Bidirectional introgression between *Pinus taeda* and *Pinus echinata*: evidence from morphological and molecular data

Jiwang Chen, C.G. Tauer, Guihua Bai, Yinghua Huang, M.E. Payton, and A.G. Holley

Abstract: The frequency of mature hybrids, including post-F_1 individuals, between loblolly (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.), detectable with a codominant nuclear marker, was studied in a sympatric population from central Arkansas. The direction of introgression was also examined. The marker revealed 10 putative hybrids from the 80 trees sampled. PCR-RFLP analysis of their *rbcL* gene showed two of the putative hybrids (HL) share loblolly pine chloroplast DNA, and eight (HS) share the shortleaf pine chloroplast DNA. The two putative HL hybrids were morphologically similar to loblolly pine, and the eight putative HS hybrids were morphologically similar to shortleaf pine. Utilizing microsatellite data, Nei's measure of genetic identity showed the putative HL hybrids to be similar to loblolly pine, and the putative HS hybrids as being similar to shortleaf pine. An inferred tree of the individuals, using simple sequence repeat data and the neighbor-joining method, also suggested that some of the putative hybrids were not F_1 individuals. Principle component analysis of the morphological characters groups the HL trees with loblolly pine and the HS trees with shortleaf pine. These results suggest bidirectional introgression occurred within the study population, and some of the putative hybrids were likely derived from early-generation backcross(es) with either shortleaf or loblolly pine.

Résumé : Les auteurs ont évalué la fréquence des hybrides matures, incluant les individus introgressés, entre le pin à encens (*Pinus taeda* L.) et le pin à courtes feuilles (*Pinus echinata* Mill.) tels que détectés à l'aide d'un marqueur nucléaire codominant. L'évaluation a été réalisée au sein d'une population sympatrique du centre de l'Arkansas. La direction de l'introgression a aussi été étudiée. Le marqueur a permis de détecter 10 hybrides présumés à partir d'un échantillon de 80 arbres. L'analyse PCR-RFLP de leur gène *rbcL* a démontré que deux de ces hybrides présumés (HL) partageaient l'ADN chloroplastique du pin à encens, alors que les huit autres (HS) partageaient l'ADN chloroplastique du pin à courtes feuilles. Les deux hybrides présumés HL étaient morphologiquement similaires au pin à encens, alors que les huit hybrides présumés HS étaient morphologiquement similaires au pin à courtes feuilles. À l'aide de données de marqueurs microsatellites, la mesure d'identité génétique de Nei a démontré que les hybrides présumés HL étaient similaires au pin à encens, alors que les hybrides présumés HS étaient similaires au pin à courtes feuilles. Les auteurs ont estimé un dendrogramme des individus à l'aide des données séquences répétées en tandem et de la méthode de liens de voisinage de Saitou et Nei. Le dendrogramme a également suggéré que certains des hybrides présumés n'étaient pas des F_1. L'analyse en composantes principales des caractères morphologiques a regroupé les arbres HL avec le pin à encens et les arbres HS avec le pin à courtes feuilles. Ces résultats suggèrent qu'une introgression bidirectionnelle a eu lieu au sein de la population étudiée, et que certains des hybrides présumés sont probablement le résultat des premiers cycles d'une introgression autant avec le pin à courtes feuilles qu'avec le pin à encens.

[Traduit par la Rédaction]

Introduction

Loblolly pine (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) are economically important species in the southeast United States, and they have widely overlapping geographic ranges. Hybridization between the two species has interested tree breeders for a long time, because the hybrids are more resistant to both littleleaf disease (*Phytoph-*

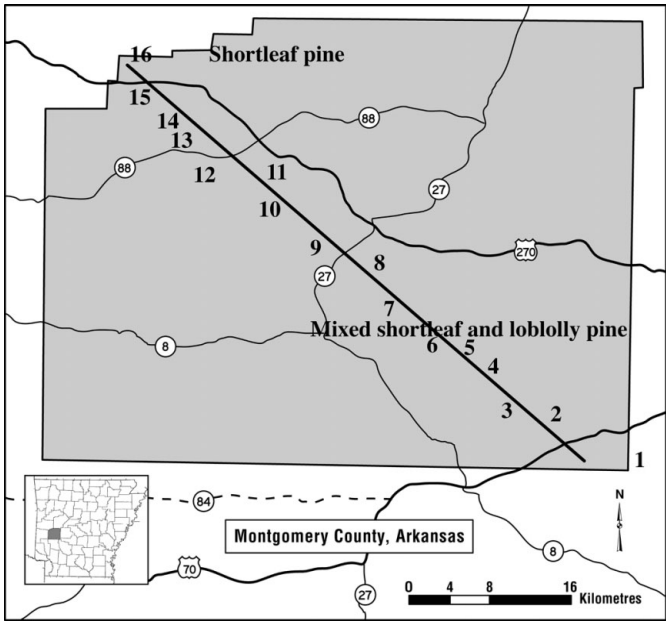
Received 26 January 2004. Accepted 12 August 2004. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 25 December 2004.

J. Chen, C.G. Tauer,[1] Y. Huang, and A.G. Holley.[2] Department of Forestry, 08C Ag Hall, Oklahoma State University, Stillwater, OK 74078, USA.
G. Bai.[3] Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK 74078, USA.
M.E. Payton. Department of Statistics, 301 Math Sciences, Oklahoma State University, Stillwater, OK 74078, USA.

