), using isoenzyme markers, reported 64.9% polymorphic loci also sampling a SSPSSS planting. Sun et al. (1999) found similar differences when they compared genetic diversity estimates obtained by isozyme, random amplification of polymorphic DNA (RAPD), and microsatellite markers in *Elymus caninus*. RAPDs 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 noncoding regions across the entire genome, while isozyme markers reflect diversity of coding regions only. However, as dominant markers RAPDs and AFLPs miss all the remaining alleles at any locus, all other alleles will be scored as nulls. As dominant markers cannot distinguish heterozygote genotype (Aa) from homozygote genotype (AA), the percentage of polymorphic loci of this study might be underestimated. The percentage of polymorphic loci could also be underestimated due to small sample sizes in this study, since polymorphic loci with low frequency would have a low probability of being sampled. The small sample size may also lead to underestimates of mean percentage of polymorphic loci for populations.

Second, AFLPs revealed greater genetic diversity among populations. In shortleaf pine, genetic diversity reported here with AFLPs, by Raja et al. (1997) and by Edwards and Hamrick (1995), with isoenzymes was 0.153, 0.089 and 0.026, respectively. In loblolly pine, genetic diversity in this study and reported by Schmidtling et al. (1999) with isoenzymes was 0.123 and 0.066, respectively. The differences may be due to the marker loci sampled in the different studies. Raja et al. (1997), Edwards and Hamrick (1995), and Schmidtling et al. (1999) used isoenzyme loci, and as most isoenzymes reflect essential biological functions in *Pinus*, most mutations at these loci would be lost due to loss of function. The accumulation of variation through mutation would be limited for such loci. Accordingly, genetic variation estimates based on isoenzyme loci would be low among populations. In contrast, considerable variation may accumulate for neutral loci. In this study, the majority (97% in shortleaf pine; 98% in loblolly pine) of AFLP loci were selectively neutral, as shown by the Ewens–Watterson neutrality test. Mutations of selectively neutral loci are not harmful or useful and probably do not change the phenotypes of the individuals, so the neutral mutated loci have no selection pressure. In the evolutionary process, without selection pressure, such loci may accumulate multiple neutral mutations within populations. These neutral mutations would result in an increased measure of genetic variation among populations. The level of variation at selected loci may differ from that of neutral loci (Nei 1987). Genetic diversity may also be overestimated due to AFLP's dominant nature. All the data for positive monomorphic loci (genotypes like AA or Aa) can be collected but the data of all the negative monomorphic loci (genotype like aa) cannot be collected since no bands occur for these loci and there is no way to decide if the loci exist or not. Ignoring the negative monomorphic loci could overestimate genetic diversity. Again, the small sample size could also affect these estimates.

Third, more markers were used in our AFLP study. This study was based on 794 AFLPs in shortleaf pine and 647 AFLPs in loblolly pine, while only 39 isoenzyme markers were studied by Raja et al. (1997), 22 isoenzyme markers by Edwards and Hamrick (1995), and 18 isoenzyme loci by Schmidtling et al. (1999). Although AFLP markers are dominant and less informative than with codominant isoenzyme markers, the large number of AFLP markers can compensate for their dominant shortcomings. The number of markers used in different studies can affect genetic diversity estimates (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 and loblolly pine's genomes, while limited isozyme loci may only represent genetic diversity in limited coding regions of the genome. Estimates based on AFLPs, mostly neutral loci, may overestimate true diversity in gene expression.

Fourth, our AFLP study did not find a clear east–west difference in genetic diversity measures in shortleaf pine and loblolly pine, although we did note a weak but significant correlation between genetic distance and geographic distance for loblolly pine. 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 loblolly pine. In our study, the differences between east and west populations for loblolly pine and shortleaf pine in percentage of polymorphic loci, observed number of alleles, effective number of alleles, and Nei's gene diversity were small. Schmidtling et al. (1999) reported only a subtle east–west difference in allozyme frequencies of loblolly pine, as well.

Isoenzyme markers represent the variation of a highly restricted number of enzyme-related genes. Only a very small fraction of the variation present in a species is observed by isozyme studies. AFLPs or RAPDs reflect variation of both coding and noncoding regions of the whole genome. Microsatellite markers are located in noncoding repetitive regions and they reflect variation of the noncoding regions of the whole genome. Therefore, AFLPs (or RAPDs), isoenzyme markers, and microsatellite markers may reflect genetic diversity of different genome regions. Since coding sequences are under selection pressure to maintain function and noncoding regions have low or no selection pressure, the coding and noncoding sequences undergo different evolutionary processes. For example, repetitive sequences change by amplification and transposition more rapidly than single copy sequences (Sun et al. 1999). To date, low correlations between results based on isozyme markers and RAPD markers have been reported (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 microsatellite 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 isoenzymes or AFLPs and microsatellites 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. As seen in Table 3, genetic diversity measures within populations were lower than within species. Raja et al. (1997) and Edwards and Hamrick (1995) reported similar estimates.

