Theor Appl Genet (2001) 103:930–943

C. Lambeth · B.-C. Lee · D. O'Malley · N. Wheeler

Polymix breeding with parental analysis of progeny: an alternative to full-sib breeding and testing

Received: 27 October 2000 / Accepted: 5 March 2001

Abstract Implementation of most traditional full-sib crossing and testing designs is logistically difficult, costly, and genetic gain is constrained by resource availability. An alternative to full-sib crossing and testing can ameliorate all these constraints. The alternative solution is to apply a pollen mix (PMX) of many male parents rather than a single pollen for each cross. PMX breeding is easy to implement, ensures good estimates of breeding values of the parent being pollinated, and provides for increased genetic gain opportunity because of the significantly increased number of effective parental combinations tested. However, PMX breeding has found limited use because inbreeding and pedigree control is lost. The development of relatively inexpensive molecular markers allows for the paternity analysis of progeny, thus allowing full pedigree control. We call this system PMX breeding with parental analysis (PMX/WPA). The feasibility of PMX/WPA was evaluated in a loblolly pine (*Pinus taeda*) third-generation population of 45 select individuals, among which there was considerable relatedness. All parental trees were genotyped at seven chloroplast and three nuclear microsatellite loci. Unique fingerprints were obtained for all 45 individuals, but unambiguous paternal determinations for progeny from a complete pollen mix would not be possible due largely to relatedness in the breeding population. The inclusion of more markers and/or the creation of polmixes and breeding groups that avoid relatedness would resolve this problem. Three PMX/WPA scenarios are described and compared with conventional full-sib breeding and testing systems: (1) paternity analysis of only the forward (next generation) selection candidates in each generation, (2) paternal genotyping of all progeny test individuals in PMX crosses and using those data to construct the effective full-sib crosses for statistical analysis, and (3) like #2 except mixing seed of the several PMX crosses for ease of greenhouse rearing and progeny testing and then doing maternal and paternal analysis.

Keywords Polymix crossing · Mating design · Genetic fingerprinting · Genetic markers · Genetic gain

Communicated by H.C. Becker

C. Lambeth (✉)
Weyerhaeuser Company, 810 Whittington Avenue, Hot Springs, Arkansas AR 71902, USA
e-mail: clem.lambeth@weyerhaeuser.com
Tel.: (501)-624-8510, Fax: (501)-624-8419

B.-C. Lee
North Carolina State University, College of Forest Resources, 840 Main Campus Drive, Raleigh, North Carolina, NC 27695-727

D. O'Malley
North Carolina State University, College of Forest Resources, 840 Main Campus Drive, Raleigh, North Carolina, NC 27695–727, USA

N. Wheeler
Weyerhaeuser Company, 505 N. Pearl, Centralia, Washington, WA 98531, USA

Introduction

For decades breeders have faced the dilemma of how to mate a select group of individuals (parents) for the next generation of genetic improvement (step 3 Table 1). The chosen mating design is usually a compromise as no single design has been found that best meets all objectives (Namkoong 1979). Some common objectives of a mating design are to:

(1) Accurately estimate the breeding value of the parents being bred for the purpose of recommending which to include in a production scenario (sometimes called backward selection) (step 7a Table 1), and for determining which crosses among high breeding-value parents will provide the best potential for the selection of individuals for the next generation of breeding.
(2) Produce a set of crosses that will allow a high genetic-gain potential from the selection of the best individuals within the best crosses for the next generation of breeding, often called forward selection (step 8 Table 1).

Table 1 Process flow for a typical genetic improvement program for forest trees

	Select Breeding Group (Parents)	Step 1
	↓	
	Graft For Breeding	Step 2
	↓	
	Crossing Among Parents	Step 3
	↓	
	Seed Collection	Step 4
	↓	
	Sow in a Nursery	Step 5
	↓	
	Field Plant Crosses (Progeny Test)	Step 6
	↓	
	Measurement and Analysis	Step 7
	↓	
	Advanced Generation Selection Among and Within Best Crosses	Step 8
	↓	
	(Repeat Cycle)	
Step 7a. Evaluate Parents (B.V.) for Commercial Production Purposes Such as an Open-Pollinated Seed Orchard		

(3) Provide a full pedigree where both male and female are known for new selections in order to control the level of inbreeding.
(4) Generate sound estimates of genetic parameters: heritability, general and specific combining ability, genetic correlations among traits and genotype-environment interaction.
(5) Do all of the above efficiently in order to keep breeding and testing costs at a manageable level.

The most obvious mating design to achieve all but objective 5 (above) would be to cross every parent with every other parent (full diallel), but that alternative is almost always rejected in operational breeding programs because it is impractical to make and test so many crosses. The number of total selections may be of the order of a few hundred and the number of possible crosses in the thousands, a number that may be prohibitively expensive to breed and test. The latter is especially true for tree species which, due to their large individual tree size, require special equipment to do breeding and a large area for testing individual plants. Consequently, forest geneticists have sought mating and testing designs that achieve their primary objectives as efficiently as possible. All designs currently in use for tree improvement programs represent a compromise because no design heretofore meets the above five objectives very effectively. The purpose of this paper is to present a novel solution to some of the problems of existing breeding and testing designs by using molecular-marker technologies as an integral component of improvement programs.

Currently used mating designs

Mating designs that have been commonly compared in the literature are *open pollination* versus controlled pollination and, for the latter, *full-sib* crossing versus *polymix* (a mixture of pollen from several males) crossing (Bridgwater 1992). The most-common method of quantitative comparisons for these mating designs in the literature is to hold the number of parents bred, and number of progeny tested from crosses among them, as constants in each generation and then determine how well each design estimates genetic parameters and their genetic gain potential. A synopsis follows.

Open pollination

Open pollination (OP) in forestry is usually assumed to take place in a progeny test where only a handful of new selections are found among several hundred or a few thousand other individuals. OP breeding gives a good estimation of parental breeding values and heritability estimates (Bridgwater 1992; Burdon and Shelbourne 1971; Cotterill 1986a, b; White 1996).

The genetic-gain potential from using the OP families for the next generation of progeny testing is weak since the selected parents are mated with a nonselected population from the previous generation (Burdon and Shelbourne 1971; Cotterill 1986a, b). This and other disadvantages, such as lack of full pedigree control and the inability to estimate specific combining ability, have resulted in limited use of OP breeding in forestry except in programs where limited resources dictate a simple and low-cost approach.

Full-sib crossing

There are many full-sib mating patterns that have been discussed and analyzed in the literature including full-diallel, cousin mating, single-pair, nested, factorial, disconnected half diallels and circular systems (Cotterill 1986a, b; Pederson 1972; Burdon and van Buijtenen 1990; van Buijtenen and Burdon 1990; Bridgwater 1992; Huber et al. 1992). Of these, the last three have been

commonly used in forest tree improvement because they provide: (1) a reasonably good parameter estimation (general and specific combining ability, heritability, breeding value of parents), (2) a foundation for reasonably good genetic gain, and (3) full pedigree control. In these designs each parent is usually mated with four to six of the other parents in the breeding population.

Full-sib system weaknesses include. (1) the amount of work involved in breeding and testing, (2) suboptimal estimation of breeding value, and (3) a weak foundation for selection for the next generation of breeding. The two latter limitations stem from the relatively few crosses for each parent. For a given population size for the progeny test of the crosses, it is generally better to have many crosses per parent for a precise parameter estimation and high genetic gain potential from forward selection rather than few crosses with many individuals per cross (Pederson 1972; White 1996). This is especially true when selection emphasis is heavily on family versus individuals within families, i.e., under low heritability situations. Few crosses per parent can result in inaccurate breeding value estimation, especially if there is significant specific combining ability (SCA) in the breeding population (van Buijtenen and Burdon 1990; Burdon and van Buijtenen 1990). Genetic gain for the next generation is limited due to the fact that the best parents among those being bred may not have been frequently mated to the other best parents due to chance. Nontheless, few crosses and many individuals per cross are usually opted for due to the high cost of making and testing the crosses. Although the full pedigree is known, the limited number of crosses usually means that there is a tendency to select from within a few good crosses, which can often mean a common parentage among those crosses and the danger of inbreeding building up in future generations. Alternatively, inbreeding concerns may force selection from more unrelated but mediocre families, thus limiting genetic gain.