[1]Corresponding author (e-mail: ctauer@okstate.edu).
[2]Present address: School of Forestry, Louisiana Tech University, P.O. Box 10138, Ruston, LA 10138, USA.
[3]Present address: USDA, Agricultural Research Station, 4008 Throckmorton Hall, Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA.

 doi: 10.1139/X04-134

Fig. 1. Sample transect across Montgomery County, Arkansas, showing approximate stand location by stand number. Mount Ida is at the intersection of highways 27 and 270.



thora cinnamomi) and fusiform rust (*Cronartium fusiforme*) and may also grow faster (Schultz 1997). Thus, it is important to be able to reliably identify the two pine species and their hybrids. Morphologically, the two pine species are different. The needles of loblolly pine are 6–9 in. (1 in. = 25.4 mm) long, usually with three yellow-green needles per fascicle, while shortleaf pine needles are 3–5 in. long, with two or three dark yellow-green slender and flexible needles per fascicle. Loblolly pine also has larger cones than shortleaf, as well as other differences. However, these characters offer limited help when the genotypes of the parents and their hybrids are compounded by environmental factors such as disease or drought stress, resulting in a wide range of phenotypic variability. This confusion is aggravated by the fact that subsequent backcrossing of the hybrids to either of the parent species results in morphological characters exhibiting progressively greater similarity to the backcrossed parent species. The limitations of morphological characters resulted in the identification of the allozyme marker *IDH* (isocitrate dehydrogenase) to identify hybrids (Huneycutt and Askew 1989). This allozyme marker shows a single band for each species, but the loblolly band migrates faster than the shortleaf, consequently, the hybrid shows two bands, one from each parent. The high frequency of *IDH* variation reported in a natural shortleaf pine population outside the natural range of loblolly pine (Raja et al. 1997) suggests either an unexpected high level of hybridization between the two species or that *IDH* is an unreliable marker. These results enticed us to look for new markers to confirm the identity of putative hybrids of the two pine species and the reliability of the *IDH* marker.

The maternity and paternity of natural crosses between the two species are still unclear. Artificial hybridization has been most successful using shortleaf pine as the female parent. Our controlled crosses with shortleaf pine as pollen parent resulted in no seed, while hybridization with loblolly pine as the pollen parent resulted in many viable seeds (12 crosses with reciprocals, unpublished data). Our results agree with the early work of Critchfield (1962). Snyder and Squillace (1966) did report a few seed from loblolly pine female parents, and more recently Richard Bryant (personal communication, 2003) gave three independent examples of success using loblolly pine as the female parent in crosses with shortleaf pine, including raising the seedling offspring and characterizing them as hybrid (intermediate) in their morphology. However, the preponderance of evidence from artificial crosses, and the natural pollen receptivity times of the female strobili of these two species (Schultz 1997), suggest that the naturally occurring hybrids probably have shortleaf pine maternity. Edwards et al. (1997) reported that shortleaf pine sired the putative hybrids they identified in two natural shortleaf–loblolly pine sympatric populations; however, they had no proof of the female parents of these putative hybrids, and morphologically, these putative hybrids could be early generation backcrosses.

Our previous study using the *IDH* allozyme marker (Raja et al. 1997) showed that 16% of the trees in a stand near Mount Ida, Arkansas, were putative hybrids. To further explore the nature and extent of hybridization in native populations, a more extensive study sampling a larger portion of the shortleaf–loblolly pine sympatric population, including the Mount Ida area, was conducted. The southeast stands of this population are mixed loblolly and shortleaf pine, while the northwest stands are only shortleaf pine. We combined a codominant nuclear DNA marker, a paternally inherited chloroplast DNA marker, and simple sequence repeat (SSR) markers to determine the frequency of occurrence of mature putative hybrid trees and the direction of introgression between the two species. The morphology of these trees was also examined. In addition, our markers enabled us to assess the value of the *IDH* marker in the identification of putative hybrids between the two species. It became clear to us early in this study that most of the putative hybrids we identified were not F_1 individuals, thus we are using the term hybrid in the more general sense, to include F_1s, early generation intercrosses, and introgressants.

Materials and methods

Plant materials

The sample population was defined as the pine stands of Montgomery County, Arkansas. Five trees in each of 16 stands were sampled on a southeast to northwest transect across the county (Fig. 1). Stands were located at approximately equal distances across the transect (average distance was 4.0 km, stands 9 and 10 were the farthest apart at 6.5 km, and stands 10 and 11 were the closest at 1.5 km). The southeast stands (Nos. 1–8) are mixed loblolly and shortleaf pine, while the northwest stands (Nos. 9–16) are only shortleaf pine. Only dominant or codominant trees, separated by at least 60 m, were sampled within each stand. Raja et al. (1997) showed that about 16% of the individuals in a shortleaf pine population near Mount Ida are putative hybrids. Their results were based on the heterozygosity of the *IDH* allozyme marker, reported by Huneycutt and Askew

(1989) to be indicative of a hybrid between shortleaf and loblolly pine. Mount Ida is located about 3 km east of stand nine, and a few kilometres north of any known stands of loblolly pine.

In addition to this study population, parents of one controlled cross, shortleaf pine (Z15, seed parent) × loblolly pine (No. 631, pollen parent), and 20 artificially produced F_1 hybrids from this cross were used to confirm the utility of our codominant DNA marker developed from the nuclear ribosomal internal transcribed spacer region.

Morphological analysis

All samples were measured for four traits previously determined to distinguish shortleaf pine and loblolly pine (Keng and Little 1961; Abbott 1974; Snyder and Hamaker 1978; Cotton et al. 1975). These traits are the number of needles per fascicle, needle length, fascicle sheath length, and cone length. Thirty-four needles were measured for the three needle traits. Previous work by Cotton et al. (1975) showed this to be a more than adequate number. The number of cones measured was dependent on the number we were able to collect, which varied from a few to over 45. The mean values and the standard deviations of these traits for the 80 trees were calculated. Since these 80 trees were ultimately assigned into four types of unequal size, a pseudo-*t* test was used to test for differences among them. Principal component analysis of the four traits for all 80 trees was performed using the SAS (SAS Institute Inc. 1990) PRINCOMP procedure. We did not include morphological data on the artificial hybrids for comparison, since these trees were grown in southern Mississippi, and variation within both species across geographic distances can be considerable (compare our data with those of Dorman 1976 and Edwards et al. 1997), and even year to year variation can be large (compare our data to those of Abbott 1974).

DNA extraction

Needles from the parent trees, artificial hybrids (F_1), and the 80 trees from the natural population were stored at –80 °C. Total DNA was extracted from needles using the CTAB protocol (Doyle and Doyle 1988).

