Tree Genetics & Genomes (2007) 3:207–214
DOI 10.1007/s11295-006-0056-y

ORIGINAL PAPER

Testing for unequal paternal contributions using nuclear and chloroplast SSR markers in polycross families of radiata pine

S. Kumar · S. Gerber · T. E. Richardson · L. Gea

Received: 6 December 2005 / Revised: 30 July 2006 / Accepted: 23 August 2006 / Published online: 28 October 2006

Abstract The lack of male pedigree control is the major limitation of an otherwise very useful and cost effective mating design, namely, the polycross. This study was conducted to investigate the relative contribution of different pollen parents to the sound-seeds stage, and also in a field progeny trial. Pollen from 15 radiata pine (*Pinus radiata* D. Don) parents was mixed in equal volume and applied to the same 15 parents, potentially allowing selfing. Samples of 8-year-old offspring were genotyped from five polycross families, and available seed from three of these five families was tested for unequal paternal contributions. The total paternal exclusion probability of five chloroplast markers and four microsatellite markers in our study was 99.1%. Overall, 81% of the offspring (both seeds and 8-year-old offspring) were assigned to 1 out of the 15 potential male parents, but a surprisingly high proportion (about 13%) was evidently fathered by pollen not included in the pollen-mix. Inconclusive evidence of unequal paternal contribution was observed in some families, but it did not influence the general combining ability (GCA) estimates appreciably, as evident from a high degree of correspondence between GCA estimates obtained from polycross and female-tester mating designs. A non-significant negative correlation was observed between the relative reproductive success (across polycross families) and predicted breeding values for diameter growth.

Keywords Molecular markers · Polycross · Pedigree reconstruction · Reproductive success · Selfing · GCA · *Pinus radiata*

Communicated by R. Johnson

S. Kumar (✉) · L. Gea
Ensis Genetics,
Private Bag 3020,
Rotorua, New Zealand
e-mail: Satish.Kumar@ensisjv.com

S. Gerber
INRA, Équipe de Génétique et Amélioration
des Arbres Forestiers,
69 route d'Arcachon,
33612 Cestas Cedex, France

T. E. Richardson
Scion,
Private Bag 3020,
Rotorua, New Zealand

Introduction

Genetic testing is the basis for any successful tree improvement programme. The main objectives of genetic testing include progeny testing, estimation of genetic parameters, advancement of the breeding population for recurrent selection and breeding, and estimation of genetic gain (Zobel and Talbert 1984). The mating designs that are used to create the progeny population determine the *type* of information that will be derived from the genetic testing, and the experimental design that is employed when the test is established in the field will determine the *quality* of the information that is obtained. Numerous mating designs were proposed for forest trees. These mating designs may be divided into two general classes:

(1) *Incomplete-pedigree designs* in which only one parent (the seed parent in forest trees) is known for a given offspring. Some common examples of such design include open-pollinated (OP) mating of parent ortets or archived ramets, and polycross (or polymix) mating.
(2) *Complete-pedigree designs* in which both parents are known for a given offspring. Some common examples

of such design include factorial designs, single-pair mating, diallel mating, and female-tester. The concept of 'female testing' assumes that the tested progeny of crosses of candidate pollen parents with four or five females give accurate estimates of the general combining ability (GCA) of pollen parents. The female-tester mating design has many advantages, which include providing comparable estimates of GCA values of top-production population parents, coping with maternal effects and capitalizing on early pollen fecundity. The major drawback of female-tester design is that we only have a limited number of available unrelated selections.

The advantage of the incomplete-pedigree designs, such as OP and polycross, is that they are cost-effective ways of providing estimates of additive genetic variances, heritabilities, and breeding values of the female parents. However, because the male parent's identity is unknown, estimates of non-additive variance and specific combining abilities are not possible. In addition, there is a risk of future inbreeding, when forwards selections from OP and polycross tests are used in later advanced-generation breeding programmes. The complete-pedigree designs, such as factorial and diallel, will yield information on the GCA of all parents, specific combining ability (SCA) of all crosses, and variance-component information for estimating narrow-sense and broad-sense heritabilities. Since both parents of each offspring are known, it is easy to control the level of inbreeding. Thus, complete-pedigree designs might appear to be ideal, but their major disadvantage is that the large number of crosses required can often make them expensive and time-consuming.