Polymix crossing

Polymix (PMX) crossing is done by mixing pollen from several males and applying the pollen to isolated females. One of the big advantages is the simplicity of crossing and the subsequent testing of relatively few crosses. Polymix crossing is sometimes considered for parental breeding value estimation only (Burdon and van Buijtenen 1990; White 1996), in which case the source of pollen may or may not be made up of the parents in the breeding group. It is sometimes given credence as a complete breeding system for breeding value (B.V.) estimation and as a foundation for the selection of individuals for advanced generation breeding (Shelbourne 1969; Burdon and Shelbourne 1971; Cotterill 1986a, b; Kerr 1998). In the latter case it is preferable to use pollen from the select group in order to keep the genetic gain potential high. PMX crossing provides excellent estimation of breeding value and GCA but not usually of SCA (Bridgwater 1992; Huber et al. 1992). Genetic gain potential for forward selection is good but not as high as that offered by the commonly used full-sib systems mentioned above. The gain potential from PMX crossing is 70% to over 90% of the potential for full-sib systems (Cotterill 1986a, b; Kerr 1998; Shelbourne 1969; van Buijtenen and Burdon 1990), depending on the magnitude of the genetic parameters, with increasing heritability and SCA favoring PMX crossing (van Buijtenen and Burdon 1990).

PMX mating systems result in less genetic gain from advanced generation selections versus full-sib systems, primarily because the paternal parent's general combining ability (GCA) is unknown (Burdon and Shelbourne 1971). When individual selections are made within the best full-sib crosses for the next generation of breeding they are commonly chosen on the basis of an index that includes the following information:

$$\text{Ind. B.V.}=\text{fem. gca}+\text{male gca}+h^2_w\ (\text{ind. dev. within full-sib cross}) \quad (1)$$

where: h^2_w=within-cross heritability.
For the PMX cross the individual value is estimated as follows:

$$\text{Ind. B.V.}=\text{fem. gca}+h^2_w\ (\text{ind. dev. within PMX cross}). \quad (2)$$

Selection within PMX crosses does not have the benefit of knowing the male GCA, but the heritability for, and variation within, a PMX cross is greater than that for full-sib crosses because a PMX cross is actually made up of the variation among many full-sib crosses as well as the variation within those crosses. On balance, lack of information on male GCA results in a reduced gain efficiency of PMX breeding and testing. Lack of full pedigree control and the possibility that a few males may be represented among selections for the next breeding generation, resulting in possible inbreeding depression, is another drawback that has limited use of PMX crossing as a stand-alone system (Bridgwater 1992).

Combined systems

Many organizations use a combination of mating designs to take advantage of the strengths of each and to offset, to some degree, the limitations of any one design (van Buijtenen and Burdon 1990). The major tree-improvement cooperatives for pine species in the southeastern U.S. use a complementary system consisting of PMX crossing for parental breeding values and full-sib crossing from which selections are made for the next generation of breeding (Lowe and van Buijtenen 1986; McKeand and Bridgwater 1992; White et al. 1993). Full-sib crossing is typically a circular or disconnected half-diallel design. While offering many advantages, these systems are costly in terms of time and resources. Furthermore, they do not overcome the above-mentioned limitations of full-sib crossing systems for the purposes of forward selection gain and inbreeding control.

A new alternative: PMX crossing with parental analysis of progeny (PMX/WPA)

Advances in molecular genetics, and specifically the development of low-cost genetic markers, open up new breeding and testing strategies that were once unthinkable. We propose the use of parental analysis to overcome the heretofore primary limitations of PMX breeding and testing, which are lack of male pedigree control and lack of male GCA information for advanced generation selection. Three scenarios will be discussed in detail. For each of these scenarios, select parents are assumed to be mated with a pollen mixture of a majority of parents from the same select group and the resulting PMX crosses planted in typical field tests.

Scenario 1 – partial population paternity analysis

A group of provisional candidates are identified for advanced generation breeding (step 8 in Table 1) using female GCA and individual performance as selection criteria. Tissue samples from all of the parents involved in the PMX crosses, as well as the advanced generation candidates, are genotyped and paternity is assigned to the provisional candidates. The list of provisional candidates is then narrowed down to a final breeding list whose fathers have an acceptable GCA and an acceptable level of relatedness based on a criteria such as group merit which weighs genetic gain potential against an estimate of inbreeding depression (Rosvall and Andersson 1999). A limitation of this system is that the initial list of provisional candidates may not include the highest value individuals in the test since male GCA is known only for the provisional candidate list. Nonetheless, the final screening based on paternity analysis should allow genetic gain to be higher than had been possible with PMX crossing systems of the past. This approach is feasible if the paternal analysis cost per individual is somewhat expensive, perhaps in the order of $30–$50/individual for some tree-improvement programs.

Scenario 2 – full population paternity analysis

Genotyping of all original parents and all individuals destined for field testing is carried out by collecting tissue samples either in the nursery (step 5) or in the field test (step 6). Sample DNA is characterized and the father of each tree is identified before data analysis (step 7). By determing the father of all individuals before analysis of the measurement data, the information on both male and female parentage can be used in advanced generation selection as is done in any other full-sib mating system. Besides the usual advantages of PMX breeding, this system has several additional advantages over other full-sib systems such as the commonly used circular and disconnected half-diallels. As long as fertilization is mostly at random, many effective crosses are generated with PMX mating, resulting in a greater genetic gain potential for advanced generation selection and more flexibility for relatedness control. Also, genetic typing of all individuals helps reduce the number of identity (pedigree) errors that are carried into future generations, a serious concern in any genetic-improvement program (Adams et al. 1988). This scenario is possible when genetic fingerprinting is inexpensive enough to genetically type all individuals, perhaps in the order of $3–5/individual in some tree-improvement programs.

Scenario 3 – full population maternity and paternity analysis (throw away the ID tags)

After the seed is processed (step 4 Table 1) equal (or known) quantities of seed from the PMX crosses are completely mixed before sowing in the nursery. The nursery crop and the field test are established without PMX cross-identification tags. Consequently, the field test is a completely randomized design. Tissue samples are taken from all field-test individuals at the time of measurement and from the original parents for genetic typing, and maternal and paternal determinations are made in the laboratory for all individuals in the progeny test. This system should be considered when both maternity and paternity determinations are possible and genetic typing is inexpensive, perhaps in the order of $3–$5/tree for some tree-improvement programs, and when the cost of maintaining the identity of crosses in the nursery and/or field test is difficult and/or expensive.

The state of pedigree analysis in tree improvement programs

Molecular-marker applications in tree improvement and tree breeding have increased notably in recent years, including such uses as clonal fingerprinting for identification and taxonomic analysis, the creation of genetic maps, the dissection of complex traits and quantitative trait loci detection, and marker-aided selection (Groover et al. 1994; O'Malley and McKeand 1994; Cervera et al., 1996; O'Malley et al. 1996; Staub et al. 1996; Dinus and Tuskan 1997). However, marker applications focused on pedigree issues in forestry are largely restricted to seed orchard studies, including estimating pollen contamination, selfing, male reproductive success, and supplemental mass-pollination success (Wheeler and Jech 1992; Stoehr et al. 1998; Walter et al. 1998). Most orchard applications rely on paternity analysis to infer pollen donors. Paternity analyses have also been applied in natural-stand progeny arrays (Dow and Ashley 1996; Ziegenhagen et al. 1998; Isagi et al. 2000).