Polymerase chain reaction - restriction fragment length polymorphism (PCR-RFLP) analysis of the nuclear ribosomal internal transcribed spacer (ITS)1 region

The primer 26S-25R (5′-TATGCTTAAACTCAGCGGGT-3′) and the modified ITS-5 primer (5′-GGGAGGAGAAGT-CGTAACAAGG-3′) (Quijada et al. 1997; Nickrent et al. 1994; White et al. 1990) were used to amplify the nuclear ITS region. Conditions for PCR amplification were the following: 10 mmol/L Tris–HCl (pH 9.0 at 25 °C), 50 mmol/L KCl, and 0.1% Triton X-100, 1.8 mmol/L $MgCl_2$, 0.16 mmol/L dNTP mix, 1.6 μmol/L of each primer, 1 unit DNA *Taq* polymerase (Promega company, ?), 1% BSA, 5% DMSO, and 20 ng of DNA in a final reaction volume of 25 μL. Cycling conditions were as follows: 70 °C for 3 min; two cycles of 94 °C for 2 min, 55 °C for 40 s, and 72 °C for 3 min. This was followed by 35 cycles of 94 °C for 30 s, 50 °C for 50 s, and 72 °C for 3 min; then 72 °C for 8 min. Restriction digests of the amplified ITS region DNA were accomplished with 500 ng (12 μL) of unpurified PCR product and the addition of the recommended enzyme buffer. Sixteen endonucleases that recognize 4-bp or 6-bp sites (*Alu*I, *Hae*III, *Hin*fI, *Mse*I, *Rsa*I, *Taq*I, *Pst*I, *Msp*I, *Pst*I, *Sac*I, *Sma*I, *Eco*RI, *Hin*dIII, *Bam*HI, *Apa*I, *Xho*I) were used separately to digest the amplified ITS fragment. Agarose gel electrophoresis (2.0%) and ethidum bromide staining were used to reveal PCR-RFLP bands.

Isocitrate dehydrogenase allozyme marker

The *IDH* allozyme marker reported by Huneycutt and Askew (1989) to identify putative shortleaf–loblolly pine hybrids is an assay that relies on megagametophyte tissue from the seed. Seed was collected from all of the 80 trees in the population sample that had mature unopened cones. Unfortunately, it was a poor seed year, and not all trees had cones with seed. The *IDH* marker was assayed (as described by Raja et al. 1997) for the 25 individuals in the population for which seed was available.

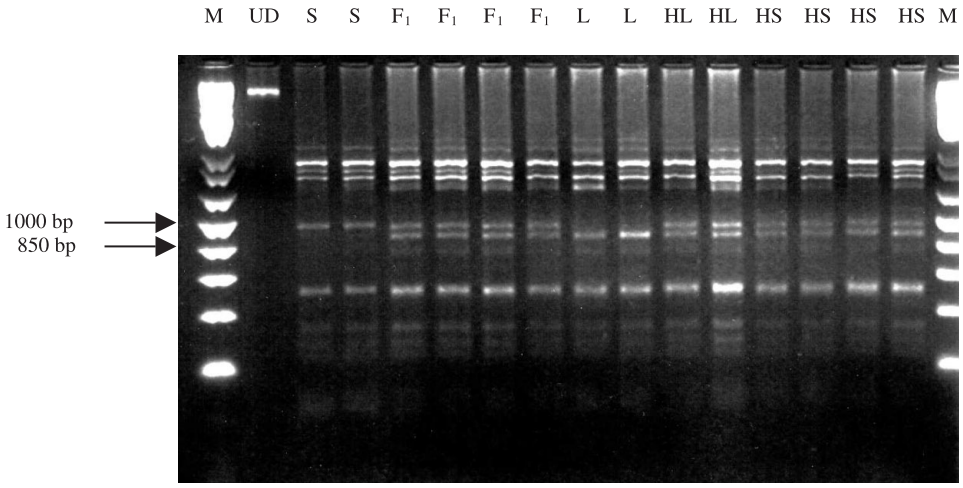
PCR-RFLP analysis of chloroplast DNA *rbcL* gene

The primer A1010 (5′-GTAGTAGGTAAACTTGAAGG-3′) and the *rbcL* 3′ primer (5′-ATTGGTAGAACGAAAGTCACT-GGA-3′) (Edwards et al. 1997) were used to amplify the chloroplast *rbcL* region (from bases 961 to 1428). Conditions for PCR amplification were similar to the amplification of the nuclear DNA ITS region, but the extension time was 2 min. Approximately 12 μL unpurified PCR product were digested with *Hin*dIII. The separation and staining of PCR-RFLP bands were the same as above.

Microsatellite amplification

Eleven genomic microsatellite markers reported polymorphic in loblolly pine were selected from Web sites formerly maintained by Dr. C. Williams while at Texas A&M University. The sequences, sizes, and polymorphisms are given at the above website. Selected primer pairs were PtTX 2001, PtTX 2006, PtTX 2033, PtTX 2119, PtTX 3016, PtTX 3024, PtTX 3025, PtTX 3035, PtTX 3117, PtTX 3125, and PtTX3104. Each 15-μL PCR reaction was composed of 35 ng of template DNA, 10 mmol/L Tris–HCl (pH 9.0 at 25 °C), 50 mmol/L KCl, and 0.1% Triton X-100, 0.16 mmol/L dNTP mix, 1.6 μmol/L of each forward and reverse primer, 0.9 unit DNA *Taq* polymerase (Promega), and 1% DMSO. The microsatellites were amplified with the following touch-down programs: two cycles of 95 °C for 2 min, T_{mi} °C for 40 s, and 72 °C for 1 min; 35 cycles of 94 °C for 30 s, T_{mv} °C for 40 s, and 72 °C for 1 min; 20 cycles at 93 °C for 30 s, 50 °C for 40 s, and 72 °C for 50 s; then 72 °C for 5 min. T_{mi} is the initial annealing temperature; different SSR primers have different initial temperatures, which are available at http://silva.tamu.edu/genetics/AlleleSizes.rtf. T_{mv} decreased 0.5 °C from the initial annealing temperature for every cycle until reaching 50 °C. The concentrations of $MgCl_2$ for different primers are also available from the above website. PCR products were first checked by 2% agarose gel electrophoresis to ensure successful amplification, then run in a 6% polyacrylamide denaturing gel to separate the microsatellite bands. A silver staining method (from Dr. Bai Guihua's laboratory, Oklahoma State University) was used to visualize microsatellite bands. The allele size was

Fig. 2. PCR-RFLP analysis of the nuclear ribosomal DNA internal transcribed spacer 1 for shortleaf pine, loblolly pine, their artificial hybrids (F_1), and putative hybrids. Lanes are as follows: M, 1-kb plus DNA marker; UD, undigested PCR product; S, shortleaf pine (Z15); L, loblolly pine (No. 631); F_1, artificial hybrids between Z15 (seed parent) × No. 631 (pollen parent); HL, putative hybrids with loblolly pine chlorolpast DNA (cpDNA); HS, putative hybrids with shortleaf pine cpDNA.



determined using a 100-bp DNA ladder (Life Technologies™, St. Paul, Minnesota).