The use of molecular marker technologies for parental analysis in breeding programmes could provide a solution to deficiencies of incomplete-pedigree designs. The lack of male pedigree control, and the lack of male GCA information, are two major limitations of the incomplete-pedigree designs. The use of pollen from a known group of parents could overcome the lack of male GCA information in polycross designs. If a pollen-mix of a known group of pollen parents is mated with each seed parent, molecular markers could then be used to identify the male parents of forwards selections. Various practical scenarios for using molecular markers to identify male parentage in polycross designs were reported previously (e.g. Lambeth et al. 2001).

Estimation of GCA from polycross designs assumes equal contribution of pollen parents. Failure of this assumption leads to biased GCA estimates. The paternal contributions in controlled polycross families were shown to be unequal for *Populus* spp. (Wheeler et al. 2006), *Picea mariana* (Rogers and Boyle 1991), *Pseudotsuga menziesii* (Apsit et al. 1989; Nakamura and Wheeler 1992), *Picea abies* (Schoen and Cheliak 1987; Skrøppa and Lindgren 1994), *Picea glauca* (Schoen and Stewart 1986) and *P. radiata* (Moran and Griffin 1985), but the equal mating hypothesis was verified for *Pinus taeda* (Wiselogel and van Buijtenen 1988). All of these studies used isozyme loci, but a few recent studies have used highly polymorphic nuclear microsatellites (nSSR) (Grattapaglia et al. 2004) and chloroplast microsatellites (cpSSR) (Plomion et al. 2001) for paternity analyses in polycross seed orchards. Chloroplast markers are non-recombining and paternally inherited in conifers, and thus prove very useful for paternity analyses (Neale et al. 1986; Wagner 1992; Cato and Richardson 1996).

All previous studies, except Wheeler et al. 2006, have tested for unequal pollen contribution using seeds from polycross families. Parental GCA estimates are obtained from young offspring and thus this hypothesis should actually be tested using offspring that were planted in progeny trials. Since only healthy and vigorous seedlings are typically planted in progeny trials and generally all seedlings do not survive by the selection age (7–9 years in radiata pine), the relative contributions of pollen parents could deviate from those observed at the sound-seeds stage. Thus, the main objective of this study was to investigate, using nSSR and cpSSR markers, the relative contributions of different pollen parents at the seed stage, and also after the polymix offspring were planted in a progeny trial. A comparison of parental GCA values obtained from polycross designs could be biased if pollen contribution is unequal. Female-tester designs, however, provide comparable estimates of GCA values and thus a comparison of female-tester and polycross design was also made in this study to check if there are significant differences in GCA estimates from these two mating designs. The implications of the results on the radiata pine breeding were also examined.

Materials and methods

Experimental material

A set of 15 radiata pine breeding population parents was used as seed parents, and controlled pollination of all 15 parents with pollen from the same 15 parents was conducted in a clonal archive in 1988. Pollens were mixed in equal volumes and applied to each seed parent, thus allowing some selfing to occur. Pollination was conducted by bagging the female flowers and introducing the pollen-mix into the bags with a hypodermic needle. Seeds were sown in the nursery, and the seedlings were planted in a

replicated field progeny trial in 1992. Surplus seeds from each cross were kept in cold storage. In 2003, the polycross families with more than 75 surplus seeds were considered for paternity analyses. There were only three families (Table 1) with enough seeds available. Offspring from a set of 5 polycross families, out of 15 families that were planted in the progeny trial in 1992, were also sampled for paternity analyses (Table 1).

Molecular marker assays

A set of four nuclear microsatellite (nSSR) markers and five chloroplast microsatellite (cpSSR) markers was used for genotyping seeds and needle samples. The four nSSR markers (NZPR0143, NZPR0290, NZPR0006, PR4.6) were previously selected specifically for their DNA-profiling applications after evaluating their reproducibility, distribution of allelic sizes, and level of polymorphism. These loci were previously mapped on a radiata pine linkage map, and were found to be unlinked (Wilcox et al. 2001; Kumar et al. 2004). The set of five cpSSR loci used in this study had previously been used in fingerprinting radiata pine (Cato and Richardson 1996). Assuming that the chloroplast genome of radiata pine does not recombine, the set of five cpSSR loci could be viewed as a single locus. For each individual, the alleles scored at these five loci were combined to derive the chloroplast haplotype.