Recently, rapid advancements in the development of highly informative markers have led to increased use of pedigree analyses for parentage assignment and the estimation of relatedness among progeny arrays for a num-

ber of species including humans (Hohoff and Brinkmann 1999), *Drosophila* (Dewoody et al. 2000), salmon (Norris et al. 2000), shrubs (Krauss 1999) and trees (Isagi et al. 2000). While the salmon studies have an applied objective of tracking pedigrees to maintain diversity and reduce inbreeding in mating lines, to our knowledge there are no reports in the literature of pedigree analysis for improving breeding and or testing efficiencies, or increasing genetic gain potential as discussed above.

The discrimination requirements of a pedigree analysis increase notably in the three PMX/WPA scenarios referred to previously. The ability to accurately assign parentage and estimate relatedness using molecular markers depends on many factors:

1. Size of the potential parent pool or parental combinations.
2. Amount of relatedness among parents in the breeding group.
3. Prior knowledge of parental genotypes (one parent known vs no parental genotypes).
4. Type of marker chosen (information content, number of loci, repeatability, inheritance, i.e., Mendelian, paternal, maternal).
5. Population gene frequencies.
6. Statistical models (exclusion vs likelihood or paternity or maternity assignments).
7. Genotype error and mutation rates.

Additionally, the cost of genotyping will significantly influence how the technique is applied.

In this study we test the diagostic power of three highly variable nuclear microsatellites and seven chloroplast microsatellites to: (1) uniquely fingerprint a select breeding population of 45 individuals of mixed pedigree, (2) to assign paternity where male parent/offspring relationships exist, and (3) to estimate known levels of relatedness. The chloroplast SSRs are non-recombining and, like the Y-chromosome markers in humans (Hohoff and Brinkmann 1999), are paternally inherited in conifers (Neale et al. 1986; Wagner 1992) making paternity assignment straightforward.

Microsatellites (SSRs, STRs) have proven very useful for pedigree analyses (Blouin et al. 1996; Norris et al. 2000), and are generally considered the most-powerful genetic marker available today (Goldstein and Pollock, 1997). Microsatellites are easily scored, reproducible, highly variable, and codominant. Most individuals in a population are heterozygous at SSR loci and few share the same genotype, making them ideal for fingerprinting and determination of parentage (O'Malley and Whetten 1997; Elsik et al. 2000).

Objectives

The objectives of this paper are:

(1) To introduce the PMX/WPA approach to plant breeding and testing.

(2) To present prototype genetic marker results and discuss the current and future genetic marker potential for the kinds of genetic typing and parental determinations called for in the three PMX/WPA scenarios.

(3) To discuss the overall applicability of the PMX/WPA breeding and testing system and its advantages and disadvantages.

A subsequent manuscript will present the results of computer simulations comparing the genetic gain potential of the three PMX/WPA scenarios with some currently used breeding and testing designs. Discussions and examples typically center on the authors' experience with tree species but the PMX/WPA scenarios may be applicable to other organisms wherein a plurality of crosses can be created through fertilization with a mix of male gametes.

Materials and methods

Fingerprinting, paternity and pedigree analysis

Plant materials

For data reported here, needle samples were collected from two groups: (1) 78 first-generation loblolly pine plus-tree selections were used in our analysis to determine population gene frequencies for seven cpSSR and three nuclear SSR loci. These trees were selected in natural stands from North Carolina, South Carolina and Georgia (Atlantic Coastal Plain provenance, ACP); (2) 45 elite 1st, 2nd and 3rd generation selections from Weyerhaeuser's breeding program were genotyped at the same suite of loci. Considerable relatedness occurred among these 45 trees, including half-sibs, full-sibs, parent-offspring, and grand-parent-offspring relationships Twenty nine of these trees were paternally unrelated.

DNA extraction

DNA was prepared from needle tissue following the CTAB procedure of Doyle and Doyle (1990) or using the DNeasy 96 DNA extraction kit (Qiagen, Valencia, Calif.). Details of extraction methods are presented in Appendix 1.

PCR-amplification

DNA fragments were PCR-amplified from loblolly pine genomic DNA samples using the 20 pairs of DNA primers designed by Vendramin et al. (1996) for amplification of chloroplast microsatellites based on the complete DNA sequence of *Pinus thunbergii* (Wakasugi et al. 1994). Complete details of the PCR methods are noted in Appendix 1.

Electrophoresis

DNA fragments were separated on an automated DNA sequencing apparatus (LI-COR 4000) using 0.25-mm gels (8% Long Ranger, FMC BioProducts), 25-cm gel plates, and electrophoresis buffer 1× TBE. Gel images were analyzed using RFLP-Scan software (Scanalytics, Billerica, Mass.). The software was used as an aid to facilitate scoring the size of the DNA fragments amplified from chloroplast DNA templates. The bands were analyzed for multiple peaks and "stutter" bands using quantitative peak profiles. Microsatellite fragments were classified by size using estimates of band migration rates relative to molecular-weight markers and size standards. Further details of electrophoretic methods are noted in Appendix 1.

Analysis

Exclusion probability for haplotypes was calculated using haplotype frequencies derived from the 78 first-generation selections in group 1 and the 29 paternally unrelated individuals in the elite second group. For haplotype j, the exclusion probability was calculated as

$1-p_j$,

where pj is the frequency of the jth haplotype. The mean exclusion probability was calculated as

$$\sum_{1}^{N_h} p_j \times (1-p_j)$$

where N_h is the total number of haplotypes.

Known genotypes and putative pedigrees for the elite population were submitted to the likelihood model KINSHIP (Goodnight and Queller 1999). Estimates of relatedness (probability of sharing an allele by descent) for all possible pairs of individuals (990 combinations) and for all potential hypothetical relationships were determined.

Table 2 Primer sequences for chloroplast and nuclear microsatellites (* denotes the IRD 41-labeled nucleotide)

Chloroplast microsatellites	Sequence
PT1254f	*CAATTGGAATGAGAACAGATAGG
PT1254r	TGCGTTGCACTTCGTTATAG
PT9383f	*AGAATAAACTGACGTAGATGCCA
PT9383r	AATTTTCAATTCCTTTCTTTCTCC
PT26081f	*CCCGTATCCAGATATACTTCCA
PT26081r	TGGTTTGATTCATTCGTTCAT
PT30204f	*TCATAGCGGAAGATCCTCTTT
PT30204r	CGGATTGATCCTAACCATACC
PT71936f	*TTCATTGGAAATACACTAGCCC
PT71936r	AAAACCGTACATGAGATTCCC
PT109567f	*TATTATCGAACAACGAGAATAATCC
PT109567r	TCACTGTCACTCTACAAAACCG
PT110048f	*TAAGGGGACTAGAGCAGGCTA
PT110048r	TTCGATATTGAACCTTGGACA
Nuclear microsatellites	
3011f	CACGACGTTGTAAAACGACAATTTGGG TGTATTTTTCTTAGA
3011r	AAAAGTTGAAGGAGTTGGTGATC
3034f	CACGACGTTGTAAAACGACTCAAAATG CAAAAGACG
3034r	ATTAGGACTGGGGATGAT
6C12Ff	CACGACGTTGTAAAACGACCCAGACAA CCCAAATGAAGG
6C12Fr	GCCAGTGCAGACACAAACAA

Results and Discussion

Fingerprinting, paternity and pedigree analysis

This study investigated the utility of molecular markers for tree-breeding applications, including a novel approach to breeding and testing that relies on paternity or full-pedigree analysis.