Microsatellite data scoring and data analysis

Microsatellite loci were selected based on their molecular sizes as given at http://silva.tamu.edu/genetics/AlleleSizes.rtf. Allele frequencies were determined by direct manual count. The frequency of each allele per locus, the observed heterozygosity (H_o), the expected heterozygosity (H_e), and deviations from Hardy–Weinberg equilibrium were computed by the program POPGENE (Yeh and Boyle 1997).

Based on the chloroplast marker data and the PCR–RFLP analysis of the ribosomal DNA ITS marker, the 80 individuals in the study population were placed in four groups: pure shortleaf pine, pure loblolly pine, hybrids with the loblolly pine chloroplast marker, and hybrids with the shortleaf pine chloroplast marker. All SSR data were then combined as four groups within one population, and the genetic identity was calculated among the four groups (Nei 1978). SSR data on one individual in the shortleaf group was not available. The relationship among 79 of the 80 individual trees was also examined using the SSR data. We inferred a tree relating all the individuals, from genetic distances estimated according to the method of Mountain and Cavalli-Sforza (1997), utilizing the neighbor-joining method of Saitou and Nei (1987).

Results

A codominant internal transcribed spacer DNA marker

From the 16 restriction enzymes used to digest the PCR-amplifed nuclear DNA internal transcribed spacer region, only *Msp*I produced polymorphic patterns among the parental species (Fig. 2). The artificial hybrids (Fig. 2) showed codominant restriction site patterns concordant with patterns of the parental species.

This diagnostic nuclear ribosomal DNA marker was used to screen the natural population materials we selected. Of the 80 samples in the natural population, 10 putative hybrids were identified, having both bands, one from each parent.

Chloroplast DNA inheritance

The *Hin*dIII-digested PCR amplified *rbcL* fragment produced polymorphic patterns, which can be used to distinguish shortleaf pine and loblolly pine (Fig. 3). All of the artificial hybrids (F_1) show the pollen parent (loblolly pine) pattern, which confirms that chloroplast DNA is paternally inherited in the cross of shortleaf pine (seed parent) × loblolly pine (pollen parent), as reported by Neale and Sederoff (1989). This diagnostic marker was also used to screen the population samples. The 10 putative hybrids identified from the natural population show two different patterns: two putative hybrids show the loblolly pine pattern and are designated HL, while the other putative hybrids show the shortleaf pine pattern and are designated HS.

Morphological data

Mean values of the morphological data for the 80 trees as placed into the four groups are shown in Table 1. The pseudo-*t* tests comparing each possible pairing of the groups are given in Table 2. The morphological data clearly distinguish loblolly pine from shortleaf pine. Loblolly pine has longer needles, cones, and fascicle sheaths, and essentially 3 needles per fascicle, while shortleaf has an average of 2.3. These data also distinguish the pure species from the putative hybrids that are morphologically similar to the other parent, but do not allow identification of those putative hybrids morphologically similar to themselves. These results are strongly supported by principle component analysis (PCA) of the four traits for all 80 trees (Fig. 4). PCA divided loblolly and shortleaf into two distinct groups using the first two principal components. The two putative HL hybrids were in the middle of the loblolly group. The putative HS hybrids were scattered across the shortleaf group. Since the putative hybrids identified from the natural population appear morphologically either similar to shortleaf pine or loblolly pine, they could be easily misclassified as pure species without

Fig. 3. PCR-RFLP of the partial *rbcL* chloroplast gene from shortleaf pine, loblolly pine, their artificial and natural hybrids. Lanes are as follows: M, 100-bp DNA marker; UD, undigested PCR product; S, shortleaf pine (Z15); L, loblolly pine (No. 631); F_1, artificial hybrids between Z15 (seed parent) × No. 631 (pollen parent); HL, putative hybrids with loblolly pine chlorolpast DNA (cpDNA); HS, putative hybrids with shortleaf pine cpDNA.

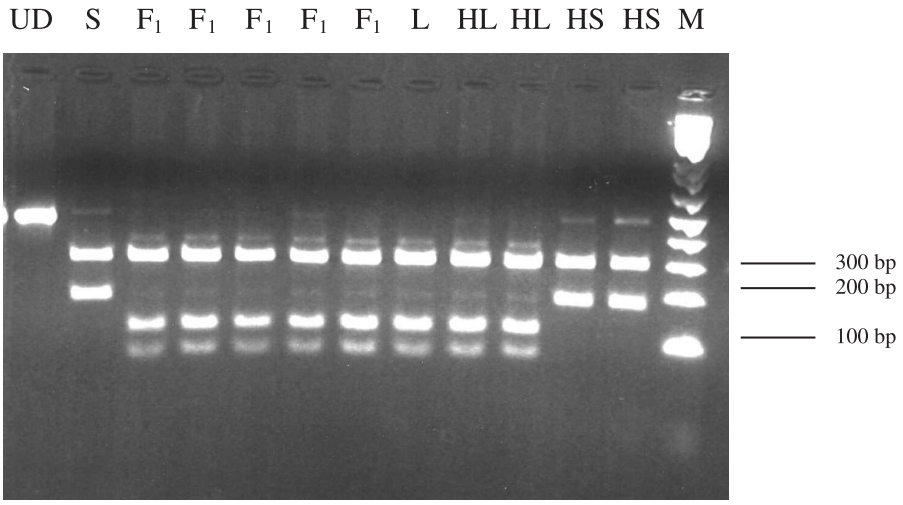


Table 1. Mean values for morphological characters of 80 samples from a natural mixed population of shortleaf and loblolly pine, grouped by molecular marker differences.