Along with the chloroplast haplotype, the two-allele genotypes of the 15 parents were obtained at four nSSR loci using DNA extracted from needle samples collected from a clonal archive. DNA extracted from the embryos from each of the three polycross families (1, 5, and 10) was used to obtain genotypes of seeds at cpSSR and nSSR loci. Megagametophyte tissues from six seeds of each of these three polycross families were also genotyped at four nSSR loci to infer maternal genotypes. Genotypes of 8-year-old offspring from each of the five polycross families were obtained at cpSSR and nSSR loci, using needle samples.

Table 1 The number of seeds and 8-year-old offspring sampled from each polycross family

	Seeds			Offspring		
Family	Sample size	Maternal mismatch	Self	Sample size	Maternal mismatch	Self
1	84	28	2	45	14	0
3	–	–	–	46	2	1
4	–	–	–	42	25	0
5	90	23	4	53	16	0
10	88	2	5	43	7	1
Total	262	53 (20%)	11 (5%)	229	64 (28%)	2 (1%)

The number of maternal mismatches and putative self in each family are also shown

Paternity analysis and hypothesis testing

Paternity analyses of seeds and 8-year-old offspring were conducted separately using the computer software FaMoz (Gerber et al. 2003). The paternity of each seed of three families, and 8-year-old offspring from five families was determined. The observed frequencies of pollen contributions were tested against those expected under equal mating, using the chi-square test of goodness of fit, in each family separately. The number of 8-year-old offspring sampled from each family is small and thus the expected number in different classes would be less than 5, which could require combining neighboring classes. However, Snedecor and Cochran (1967, page 235) suggested that chi-square test of *single* classification is accurate enough if the smallest expectation is at least one. Since this condition was met in each polycross family in our study, different classes were not combined. To test the hypothesis that the ratios of progeny sired by different pollen parents were homogeneous among polycross families, the *G*-test (Sokal and Rohlf 1981) was applied by assuming the different polycross families as replicates. In *multiple* classification tests, Yate's correction is generally used when sample sizes are small. However, following the suggestions of Sokal and Rohlf (1981), we used Williams' correction (Williams 1976).

Comparison of GCA estimates

The 8 (out of 15) parents that were used as seed parents in polycross families were also crossed individually to a set of five different female parents ("female-tester" design), and the offspring resulting from these two mating designs were planted together in a replicated experiment at two sites (Kinleith and Kaingaroa) in the central North Island of New Zealand. Estimates of parental means from the two mating designs were compared for a set of seven parents (including the five shown in Table 1) that were common to the two field trials. A key point to note here is that the same parents were used as female parents in the polycross design, and as male parents in the female-tester design. It was previously reported (e.g. Kumar and Wu 2003) that maternal effects could exist for some parents, but estimated correlation between GCA estimates obtained with and without taking into account maternal effects were almost perfect for growth and form traits. Thus, a comparison of parental GCA estimates from the two mating designs should be feasible using the following model:

$$\text{Phenotype} = \mu + R_i + F_j + M_k + FM_{jk} + \text{error}$$

where μ is the grand mean, R_i is the effect of replicate, F_j is the effect of family, M_k is the effect of mating design, FM_{jk} is the interaction of parent and mating design, and error is the residual effect. In this model, replicate and mating design effects were treated as fixed effects while all other effects were assumed random. This model was implemented using SAS PROC GLM (SAS Institute Inc. 1989).

Results and discussion

Allele frequencies and maternal verification

Most of the individuals produced reliable two-allele genotypes at the nSSR loci. The number of alleles, in the parental population (15 parents), at the nSSR loci NZPR0143, NZPR0290, NZPR0006, PR4.6 were 9, 13, 9 and 10, respectively, with allele frequencies varying between 0.03–0.33, 0.03–0.17, 0.03–0.33 and 0.04–0.32, respectively. Maternal alleles inferred from genotyping megagametophytes from three polycross families (1, 5, and 10) matched with the two-allele maternal genotypes obtained from genotyping their needle samples from the clonal archive.