(1) cpSSRs: all of the 20 microsatellite primer pairs designed by Vendramin et al. (1996) to amplify simple sequence repeat regions in the chloroplast genome of *Pinus thunbergii* amplified similar size DNA fragments in *Pinus taeda.* Of these, seven showed a simple pattern of inheritance that conformed with the expectation of uni-parental inheritance from the pollen parent, and were considered sufficiently variable to be used in this study (Table 2). Variation generally consisted of a series of single base pair differences. Unlike nuclear SSRs, the number of alleles per locus in these conifer cpSSRs was rather limited, varying from two to four. Fifty different 7-locus haplotypes were detected in the natural and elite populations surveyed in this study. Fourteen additional haplotypes were detected in open-pollinated progeny from an ACP seed orchard (data not shown). Twenty of the 64 haplotypes were observed in the elite study population (Table 3). Of these, ten occurred in only one selection, five were shared between two or more selections related by descent, four were shared between related and unrelated selections, and one was shared only among unrelated individuals (Table 3). Haplotype frequencies among the 29 paternally unrelated individuals in the elite group ranged from 0.034 to 0.103, and among unrelated individuals in the combined groups from 0.009 to 0.112. Combined, the five most-common haplotypes in these two groups occurred 42.4% of the time.

The mean exclusion probability for the multi-locus haplotype in the elite group was 92.1% (the probability of excluding a non-sire), but this calculation assumes random mating among unrelated individuals. The 45 individuals in the elite group represent selections from three separate generations, many of which are related as grandparent-offspring, parent – offspring, half-sibs or full-sibs, both maternal and paternal (Table 3).

In Weyerhaeuser's pine breeding-program databank, parentage gender is not maintained (because reciprocal effects are virtually non-existent). As a consequence, chloroplast haplotypes were used to infer paternity/maternity for all levels of relationships (Table 3). In all cases but one, paternity could be verified with double exclusion (i.e., at least two cpSSR loci excluded one parent as being the father). In the exceptional case, accessions #5 and #6, parents of accession #4, shared the same haplotype and it was impossible to determine maternity or paternity for either. For the 45 individuals studied, there were 16 known cases of paternal haplotype-identity by descent.

(2) Nuclear microsatellites: one bi (3034)- and two tri (6c12f, 3011) -nucleotide repeat nSSR loci were genotyped for all 45 elite selections (Table 3) and the natural population sample. The number of alleles per locus varied from 12 to 28 (select and population groups combined), with about 25% of the 54 total alleles occurring only once in the combined populations. The frequency of the most-common allele at each locus varied from 0.22 to 0.42. For two of the loci (3034 and 6c12f), gene frequencies in the test population and the population sample were similar. However, for locus 3011 they were dramatically different. Of the 28 alleles detected, only eight were shared by both populations. The natural population

Table 3 Accession number, known relatives within the breeding population (m=maternal parent and p=paternal parent), and molecular-marker genotypes for an experimental elite breeding population of loblolly pine. Genotypes include a 7-locus haplotype derived from chloroplast microsatellites and three diploid nuclear microsatellites

Accession	Half-sibs	Full-sibs	Known parent	Known grand-parent	Genotype			
					Haplo-type	6c12f	3011	3034
7					1	161/176	150/204	216/216
30	32 m		33 m		4	173/176	135/238	220/212
17	19p,21 m				5	173/176	150/243	220/216
19	17p,21 m				5	161/176	189/243	218/216
3					5	170/173	150/150	230/228
8					7	182/191	231/255	228/216
15	16 m				8	173/191	204/243	228/228
41					16	0	0	220/228
12		13p	14 m		19	164/176	150/243	228/216
13		12p	14 m		19	197/173	162/238	228/216
18		20p,22p			19	173/176	150/150	224/214
20		18p,22p			19	164/173	135/150*	224/214
22		18p,20p			19	173/176	150/162	228/224
24	26 m	25p	27 m	45 m	19	167/182	201/243	216/216
25	26 m	24p	27 m	45 m	19	167/182	201/204	226/216
1					28	167/173	150/232	224/220
32	30 m		33 m		29	164/176	168/238	226/212
35			43p		34	164/170	150/150	226/224
43					34	164/173	150/168	228/226
14					35	173/173	186/201	216/216
16	15 m				35	173/173	135/150	226/220
37					35	173/173	150/168	226/216
38					36	161/176	150/204	228/224
4			5,6 m/p		37	164/164	150/150	224/216
5					37	164/164	150/150	224/222
6					37	164/173	150/150	216/216
2					37	161/173	204/261	226/220
10	9 m		11p		42	164/173	150/201	226/216
11					42	167/173	135/201	226/226
34			42p		45	164/191	150/243	226/214
21	17 m,19 m		23p		45	161/173	150/189	222/214
23					45	173/173	177/177	216/214
42					45	164/173	150/216	218/214
9	10 m		11 m		45	173/173	201/223	226/216
26	24 m,25 m		27p	45 m	48	167/176	150/204	228/226
27			45 m		48	167/182	204/243	226/216
33					53	176/176	150/238	228/212
44					57	173/191	135/150	228/228
36	39p	40p	45p		60	182/191	162/243	226/220
39	36p,40p		45p		60	173/182	135/243	226/226
40	39p	36p	45p		60	182/191	243/243	226/220
45					60	182/191	243/243	226/224
28	31p	29p	35 m	43p	63	164/173	150/204	228/224
29	31p	28p	35 m	43p	63	164/167	150/204	228/226
31	28p,29p		34 m	42p	63	164/164	150/150	216/214

group had 12 private alleles, the elite population sample had eight. The most common allele in the natural population (177; 0. 52%) occurred in only one individual in the study population. The two populations originated from geographically distinct areas in the eastern range of the species in the southeast United States. Williams et al. (2000) have demonstrated that populations of *P. taeda* from the eastern and western portions of the species-range possess some private alleles, but the differences do not seem to be of the order noted here.

Twenty one putative cases of either maternal- or paternal – offspring relatedness occurred in this elite population. In three of these cases, observed nSSR genotypes were not consistent with expected genotypes. Accession 21 shared an appropriate haplotype with the putative father, accession 23, but did not share a common allele at nSSR 3011. Accessions 12 and 13, putative full-sibs, did not share a common allele with the putative mother (14), at one locus in the case of #13, and at two loci in the case of #14 (Table 3). Such inconsistencies could have arisen by (1) inaccurate genotyping, (2) mutations, or (3) an inaccurately labeled pedigree. The latter case is known to occur with a higher than acceptable frequency in tree breeding (Adams et al. 1988). The routine use of markers in confirming parentage in progeny tests, as suggested in the PMX/WPA scenarios, would greatly reduce this occurrence.

(3) Fingerprinting: the combination of paternally inherited cpSSR haplotypes and three Mendelian nSSR genotypes provided a unique fingerprint for each of the 45

selections in this study, despite the appreciable level of relatedness.

(4) Paternity assignment: the ability to assign paternity unambiguously to offspring is the core requirement for the successful adoption of PMX/WPA. Statistical approaches to assigning paternity range from straightforward genetic exclusion to sophisticated likelihood models (Smouse and Meagher 1994; Marshall et al. 1998; Goodnight and Queller 1999).

Exclusion, where technically and economically feasible, would be desirable in breeding populations where the tracking of pedigree and the level of inbreeding are critical. Based on the results of the elite population studied here, it would not be possible to unambiguously assign the paternity of progeny generated from a complete mix of the individuals, largely as a result of the considerable relatedness observed in the population. In two instances it would be impossible to distinguish between progeny of two unrelated trees (Table 3, Accessions 16 and 37 and Accessions 12 and 22) with the markers used in this study. Complete exclusion of progeny derived from a polymix of this test population could be obtained in the following three ways:

(i) Increase the number of cp markers. The addition of one or more cpSTS markers (Stoehr et al. 1998) would virtually eliminate the opportunity for unrelated individuals to share haplotypes.