	Group[a] mean value (standard deviation)			
Characters	L(16)[b]	HL(2)[b]	HS(8)[b]	S(54)[b]
Needle number/fascicle	3.0(0.17)	3.0(0)	2.4(0.30)	2.3(0.08)
Needle length	18.0(5.23)	19.5(0.28)	10.7(0.90)	10.2(6.64)
Cone length	6.9(4.53)	6.2(0.20)	4.8(0.60)	4.3(0.60)
Fascicle sheath length	1.9(0)	1.9(0.18)	1.4(0.25)	1.3(0.50)

[a]Group abbreviations: L, loblolly pine; HL, putative hybrids with loblolly pine chlorolpast DNA (cpDNA); HS, putative hybrids with shortleaf pine cpDNA; S, shortleaf pine.
[b]Numbers in parentheses following groups are sample sizes.

Table 2. Pseudo-*t* test *p* values for all paired comparisons of the four molecular marker groups from the 80 samples from a natural mixed population of shortleaf and loblolly pine.

	Group[a] pairs					
Characters	S–HS	S–HL	S–L	HS–HL	HS–L	HL–L
Needle number/fascicle	0.469	0.000	0.000	0.001	0.001	0.664
Needle length	0.525	0.000	0.003	0.000	0.000	0.007
Cone length	0.037	0.179	0.000	0.226	0.000	0.417
Fascicle sheath length	0.153	0.131	0.000	0.100	0.001	0.942

Note: If $p < 0.05$, the two groups are significantly different for this character. See Table 1 for group abbreviations.

utilizing molecular marker data. These morphological data suggest that the putative hybrids are introgressants.

Genetic diversity and relationships among the four groups

Observed (H_o) and expected (H_e) heterozygosity are given in Table 3. Genetic diversity levels for the four groups were high in this natural population. A total of 104 SSR alleles were found for 79 of the 80 individuals. The 16 loblolly pine trees share 78 alleles, the 53 shortleaf pine have 95 alleles, the two putative hybrids morphologically similar to loblolly pine have 41 alleles, and the eight shortleaf-like putative hybrids have 66 alleles. These allele numbers are not a good indicator of genetic variation, because the sample sizes of the four groups are different.

Genetic identity between loblolly pine and the loblolly-like putative hybrids (HL) was 0.9370, and between shortleaf pine and the shortleaf-like putative hybrids (HS), it was 0.9742 (Table 4). The close identity of loblolly-like putative hybrids (HL) to loblolly pine, and of the shortleaf-like putative hybrids (HS) to shortleaf pine suggests introgressants rather than F_1s.

The neighbor-joining tree of 79 of the 80 individuals (Fig. 5) divided the individuals into mostly loblolly on one end of the tree, and the rest of the tree was mostly shortleaf. There

Fig. 4. Plot of the first and second principal components (Prin1 and Prin2, respectively) from a principal component analysis of four morphological characters of 80 trees from a sympatric population of shortleaf and loblolly pine in Montgomery County, Arkansas. 1, shortleaf pine; 2, putative hybrids with shortleaf pine chlorolpast DNA (cpDNA); 3, putative hybrids with loblolly pine cpDNA; 4, loblolly pine.

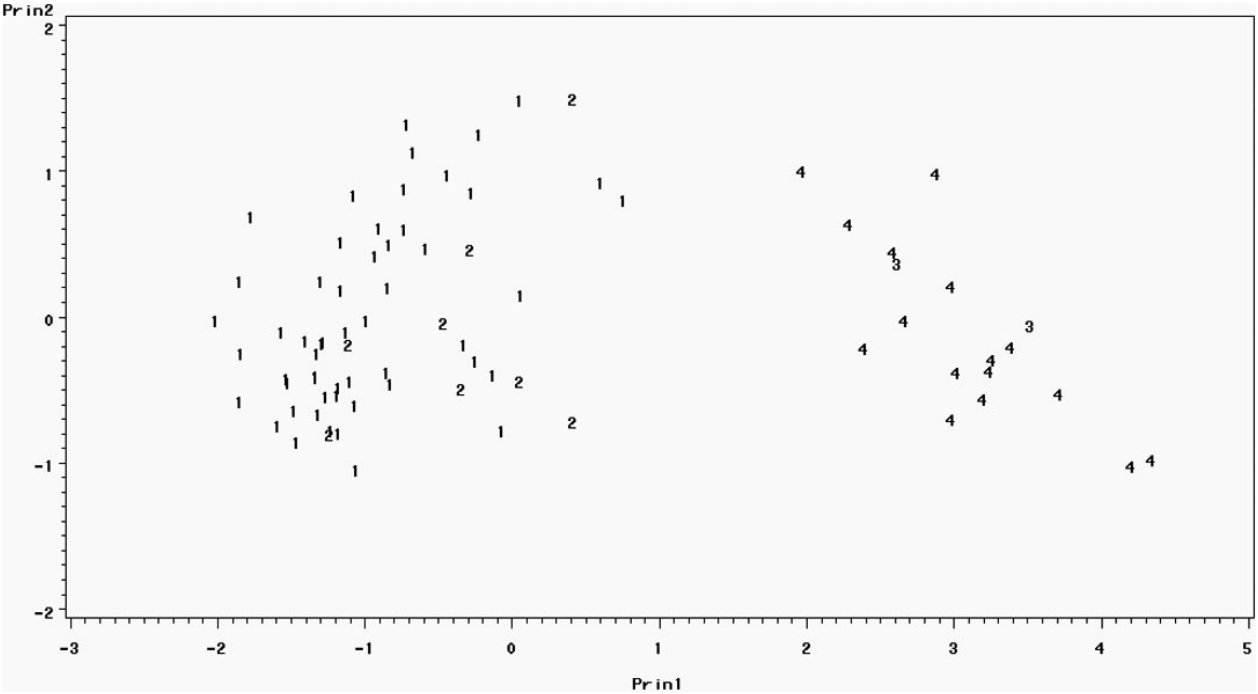


Table 3. Observed (H_o) and expected (H_e) heterozygosity estimates based on 11 microsatellite loci assayed for 79 of 80 trees from a natural mixed population of shortleaf and loblolly pine.