In shortleaf pine, all the genetic diversity measures in the west region were slightly higher than those in the east region (Table 3). This same trend was observed by Raja et al. (1997). However, Edwardz and Hamrick's (1995) results were different. In their study, all the genetic diversity measures within the east region, except expected heterozygosity (He), were slightly higher than those in the west region. Since the differences between east and west regions are small, Edwardz and Hamrick's (1995) conclusion that the east and west regions have similar levels of genetic diversity seems reasonable.

All the studies, whether based on AFLPs, isoenzymes, or microsatellite markers, revealed some common results concerning the genetics of shortleaf pine and loblolly pine. These are: (1) high genetic diversity exists in shortleaf pine and loblolly pine, and most of the genetic diversity is within populations; (2) there is no correlation between population genetic distances and geographic distances in shortleaf pine; and (3) genetic differences between the east and the west regions are minimal (although the microsatellite study in loblolly pine did not agree on this point). Since AFLPs, isoenzyme, and microsatellite markers reflect variation in different parts of the genome, it may be best to combine them to get a comprehensive estimate of the genetic diversity for any organism.

Acknowledgements This study is supported by the USDA Forest Service, Southern Research Station, Cooperative Agreement SRS 05-CA-11330126-168 and by the Oklahoma State University Agricultural Experiment Station. We thank Larry Lott (Southern Institute of Forest Genetics) and personnel of the Oklahoma State University Kiamichi Forestry Research Station for assistance in locating and collecting needle samples.

References

Al-Rabab'ah MA, Williams CG (2002) Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. For Ecol Manag 163:263–271

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

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

Edwards MA, Hamrick JL (1995) Genetic variation in shortleaf pine, *Pinus echinata* Mill. (Pinaceae). For Genet 2:21–28

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

Hartl DL, Clark AG (1989) Principles of population genetics, 2nd edn. Sinauer, Sunderland

Heun M, Murphy JP, Phillips TD (1994) A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. Theor Appl Genet 87: 689–696

Huneycutt M, Askew GR (1989) Electrophoretic identification of loblolly pine shortleaf pine hybrids. Silvae Genet 38:95–96

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

Manly BFJ (1985) The statistics of natural selection. Chapman and Hall, London, pp 272–282

Messmer MM, Melchinger AE, Woodman WL, Lee EA, Lamkey KR (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, Lee M (1994) RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variation in oats. Theor Appl Genet 87:947–956

Muluvi GM, Sprent JI, Soranzo N, Provan J, Odee D, Folkard G, McNicol JW, 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, pp 187–192

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

Raja RG, Tauer CG, Wittwer RF, Huang Y (1998) Regeneration methods affect genetic variation and structure in shortleaf pine (*Pinus echinata* Mill.). For Genet 5:171–178

Remington DL, O'Malley DM (2000) Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. Genetics 155:337–348

Remington DL, Whetten RW, Liu BH, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in P*inus taeda*. Theor Appl Genet 98:1279–1292

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, 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 RC, Carroll E, LaFarge T (1999) Allozyme diversity of selected and natural loblolly pine populations. Silvae Genet 48:35–45

Schultz RP (1997) Loblolly pine: the ecology and culture of loblolly pine (*Pinus taeda* L). US Dept. of Agric., Ag. Handb. 713

Smith OS, Smith JSC, Bowen SL, Tenborg RA (1992) Numbers of RFLP probes necessary to show associations between lines. Maize Genet Coop News Lett 66:66

South DB, Buckner ER (2003) The decline of southern yellow pine timberland. J For 101:30–35

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

Tauer CG (1980) Twenty-year results of a shortleaf pine seed source study in Oklahoma. OSU Ag Exp Sta Bull B-752, pp 12

Wells OO (1973) Variation among shortleaf pines in a Mississippi seed source planting. USDA For Serv Res Note 3-SO-162, pp 8

Wells OO, Lambeth CC (1983) Loblolly pine province test in southern Arkansas 25th year results. South J Appl For 2:71–75

Wells OO, Wakley PC (1970) Variation in shortleaf pine from several geographic sources. For Sci 16:415–423

Wells OO, Switzer GL, Schmidtling RC (1991) Geographic variation in Mississippi loblolly pine and sweet gum. Silvae Genet 40:105–119

Xu S, Tauer CG, Nelson CD (2008) Natural hybridization within seed sources of shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.). Tree Genetics and Genomes (in press)

Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot 129:157