(ii) Increase the number of nuclear markers. An increase in the number of microsatellites or alternative markers such as SNPs (Chakraborty et al. 1999; Krawczak 1999) or AFLPs (Krauss 1999; Mueller and Wolfenbarger 1999) would separate both unrelated and related individuals. Gerber et al. (1999) showed that four to six highly polymorphic nSSRs, such as those used in this study, were sufficient to produce paternal exclusion of 99.9% in a natural population of oak. Approximately 45 AFLP or SNPs loci, with a low frequency of ”presence” alleles, provide similar exclusion probabilities (Chakroborty et al. 1999; Gerber et al. 1999). Since first-order relatives always share at least one allele at every locus, paternity exclusion of related individuals is dependent on the number of loci available, their level of polymorphism, and population substructure.

(iii) Alternatives to the addition of new markers : complete paternal exclusion can be insured by (1) complete elimination of paternally related individuals from the breeding population, or (2) the creation of polymixes that avoid mixing pollens known to share the same haplotype. Two approaches to the latter option exist. The maximum number of pollen parents can be retained by creating several polymixes, each one customized for a separate group of paternally related individuals. For the test population used here, assuming that haplotype exclusion of all unrelated individuals could be achieved, six polymixes would be necessary, varying from 42–44 pollen parents each. Multiple polymixes require more pollen, increase chances of error and require additional record keeping. An alternative is a single polymix for all trees, excluding all paternally related individuals but one for each related group in order to facilitate application of PMX/WPA scenarios #1 and #2. In this example, the polymix would consist of 27 individuals, an entirely acceptable number for breeding value analysis and sufficient males to ensure a broad representation of males in the resulting progeny. Similarly, elimination of another three individuals from the pollen mix would exclude maternal relatedness which would facilitate the application of PMX/WPA scenario #3.

Likelihood models are used to statistically assign paternity, parentage, or other levels of relatedness (Marshall et al. 1998; Goodnight and Queller 1999). However, such models are generally designed to deal with natural populations and assume random mating and non-relatedness. Likelihood approaches may be quite flexible and make full utilization of the data, accounting for such factors as allele frequencies and the potential for genotyping errors or mutations (by modeling double-exclusion requirements; Marshall et al. 1998). Furthermore, paternal assignments may be made with relatively few loci making the approach cost-effective. The most significant shortcoming of the likelihood approach is that the probability of assigning paternity incorrectly, given relatededness, can be non-zero. Valuable insight into paternity assignment is provided by Double et al. (1997). They provide equations for calculating single-locus exclusion probabilities when related males compete for mating, as is the case in our elite populations. They also model exclusion probabilities based on the number of loci used and their level of polymorphism. For instance, to achieve an exclusion probability in excess of 90% when one, five or ten first-order relatives are competing, five, eight and nine loci, respectively, each with ten equally frequent alleles are required. If needed, this might be both technically and economically feasible, but would require significant marker development.

(5) Pedigree analysis: the PMX/WPA scenario #3 requires that both maternity and paternity be determined unambiguously for all progeny. Likelihood models provide one approach to satisfying this need. In this study, the analytical software KINSHIP (Goodnight and Queller 1999) was used to estimate relatedness (probability of sharing an allele by descent) for all possible pairs of individuals (990 combinations) and for all potential hypothetical relationships. Based on known relationships, the program-estimated averages were reasonably accurate (Table 4), but variation about the mean was significant. Similar dispersion was noted in salmon (Norris et al. 2000). Furthermore, high estimates of relatedness were obtained for a number of unrelated pairs. Without the addition of many more loci, it is apparent that likelihood approaches to assigning parentage for both parents would not be precise enough to meet the needs of this scenario.

It should be noted that single-locus or haplotype maternal exclusion can be developed relatively easily for conifers using maternally inherited mitochondrial DNA markers (Wagner 1992). Though variation is known to exist in the mtDNA of conifers (Deverno et al. 1993;

Table 4 Expected (R_e) and average (standard error in parentheses) observed estimates of relatedness (probability of sharing an allele by descent) for known pedigree relationships in the elite test population

Grandparent – offspring	Parent – offspring	Full-sibs	Half-sibs
N=6	N=21	N=7	N=12
R_e=0.25	R_e=0.50	R_e=0.50	R_e=0.25
0.31 (0.10)	0.44 (0.05)	0.53 (0.09)	0.31 (0.06)

Dong and Wagner 1993; Aagard et al. 1995), a suite of highly informative markers is not currently available for pine. Notable concerns for the development of mt markers would include high rates of mutation in hot spots and heteroplasmy (Wagner et al. 1991; Hipkins et al. 1994).

General applicability of PMX/WPA

Advantages of the PMX/WPA scenarios (advantages may depend on the chosen scenario) as compared to currently used full-sib breeding and testing systems include:

(1) Logistical simplicity:

(i) Breeding is simpler, easier and less expensive (Cotterill 1986a, b) because there are fewer crosses to make and fewer constraints on parental combinations. It is even conceivable that the breeding could be completed a year earlier in conifer species with the simpler PMX design. These advantages are especially true if only one pollen mix is used.
(ii) Testing is simpler because there are fewer crosses to test. Testing may also be less expensive if the same precision on genetic information such as breeding values or gca and sca estimation (scenarios #2 and 3) can be achieved with fewer total individuals in the progeny test due to a more-efficient breeding design.

(2) Improved estimation of breeding values:

Breeding values are more reliably estimated because each parent gets crossed with more of the other parents (Burdon and van Buijtenen 1990; Bridgwater 1992; White 1996). This assumes that several male parents will fertilize each female.

(3) Increased gain:

Gain from forward selection will be greater because there is a greater likelihood of best-parent by best-parent matings. In general, for a given number of individuals tested, gain is greater with many crosses and few individuals per cross rather than few crosses and many individuals per cross, especially for low-heritability traits (Pederson 1972, White 1996). A PMX cross is effectively a collection of several full-sib crosses and, if fertilization by pollen parents is random, then random-number generation indicates that nearly all of the possible crosses among all parents will be realized under most conventional testing designs. This improved gain also comes about because there is better opportunity to control inbreeding (i.e.), there are more choices of high value crosses from which selections can be made that can meet desired coancestry constraints.

(4) Maintaining genetic integrity:

Molecular marker genotypying of all selected progeny provides a quality control check on genetic identity. In most breeding programs if identity errors occur they would go undetected which could possibly lead to later inbreeding complications or nonrealization of expected gains. Studies have revealed that identity errors in breeding and production programs are common (Adams et al. 1988); such errors could strongly affect genetic-parameter estimation and seriously affect genetic gain if error rates exceed 2% (Ericsson 1999).

(5) Tie-in with other molecular genetics objectives:

Genetic fingerprinting of progeny and parents for PMX/WPA applications could readily tie in with marker-assisted selection, association genetics and transformation genetics.

The value of PMX/WPA depends greatly on whether or not genotyping be done at a low enough cost to make it competitive with conventional full-sib breeding and testing. We believe that costs are already low enough to make PMX/WPA scenario #1 a viable option since only a small percentage of the total progeny test population would be genotyped. To make scenarios #2 and #3 viable, genotyping costs would likely need to be much lower than today's costs. However, development of new marker technologies (e.g., AFLP and single nucleotide polymorphisms or SNPs) along with automated DNA sample analyses (e.g., DNA chips or microarrays) could lead to applicability of scenarios #2 and #3 within the next few years.