Group[a]	Sample size	No. of alleles	H_o	H_e[b]
L	16	78	0.5631(0.2463)	0.5606(0.2100)
HL	2	41	0.8611(0.2304)	0.6852(0.1512)
HS	8	66	0.7583(0.2114)	0.6550(0.1949)
S	53	95	0.4605(0.2023)	0.5540(0.2210)

Note: Data for H_o and H_e are mean (standard deviation).
[a]See Table 1 for group abbreviations.
[b]Values were computed using Levene (1949), which is the same as Nei's (1978) unbiased heterozygosity.

Table 4. Nei's (1978) unbiased measures of genetic identity among the four molecular marker based groups within the natural mixed population of shortleaf and loblolly pine.

Group[a]	L	HL	HS	S
L	—	0.9370	0.8474	0.8808
HL		—	0.8734	0.8279
HS			—	0.9742
S				—

[a]See Table 1 for group abbreviations.

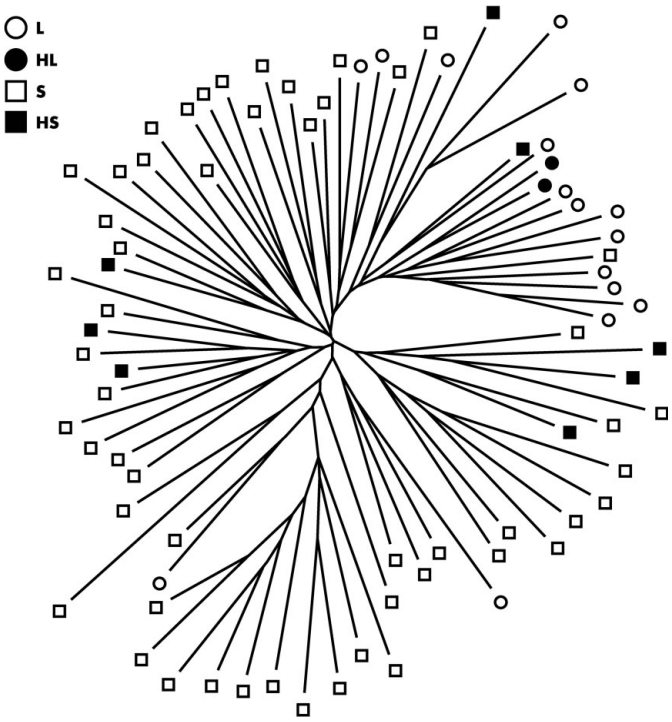
was some mixing of the two species at the interface of the groups. The two putative HL hybrids fell near the center of the loblolly grouping, and were not intermediate. The putative HS hybrids were scattered across the shortleaf grouping, from the interface of the two species to the center of the shortleaf group. The two putative HS hybrids that did plot to the interface of the two species were not intermediate morphologically (Fig. 4). Separate from the interface of the two species, there were two loblolly in the shortleaf group, and one shortleaf in the loblolly group.

Discussion

Ten putative hybrid individuals were found among the 80 trees from the stands sampled in the transect across Montgomery County, Arkansas, utilizing the ITS nuclear marker. There was no apparent pattern to their occurrence relative to the natural distribution of the two species across the sample transect. One putative hybrid individual was found in each of stands 1, 5, 6, 12, 13, and 15, and two were found in stands 4 and 9. Both of the putative HL hybrids were found in stand four. The presence of putative hybrids in the shortleaf populations of considerable distance (up to 35 km) from any loblolly pine supports our conclusion that at least some of these are backcross generation hybrids.

The codominant DNA marker found from PCR-RFLP analysis of the nuclear ribosomal internal transcribed spacer region was reported for the first time to distinguish loblolly and shortleaf pine. The nuclear ribosomal DNA ITS region in the genus *Pinus*, more than 3000 bp (Liston et al. 1999), includes ITS1, 5.8s rDNA, and ITS2. Our previous data (Chen et al. 2002) show that shortleaf and loblolly pine share the same nucleotide sequences of 5.8s rDNA and ITS2. The difference between the two species in the rDNA ITS region is located in the ITS1 region, which was found by PCR-RFLP analysis.

Fig. 5. Unrooted neighbor-joining tree from 11 microsatellite loci showing genetic similarities among 79 trees from a sympatric population of shortleaf and loblolly pine in Montgomery County, Arkansas. S, shortleaf pine; L, loblolly pine; HL, putative hybrids with loblolly pine chlorolpast DNA (cpDNA); HS, putative hybrids with shortleaf pine cpDNA.



This ITS1 PCR-RFLP marker identified 10 putative hybrids from the population sample of 80 trees. The two putative hybrids with loblolly pine chloroplast DNA were morphologically similar to loblolly pine (HL), suggesting they may be derived from early generation backcrosses with loblolly pine. The other putative hybrids have shortleaf pine chloroplast DNA and are morphologically similar to shortleaf pine (HS), suggesting they may be derived from early generation backcrosses with shortleaf pine. We attempted to develop mitochondrial DNA markers to confirm that the putative hybrids are from backcrosses with shortleaf pine or loblolly pine, but the three mitochondrial DNA (mtDNA) markers we examined, which show high variation in most of the genus *Pinus*, either show the same patterns between shortleaf and loblolly pine or show variability within species (Chen 2001).

We also used the ITS1 marker to evaluate 12 putative hybrids previously found in a natural shortleaf population. These trees were identified as putative hybrids (Raja et al. 1997) because they were heterozygous at the *IDH* locus. Our nuclear ITS1 marker confirmed them as hybrids. In the present study, 7 of the 10 putative hybrids identified using the nuclear marker were heterozygous for the *IDH* locus, but one showed the loblolly band only, and one showed the shortleaf band. For one putative hybrid, we had no data at the *IDH* locus. Similarly, one individual from the shortleaf portion of the population (based on the ITS1 region) was heterozygous at the *IDH* locus (we had IDH data on only 9 of the 10 hybrids, and 16 individuals from the rest of the population, because of the poor seed crop the year we sampled). Since our data indicated that the putative hybrids identified with nuclear markers are generally morphologically similar to one parent or the other, and not intermediate as expected for an F_1, these putative hybrids are probably of backcross or intercross origin. As such, one would expect segregation at the *IDH* locus, resulting in some of these early generation hybrids being homozygous at the *IDH* locus. If all 10 hybrids we identified with the ITS1 marker were first-generation backcrosses, we would expect 50% of them to be homozygous for the *IDH* marker. However, only two of nine (22%) were homozygous. This would suggest that some of the 10 hybrids may be F_1 individuals, but the sample size is too small to discount chance alone (not testable) to explain the number we observed, and we do not know if selection among the hybrids occurs.