Table 5 provides a subjective summary of comparisons between PMX/WPA and other breeding and testing systems for many typical tree-improvement programs, assuming the above advantages are achievable. These subjective ratings need verification through experimentation and computer simulations; nonetheless, the relative scores suggest that the PMX/WPA breeding and testing scenarios deserve careful study.

Potential disadvantages of the PMX/WPA scenarios include:

Table 5 Comparisons of open-pollinated, full-sib and polymix breeding and testing systems including PMX/WPA scenarios one through three. Subjective ratings are relative to each other where: 1=worst through 5=best. Ratings assume that the PMX/WPA parental-analysis cost targets of $30/tree for scenario one and $3/tree for scenarios two and three are achievable. Comparisons are for a 40-parents per generation breeding line

Item	Open pollinated	Full-sib (circular)	PMX	PMX/WPA Scenario 1	PMX/WPA Scenario 2	PMX/WPA Scenario 3
Number of crosses to make	0	80	40	40	40	40
Number of families to test	40	80	40	40	40	40
Simplicity of breeding	5	1	3	3	3	3
Simplicity of testing	5	1	5	5	5	5
Simplicity of pedigree analysis	5 (None)	5 (None)	5 (None)	3	2	1
Simplicity score	**15**	**7**	**13**	**11**	**10**	**9**
Parental breeding Value and GCA Estimation	5	1	5	5	5	5
SCA variance Estimation	1	3	1	1	5	5
SCA of Specific crosses	1	5	1	1	3	3
Advanced generation selection gain	1	3	2	4	5	5
Inbreeding control	1	3	1	4	5	5
Pedigree error control	1	3	1	5	5	5
Value score	**10**	**18**	**11**	**20**	**28**	**28**
Total	**25**	**25**	**24**	**31**	**38**	**37**

(1) Differential reproductive success:

A key assumption for some of the advantages of the PMX/WPA breeding and testing system is that several males are effective in fertilization. If few males dominate fertilization events, benefits are significantly reduced. Unequal or reduced male contribution could occur if: (1) few males are used in the pollen mix, (2) some of the pollen lots are inviable, or (3) there is genetically determined differential male reproductive success in fertlization.

These problems appear to be negligible or easily managed. We recommend using several males (>20) in PMX/WPA crossing schemes. Careful pollen collection, processing, storing and testing will ensure all pollen lots are equally viable. Regarding (3), there have been a number of studies in trees on the relative reproductive success of male pollen in mixtures (Tobolski 1983; Lindgren and Yazdani 1988, Moran and Griffin 1985, Apsit et al. 1989; Rogers and Boyle 1991, Nakamura and Wheeler 1992, Skroppa and Lindgren 1994). Though most of these studies were done with few (two to six) pollen parents in the mixture, differential paternal parent contributions were slight. Only the two spruce (*Picea*) species showed very meaningful departures from expectation in the effective fertilization by different males (Rogers and Boyle 1991; Skroppa and Lindgren 1994). Fowler (1987) points out that only a few pollen grains (four for *P. taeda*) can occupy the micropyle in a competitive position for fertilization of the ovule in many coniferous species. This fact reduces the probability that any one pollen parent will dominate in fertilization because it is unlikely that pollen grains of the same parent would occur in a high percentage of the micropyles if several parents are used in the polymix. There was no significant departure from expectation in male parental representation in the progeny when a nine-parent pollen mix was applied to four females in loblolly pine (Wiselogel and van Buijtenen 1988).

(2) Inbreeding:

In certain instances some of the pollen in a mix could be ”self” pollen or pollen from a closely related individual. The latter can occur in advanced generations of breeding in closed breeding lines of small population size. This could result in a condition where the pollen mix is more-closely related to some female parents than it is to others and the related pollen fertilizes differentially, or when fertilization results in differential levels of inbreeding among the PMX progeny groups. Different levels of inbreeding among crosses can occur in full-sib mating designs as well, especially in advanced generations. When the full parentage (and pedigree) of all individuals in the progeny test is known, as is the case for full-sib breeding and in PMX/WPA scenarios #2 and #3, it is possible to analyze the effects of the differential inbreeding levels and account for it in the calculation of breeding values. In PMX/WPA scenario #1 the male parentage is determined for only the few individuals who are candidates for the next generation of breeding so that such adjustments are not possible. This potential bias in breeding values depends on how successful the related pollen is at producing viable progeny and on the differential success of ”self” fertilizations. Nonetheless, the degree of this differential relatedness of the pollen mix with different females does not seem to be a serious problem for most current tree-breeding lines since parents tend not to be closely related and because self pollen would be only a small portion of the pollen mix. Furthermore, self pollen in a mix with unrelated pollen (and probably closely related pollen as well) does not compete well to produce

viable progeny for testing, presumably due to genetic load (embryonic lethal alleles occuring in a homozygous state; Lindgren and Yazdani 1988). Two possible solutions to this potential "related pollen" problem include: (1) create large polymixes with many males in the pollen mix, (2) create custom pollen mixes for each group of unrelated females leaving out self and closely related pollens.

(3) Specific cross SCA:

PMX breeding in scenario #1 does not permit estimates of specific combining ability (SCA) and is therefore appropriate in programs that exploit primarily general combining ability (GCA). In scenarios #2 and #3 estimates of SCA variance at a population level are well estimated even with limited numbers of trees per PMX cross. However, it would be necessary to produce several progeny per cross in order to generate SCA on a specific cross basis. Specific-cross SCA may be valuable in the commercial production of full-sib crosses but GCA is the primary source of gains from one generation to another in breeding populations.

Conclusions

Polymix breeding, combined with parental analysis of the progeny (PMX/WPA) using genetic markers, appears to be a viable alternative to currently used full-sib breeding and testing systems. Potential key advantages are simpler and less-expensive breeding and testing, greater genetic gain and an elimination of identity errors. PMX/WPA scenario #1 requires paternal analysis of a small percentage of the progeny test population that have a high breeding value based on the breeding value of the mother and the performance of the individual progeny. Based on the paternal analysis only those from high breeding value fathers and those that help in coancestry control are used for breeding for the next generation. PMX/WPA scenario #2 requires paternity analysis of all PMX progeny such that both male and female are known for selection work. This effectively yields many crosses for genetic parameter estimation and provides flexibility for high genetic gain and inbreeding control. PMX/WPA scenario #3 requires full parental analysis of all PMX progeny. This approach provides the same benefits of scenario #2 but might be desirable where maintaining parental identity in tests is difficult or expensive.

Successful implementation of a PMX/WPA program is dependent on marker technologies that provide complete parental exclusion. The preliminary data reported here suggest that relatively few, carefully chosen markers (7–10) may be adequate to meet most breeding population needs for scenarios #1 and #2 today. Although scenarios #2 and #3 may be prohibitively expensive at this point in time, new marker technologies and automation of DNA sample analysis should make them feasible within the next few years. The actual number of markers required depends on many factors including population size, marker informativeness, gene frequencies, genotyping error rates, etc. Relatedness among pollen parents and female parents in the breeding population greatly increases the difficulty of determining the parentage of PMX progeny. Relatedness within the breeding population can be accomodated by creating two or more customized pollen-mix lots that improve paternal exclusion probabilities.

Appendix 1: detailed material and methods

DNA extraction

For the CTAB method, 100 mg of needle tissue was placed in a 1.5-ml microfuge tube, frozen with liquid nitrogen, and ground to a fine powder using a plastic pestle. The needle sample was suspended in 700 µl of 2× CTAB buffer, with 200 µg of Proteinase K. The procedure yielded approximately 10 µg. The DNA samples were resuspended in 100 µl of TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). For the DNeasy method, 50 mg of needle tissue was placed in a 2-ml screw cap tube. The tube contained a ceramic ball and cylinder (BIO101, Carlsbad, Calif.). AP1 buffer was added (600 µl) with RNaseA (100 mg/ml, 1.5 ul per sample) and Reagent DX (detergent, 1.5 µl per sample). The samples were homogenized for 45 s by agitation (up and down with twisting, at high speed, setting 4.5) in the FastPrep instrument (BIO101, Carlsbad, Calif.). The samples were processed 12 at a time and held on ice until 96 were homogenized. The homogenates were incubated at 65°C for 30 min, then AP2 buffer (protein precipitation, 200 µl) was added and the tubes were shaken for 15 s. The tubes were then chilled on ice for five minutes before being centrifuged for 5 min at 5600 g. Supernatents (400 µl) were transferred to 96 2-ml tubes (mtp format), then AP3 buffer was added (600 µl) and mixed by pipetting up and down several times. The solutions (approximately 1 ml) were transferred to a DNeasy 96 plate and centrifuged at 5600 g for 4 min. AW buffer was added (800 ul) then the plate was centrifuged at 5600 g for 15 min. The DNA was eluted from the plate by adding 100 µl of AE buffer, incubating for 1 min, then centrifuging at 5600 g for 2 min. This step was repeated for maximum yield, resulting in 15 to 20 µg of DNA in 200 µl. The concentration of loblolly pine samples was estimated by visual comparison using the fluorescence of ethidium-stained lambda DNA concentration standards on 0.8% agarose gels illuminated with UV-light.

PCR-amplification

The sense primer was fluorescently labeled with IRD-41. The PCR-reaction was carried out in 5 µl using 40 ng of loblolly pine genomic DNA as a template. The PCR reaction protocol (Jang, Genetic Analysis Bulletin #30,

LI-COR, Inc., Lincoln, Nebraska) used PCR buffer and 1 unit of *Taq* polymerase (Boehringer Mannheim, 10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3), 2 pmol of reverse (anti-sense) primer, 0.1 pmol of IR-labeled forward (sense) primer, 0.1 mM of dNTP (using dGTP). The reactions were run in 96-well polycarbonate microtitre plates (25 µl conical-bottom wells). The thermocycler temperature program was (94°C, 30 s; 58°C, 30 s; 72°C, 1 m) repeated ten times, dropping the annealing temperature by 1 degree each cycle, followed by 23 cycles using the original 58°C annealing temperature. The stop solution contained 95% formamide, 0.1 mM EDTA, 0.1% bromophenol blue, and 5 µl of stop solution was added to each reaction. Multiplexed PCR reactions contained primer pairs for four microsatellites (0.1 pmol of each labeled reverse primer, and 2.0 pmol of each forward primer).

DNA fragments containing nuclear microsatellite sequences were PCR-amplified from loblolly pine genomic DNA using three pairs of DNA primers. The DNA sequences for two primer pairs, PtTX3011 and PtTX3034, were obtained from the Texas A&M University web page of Dr. Claire Williams (http://forestry.tamu.edu/genetics/primer.txt). The other primer pair was designed from the loblolly pine EST sequence 6c12f (GenBank accession AA556811; see Table 3). This EST was found by BLAST search of EST sequences in GenBank using ten repeats of the trinucleotide CAG. The DNA sequence from the M13–29 forward sequencing primer was incorporated at the 5′ end of each forward primer sequence (CACGACGTTGTAAAACGAC, tailed primer strategy, Oetting et al. 1995). The protocol is described at http://biosupport.licor.com/support/RnP/protocols/Tail-Prim.shtml. An M13 primer with a 5′ infrared fluorophore label, IRD41, was used as the only source of labeled primer for detection for all nuclear microsatellites. The PCR-reaction was carried out in 10 µl using 40 ng of loblolly pine genomic DNA as template. The PCR reaction used PCR buffer and 1 unit of *Taq* polymerase (Boehringer Mannheim, 10 mM Tris-HCl, 1.5 mM MgCl2, 50 mM KCl, pH 8.3), 2 pmol of anti-sense primer, 0.1 pmol of IR-labeled sense primer, 0.1 mM of dNTP (using dGTP). The reactions were run in 96-well polycarbonate microtitre plates (25 µl conical-bottom wells). The thermocycler temperature program was (94°C, 10 s; 65°C, 30 s; 72°C, 1 m) repeated 13 times, dropping the annealing temperature by 0.7 degrees for each cycle, followed by 23 cycles using (94° C, 10 s; 56°C, 30 s; 72°C). The stop solution contained 95% formamide, 0.1 mM EDTA, 0.1% bromophenol blue, and 5 µl of stop solution was added to each reaction.

Electrophoresis

The gels were run at 50°C with constant power (70 W), and with the scanner collecting 16-bit video data with motor speed 3. The IRD-labeled STR molecular-weight marker contained 15 DNA fragments that ranged from 50 to 350 bp (LI-COR). Loblolly pine accession 7–56 (NCSU-Industry Cooperative Tree Improvement Program, Raleigh) was used as a genotyping control sample for the cpDNA microsatellites. A 48-well comb was used, and each run contained three molecular-weight marker standards, six control sample lanes, and 40 test sample lanes. For multiplexed runs, aliquots of two PCR reactions were loaded in each lane.

Note The experiments described in this manuscript were done in compliance with the laws of the United States of America where they were conducted.

References

Aagard JE, Vollmer SS, Sorensen FC, Strauss SH (1995) Mictochondrial DNA products are frequent among RAPD profiles and strongly differentiated among races of Douglas-fir. Mol Ecol 4:441–447

Adams WT, Neale DB, Loopstra CA (1988) Verifying controlled crosses in conifer tree-improvement programs. Silvae Genet 37:147–152

Apsit VJ, Nakamura RR, Wheeler NC (1989) Differential male reproductive success in Douglas fir. Theor Appl Genet 77:681–684

Blouin MS, Parsons M, Lacaille Y, Lotz S (1996) Use of microsatellite loci to classify individuals by relatedness. Mol Ecol 3:393–401

Bridgwater FE (1992) Mating designs. In: Fins L, Friedman S, Brotschol J (eds). Handbook of quantitative forest genetics. Kluwer Academic Publishers, Dordrect, The Netherlands, pp 140–194

Buijtenen JP van, Burdon RD (1990) Expected efficiencies of mating designs for advanced-generation selection. Can J For Res 20:1648–1663

Burdon RD, Shelbourne CJA (1971) Breeding populations for recurrent selection: conflicts and possible solutions. NZ J For Sci 1:174–193

Burdon RD, van Buijtenen JP (1990) Expected efficiencies of mating designs for reselection of parents. Can J For Res 20:1664–1671

Cervera MT, Gusmao J, Steenackers M, Van Gysel A, Van Montagu M, Boerjan W (1996) Application of AFLP™ – based markers to breeding of *Populus* spp. Plant Growth Regul 20:47–52

Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. Electrophoresis 20:1682–1696

Cotterill PP (1986a) Breeding strategy: don't underestimate simplicity. Proc IUFRO Conf on Breeding Theory, Progeny Testing and Seed Orchards. Williamsburg, Virginia, USA, pp 144–149

Cotterill PP (1986b) Genetic gains expected from alternative breeding strategies including simple low cost options. Silvae Genet 35 (5–6):212–223

Deverno L, Charest PJ, Bonen L (1993) Inheritance of mitochondrial DNA in the conifer *Larix*. Theor Appl Genet 86:383–388

Dewoody JA, Dewoody YD, Fiumera AC, Avise JC (2000) On the number of reproductives contributing to a half-sib progeny array. Genet Res 75:95–105

Dinus RJ, Tuskan GA (1997) Integration of molecular and classical genetics: a synergistic approach to tree improvement. In: Klopfenstein et al. (eds) Micropropagation, genetic engineering and molecular biology of *Populus*. USFS Gen Tech Rep RM-GTR-297, pp 220–235