We can conclude that the *IDH* locus appears to be reliable in identifying F_1 hybrids of these two pine species, but that reliability does not extend to identifying all later generation hybrids. The same is true for the nuclear marker we developed. These markers used in combination should allow the identification of most of the naturally occurring hybrids between loblolly pine and shortleaf pine, but additional nuclear markers would be helpful, as demonstrated by Bobola et al. (1996).

Both the observed and expected heterozygosities for the HL and HS putative hybrids were larger than for shortleaf pine or loblolly pine, which is indicative of the hybrid nature of these individuals. That H_o exceeds H_e for the putative hybrids may simply reflect our small sample size, but may also reflect mating among related individuals within each species. Both H_o and H_e for loblolly pine was greater than for shortleaf pine, which is consistent with the report of Edwards et al. (1997), utilizing allozyme data, but this result may also reflect the fact that the microsatellites were originally selected because they were polymorphic in loblolly pine.

Schultz (1997) reported that F_1 hybrids are intermediate for the needle characters we measured, while hybrids backcrossing with loblolly pine are morphologically more similar to loblolly pine, and hybrids backcrossing with shortleaf pine are morphologically more similar to shortleaf pine. Our identified putative hybrids are generally similar to either shortleaf pine or loblolly pine, rather than intermediate. PCA analysis of the 80 individuals for the four traits shows that the two species are distinct from each other, but that the putative hybrids we found are not intermediate (Fig. 4). The putative hybrids generally fall into the middle of either the loblolly or shortleaf species group. Combining the morphological and molecular data of the putative hybrids found in this natural population sample, we conclude that some of the HL putative hybrids may be from backcrosses of early generation hybrids with loblolly pine, while some of the HS putative hybrids may be from backcrosses of early generation hybrids with shortleaf pine.

The individual tree neighbor-joining analysis, utilizing microsatellite data (Fig. 5), shows that in the population sampled the two species are generally separate. There is some overlap, and the overlap region included two putative HS hybrids. It is possible these putative HS hybrids are F_1 trees, but the remaining eight hybrids are either in the center of the loblolly

(both HL trees) or shortleaf populations (the remaining six HS trees), and not intermediate, as would be expected for F_1 trees. This observation, along with the morphology data and the *IDH* data suggest that some of the hybrids we identified are early generation backcross or intercross trees. There were three trees that did not follow the general pattern observed; two loblolly in the middle of the shortleaf and one shortleaf in the loblolly grouping. These may simply be extreme variants, but they could well be hybrids we could not detect with our marker. One of the loblolly, for example, was from stand 11, about 10 km from any known stands of loblolly pine.

At the species level, the morphological characters clearly separate the two species, while SSR data were not as effective. However, the morphological data fail to identify the hybrids and (or) introgressants, while the molecular markers are quite effective in hybrid identification. Intuitively, this might be expected, and has been reported recently by Aldrich et al. (2003) and Perron and Bousquet (1997). Perron and Bousquet (1997) effectively demonstrate how morphological analyses can lead to an underestimate of the level of hybridization in populations. Our molecular marker results, showing about 14% of the population to be hybrid, agrees with the 16% estimate of Raja et al. (1997) based on isozyme data and suggests that the conclusion of Cotton et al. (1975) utilizing morphological data alone (i.e., that introgression was not widespread in shortleaf–loblolly pine stands in east Texas) may reflect such an underestimate. However, since we have sampled a single population, further study is suggested.

Concordance of morphological data with molecular marker data might also be influenced by the type of marker used. Yeh and Arnett (1986) and Wheeler and Guries (1987), using isozymes, and Bennuah et al. (2004), using STS markers based on expressed sequence polymorphisms, reported good concordance between their markers and morphological data. In contrast, our SSR data, as well as that of Aldrich et al. (2003) and the RAPD data of Ye et al. (2002) showed less concordance. Do expression-based markers better relate to phenotype?

Our results have immediate implications for loblolly pine breeding programs across the south. Since F_1 hybrids do tend to grow fast, it is likely some of these trees have been selected into improvement programs. However, the progeny from these trees would be expected to show wide segregation, and it is probable that such trees would be rouged. Simple screening with the *IDH* marker and (or) the ITS1 marker would allow identification of most such trees and eliminate the expense of progeny testing them. As an example, the Oklahoma loblolly pine tree improvement program's original selections included five putative hybrids among the 22 selections (based on our markers, unpublished data, and Abbott 1974). Perhaps a more important question is what is the effect of conversion of many natural shortleaf stands across the south to loblolly pine on the remaining shortleaf pine population? Will the increase in loblolly pine pollen combined with a reduction in shortleaf pollen in the regional pollen cloud lead to a significant increase in hybrid production, at the expense of shortleaf pine genetic integrity? Work reported by Wheeler and Guries (1987) suggests this can happen naturally, and Fady et al. (2003) demonstrate how man's activities can lead to hybridization events.

At this point the extent of hybridization between these two species across their sympatric range is not clear. Wagner et al. (1991) found hybrids between *Pinus banksiana* and *P. contorta* tend to cluster in specific microenvironments. Since our sample is from the interface of the sympatric range with the allopatric range of shortleaf, there could be an interface environment that favors the hybrid; however, Perron and Bousquet (1997) tend to discount this explanation for tree species. Edwards et al. (1997) reported only 7% hybrids in two populations in Georgia in the center of the sympatric region, and that all these hybrids were juvenile. It is quite possible the level of hybridization we observed in Montgomery County, Arkansas, is a localized event.

The frequency of hybrids in the natural loblolly and shortleaf pine populations should also be of concern to anyone studying the molecular genetics of these species. Proper interpretation of results from such studies requires that the scientists be positive that they are working with pure species, if that is their intent. The implications of these data in regard to genetic diversity and past and future evolutionary processes are intriguing but require further, more extensive studies.

Acknowledgements

The authors thank the USDA Forest Service, Southern Institute of Forest Genetics, and Dr. Bruce Bongarten, Warnell School of Forest Resources, The University of Georgia, for the supply of needles from parents of the controlled cross and 20 F_1 hybrids from this cross. Thanks go to Dr. Peiguo Guo at Oklahoma State University and Dr. Zhou Yi at Texas A&M University for technical assistance. This study was supported by the Oklahoma State University Agricultural Experiment Station.

References

Abbott, J.E. 1974. Introgressive hybridization between shortleaf and loblolly pine in southeast Oklahoma. M.S. thesis, Oklahoma State University, Stillwater, Okla.

Aldrich, P.R., Parker, G.R., Michler, C.H., and Romero-Severson, J. 2003. Whole-tree silvic identifications and the microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest. Can. J. For. Res. **33**: 2228–2237

Bennuah, S.Y., Wang, T., and Aitken, S.N. 2004. Genetic analysis of the *Picea sitchensis* × *glauca* introgression zone in British Columbia. For. Ecol. Manage. **97**: 65–77.

Bobola, M.S., Eckert, R.T., Klein, A.S., Stapelfeldt, K., Smith, D.E., and Guenette, D. 1996. Using nuclear and organelle DNA markers to discriminate among *Picea rubens*, *Picea mariana*, and their hybrids. Can. J. For. Res. **26**: 433–443

Chen, J.W. 2001. Introgression between shortleaf pine (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.). Ph.D thesis, Oklahoma State University, Stillwater, Okla.

Chen, J.W., Tauer, C.G., and Huang, Y. 2002. Nucleotide sequences of the internal transcribed spacers and 5.8S region of nuclear ribosomal DNA in *Pinus taeda* L., and *Pinus echinata* Mill. DNA Sequencing J. Sequencing Mapping, **13**: 129–131.

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

Critchfield, W.B. 1962. Hybridization of the southern pines in California. Proceedings of a Forest Genetics Workshop, Macon,

Georgia, 25–27 October 1962. Publ. No. 22 of the Southern Forest Tree Improvement Committee. pp. 40–48.

Dorman, K.W. 1976. The genetics and breeding of the southern pines. U.S. Dep. Agric. Agric. Handb. 471.

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

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

Fady, B., Lefèvre, F., Reynaud, M., Vendramin, G.G., Dagher-Kharrat, M.B., Anzidei, M., Pastorelli, Savouré, R.A., and Bariteau, M. 2003. Gene flow among different taxonomic units: evidence from nuclear and cytoplasmic markers in *Cedrus* plantation forests. Theor. Appl. Genet. **107**: 1132–1138.

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

Keng, H., and Little, E.L., Jr. 1961. Needle characteristics of hybrid pines. Silvae Genet. **10**: 131–146.

Levene, H. 1949. On a mating problem arising in genetics. Ann. Math. Stat. **20**: 91–94.

Liston, A., Robinson, W.A., Piñero, D., and Alvarez-Buylla, E.R. 1999. Phylogenetics of *Pinus* (Pinaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. Mol. Phylogenet. Evol. **11**: 95–109.

Mountain, J.L., and Cavalli-Sforza, L.L. 1997. Multilocus genotypes, a tree of individuals, and human evolutionary history. Am. J. Hum. Genet. **61**: 705–718.

Neale, D.B., and Sederoff, R.R. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. Theor. Appl. Genet. **77**: 212–216.

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

Nickrent, D.L., Schuette, K., and Starr, E. 1994. A molecular phylogeny of *Arceuthodium* (*Viscaceae*) based on nuclear ribosomal DNA internal transcribed spacer sequences. Am. J. Bot. **81**: 114.

Perron, M., and Bousquet, J. 1997. Natural hybridization between black spruce and red spruce. Mol. Ecol. **6**: 725–734.

Quijada, A., Liston, A., Robinson, W., and Alvarez-Buylla, E. 1997. The ribosomal ITS region as a marker to detect hybridization in pines. Mol. Ecol. **6**: 995–996.

Raja, R.G., Tauer, C.G., Wittwer, R., and Huang, Y. 1997. An isoenzymatic study of the genetic structure in natural populations of shortleaf pine, *Pinus echinata* Mill. Can. J. For. Res. **27**: 740–749.

SAS Institute Inc. 1990. SAS procedures guide, version 6 [computer manual]. 3rd ed. SAS Institute Inc., Cary, N.C.

Saitou, N., and Nei, M. 1978. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**: 406–425.

Schultz, R.P. 1997. Loblolly pine: the ecology and culture of loblolly pine (*Pinus taeda* L.). U.S. Dep. Agric. Agric. Handb. 713. pp. 7–17 and 7–18.

Snyder, E.B., and J.M. Hamaker. 1978. Needle characteristics of hybrids of some species of southern pine. Silvae Genet. **27**: 184–188.

Snyder, E.B., and Squillace, A.E. 1966. Cone and seed yields from controlled breeding of southern pines. USDA For. Serv. Res. Pap. SO-22.

Wagner, D.B., Sun, Z.X., Govindaraju, D.R., and Dancik, B.P. 1991. Spatial patterns of chloroplast DNA and cone morphology variation within populations of a *Pinus banksiana-Pinus contorta* sympatric region. Am. Nat. **138**: 156–170.

Wheeler, N.C., and Guries, R.P. 1987. A quantitative measure of introgression between lodgepole and jack pines. Can. J. Bot. **65**: 1876–1885.

White, T.J., Bruns, T., Lee, S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* PCR protocols: a guide to methods and applications. *Edited by* M. Innis, D. Gelfeand, J. Sninsky, and T. White. Academic Press, San Diego, Calif. pp. 315–322.

Ye, T.Z., Yang, R.C., and Yeh, F.C. 2002. Population structure of a lodgepole pine (*Pinus contorta*) and jack pine (*P. banksiana*) complex as revealed by random amplified polymorphic DNA. Genome, **45**: 530–540.

Yeh, F.C., and Arnott, J.T. 1986. Electrophoretic and morphological differentiation of *Picea sitchensis*, *Picea glauca*, and their hybrids. Can. J. For. Res. **16**: 791–798.

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