Dong J, Wagner DB (1993) Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta.* Theor Appl Genet 86:573–578

Double MC, Cockburn A, Barry SC, Smouse PE (1997) Exclusion probabilities for single-locus paternity analysis when related males compete for mating. Mol Ecol 6:1155–1166

Dow BD, Ashley MV (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak *Quercus macrocarpa.* Mol Ecol 5:615–627

Doyle JJ, Doyle JI (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Elsik CG, Minihan VT, Hall SE, Scarpa AM, Williams CG (2000) Low-copy microsatellite markers for *Pinus taeda* L. Genome 43:550–555

Ericsson T (1999) The effect of pedigree error by misidentification of individual trees on genetic evaluation of a full-sib experiment. Silvae Genet 48:239–242

Fowler DP (1987) In defense of the polycross. Can J For Res 17:1624–1627

Gerber S, Mariette S, Streiff R, Bodenes C, Kremer A (1999) Comparison of microsatellites and AFLP markers for parentage analysis. In: Gillet EM (ed) Which DNA marker for which purpose. Final Compendium of the Research Project "Development, optimisation and validation of molecular tools for assessment of biodiversity inforest trees" in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity, pp 1–7

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J Hered 88:335–342

Goodnight KF, Queller DC (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. Mol Ecol 8:1231–1234

Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchell-Olds T, Sherman B, Vujcic C, Williams C, Neil D (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138:1293–1300

Hipkins VD, Krutovskii KV, Strauss SH (1994) Organelle genomes in conifers: structure, evolution and diversity. For Genet 1:179–189

Hohoff C, Brinkmann B (1999) Human identity testing with PCR-based systems. Mol Biotech 13:123–136

Huber DA, White TL, Hodge GR (1992) The efficiency of half-sib, half-diallel and circular mating designs in the estimation of genetic parameters in forestry: a simulation. For Sci 38:757–776

Isagi Y, Kanazashi T, Suzuki W, Tanaka H, Abe T (2000) Microsatellite analysis of the regeneration process of *Magnolia obovata* Thunb. Heredity 84:143–151

Kerr RJ (1998) Asymptotic rates of response from forest tree breeding strategies using best linear unbiased prediction. Theor Appl Genet 96:484–493

Krauss SL (1999) Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). Mol Ecol 8:217–226

Krawczak M (1999) Informativity assessment for biallelic single nucleotide polymorphisms. Electrophoresis 20:1676–1681

Lindgren D, Yazdani R (1988) Paternal contributions following artificial pollination of *Pinus sylvestris* (L.). Scand J For Res 3:299–304

Lowe WJ, van Buijtenen JP (1986) The development of a sublining system in an operational tree improvement program. In: Proc IUFRO Conf on Breeding and Testing Theory, Progeny Testing and Seed Orchards. Williamsburg, Virginia, pp 98–106

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655

McKeand SE, Bridgwater FE (1992) Third-generation breeding strategy for the North Carolina State University Industry Cooperative Tree Improvement Progoram. In: Proc IUFRO Resolving Tropical Forest Resource Concerns Through Tree Improvement, Gene Conservation, and Domestication of New Species. Cali, Colombia, pp 234–240

Moran GF, Griffin AR (1985) Non-random contribution of pollen in polycrosses of *Pinus radiata* D. Don. Silvae Genet 34:117–121

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. Tree 14:389–394

Namkoong G (1979) Introduction to quantitative genetics in forestry. U.S.D.A. Forest Service Technical Bulletin No 1588

Neale DB, Wheeler NC, Allard RW (1986) Paternal inheritance of chloroplast DNA in Douglas-fir. Can J For Res 16:1152–1154

Nakamura RR, Wheeler NC (1992) Pollen competition and paternal success in Douglas-fir. Evolution 46: 846–851

Norris AT, Bradley DG, Cunningham EP (2000) Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar)* using microsatellite markers. Aquaculture 182:73–83

Oetting WS, Lee HK, Glanders DJ, Wiesner GL, Sellers TA, King RA (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30:450–458

O'Malley DM, McKeand SE (1994) Marker assisted selection for breeding value in forest trees. For Genet 1:207–218

O'Malley DM, Whetten R (1997) Molecular markers and forest trees. In: Caetano-Anolles G, Gresshoff PM (eds) DNA markers: protocols, applications and overviews. John Wiley and Sons, New York, pp 237–257

O'Malley DM, Grattapaglia D, Chaparro JX, Wilcox PL, Amerson HV, Liu B-H, Whetten R, McKeand S, Kuhlmann EG, McCord S, Crane B, Sederoff R (1996) Molecular markers, forest genetics, and tree breeding. In: Perry Gustafson J, Flavell RB (eds) Genomes of plants and animals: 21st Stadler Genetics Symposium. Plenum Press, New York, pp 87–102

Pederson DG (1972) A comparison of four experimental designs for the estimation of heritability. Theor Appl Genet 42:371–377

Rogers DL, Boyle TJB (1991) Unequal paternal contributions in black spruce polycross seedlots. Heredity 67:373–379

Rosvall O, Andersson EW (1999) Group merit selection compared to conventional restricted selection for trade-offs between gain and diversity. For Genet 6:1–14

Shelbourne CJA (1969) Tree breeding methods. Technical Paper No 55, New Zealand Forest Service

Skroppa T, Lindgren D (1994) Male fertility variation and non-random segregation in pollen mix crosses of *Picea abies*. For Genet 1:13–22

Smouse PE, Meagher TR (1994) Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L) Gray (Liliaceace). Genetics 136:313–332

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. HortScience 31:729–740

Stoehr MU, Orvar BL, Vo TM, Gawley JR, Webber JE, Newton CH (1998) Application of a chloroplast DNA marker in seed orchard management evaluations of Douglas-fir. Can J For Res 28:187–195

Tobolski JJ (1983) Isozyme identification of competing pollen parents in seed from mixed pollen crosses between Austrian pines and Austrian×Japanese red pines. Proc 28th Northeast Tree Improvement Conf, Durham, New Hampshire, pp 262–273

Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. Mol Ecol 5:111–114

Wagner DB (1992) Nuclear, chloroplast, and mitochondrial DNA polymorphisms as biochemical markers in population genetic analyses of forest trees. New For 6:373–390

Wagner DB, Dong J, Carlson MR, Yanchuk AD (1991) Paternal leakage of mitochondrial DNA in *Pinus*. Theor Appl Genet 82:510–514

Wakasugi TJ, Tsudzuki SI, Nakashima K, Tsudzuki T, Sugiura M (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc Natl Acad Sci USA 91:9794–9798

Walter C, Carson SD, Menzies MI, Richardson T, Carson M (1998) Review: application of biotechnology to forestry – molecular biology of conifers. World J Microbiol Biotech 14:321–330

Wheeler NC, Jech KS (1992) The use of electrophoretic markers in seed orchard research. New For 6:311–328

White T (1996) Genetic parameter estimates and breeding value predictions: issues and implications in tree improvement programs. In: Proc QFRI-IUFRO Conf, Tree Improvement for Sustainable Tropical Forestry, Caloundra, Australia pp 110–117

White TL, Hodge GR, Powell GL (1993) Advanced-generation breeding strategy for slash pine in the southeastern United States. Silvae Genet 42:359–371

Williams CG, Elsik CG, Barnes RD (2000) Microsatellite analysis of *Pinus taeda* L. in Zimbabwe. Heredity 84:261–268

Wiselogel AE, van Buijtenen JP (1988) Probability of equal mating in polymix pollinations of loblolly pine. Silvae Genet 37 (5–6):184–187

Ziegenhagen B, Scholz F, Madaghiele A, Vendramin GG (1998) Chloroplast microsatellites as markers for paternity analysis in *Abies alba*. Can J For Res 28:317–321