Genotypes of many seeds and 8-year-old offspring did not conform to their nominal maternal genotypes at one or more nSSR loci. All such cases were called 'maternal mismatch' and their frequencies are shown in Table 1. Various factors, including inaccurate genotyping, mutations, and mislabeling, could have caused 'maternal mismatch'. As SSR mutation rates in most plants are sufficiently low, and mismatches were apparent only at one or two loci, it is likely that genotyping error could be the prime cause of such discrepancies in this study. Since the pollen-mix included pollen from the mother trees as well, there is a possibility of obtaining selfed offspring. An embryo was considered 'self' if it inherited only the maternal alleles at all nSSR loci. Since there was no parent that shared one or both alleles with the seed parent at all four nSSR loci, the criteria used for identifying selfed embryos in our study should be robust. All seeds or 8-year-old offspring identified as 'maternal mismatch' or 'self' were excluded from paternity analyses. The numbers of such cases in each polycross family is shown in Table 1.

The estimated frequencies of selfed seeds, calculated after discarding maternal mismatches, were 3.6%, 6.0% and 5.8% in families 1, 5 and 10, respectively. In the progeny trial, however, the frequency of selfed offspring varied from 0% to about 3% (Table 1). In addition to the typically low germination rate of selfed seeds, the post-germination inbreeding depression would lead to the culling of most of the selfed offspring at the nursery level. Our results from three families suggest that the frequency of selfed offspring planted in progeny trials or commercial plantations would be very low (approximately less than 3%). Estimated selfing rate in our study could, however, differ from those observed in open-pollination in natural stands (e.g. Vogl et al. 2002) of radiata pine.

Unlike nSSR, the number of alleles per locus at the five cpSSR loci was rather limited, varying from two to three. Eight different 5-locus cpSSR haplotypes were observed in the parental population with some parents having identical haplotypes (Table 2).

Paternity analyses and test for unequal pollen contribution

All cases of 'maternal mismatch' and 'self' (shown in Table 1) were excluded from paternity analyses, and thus sample size was reduced considerably for some families, especially family 4. We assigned paternity in two ways, first by using only the cpSSR loci and then combining both cpSSR and nSSR loci. Results for each approach are described separately:

(1) *cpSSR*: The observed frequencies of the cpSSR haplotypes were compared with those expected based on the paternal haplotype (shown in Table 2). Probabilities of getting observed data under the null hypothesis of equal contribution, using chi-square test with 7 *df*, are shown in Table 3. Results showed that the observed contribution of different pollen parents in seeds was significant in family 1 (p=0.012) and family 10 (p=0.015). In the 8-year-old offspring samples in all five families, paternal contributions did not differ significantly from those expected under the null hypothesis (Table 3). Since some pollen parents shared the same cpSSR haplotype, testing for equal contribution using only the cpSSR haplotype frequencies has some drawbacks. If any two pollen parents have the same cpSSR haplotype and the female parent is incompatible with either one, then the use of only cpSSR haplotype frequencies could have biased the results in favour of the null hypothesis. If the contaminant pollen shares the

Table 2 Frequency of chloroplast haplotypes in the parental population

cpSSR haplotype	Number of parents	Frequency
21122	4	0.2667
21212	1	0.0667
21221	3	0.2
21222	3	0.2
21223	1	0.0667
31222	1	0.0667
31223	1	0.0667
42222	1	0.0667

Table 3 Chi-square test of goodness of fit of observed frequencies to that expected given the equal representation of 15 pollen parents in seeds and 8-year-old offspring

	cpSSR		cpSSR and nSSR combined	
Family	Seeds	Offspring	Seeds	Offspring
1	0.012	0.136	0.511	0.355
3	–	0.209	–	0.058
4	–	0.065	–	0.056
5	0.075	0.287	0.316	0.471
10	0.015	0.871	0.012	0.889

Probabilities of getting the observed data under the null hypothesis are shown separately for cpSSR loci alone, and for cpSSR and nSSR combined

haplotype with any parent in the pollen-mix, then the results could be biased in favour of the alternative hypothesis.

(2) *cpSSR and nSSR combined*: In majority of the cases, it was possible to assign a unique pollen parent to each progeny (Table 4). The number of progeny (both seeds and 8-year-old offspring) that were assigned to specific males varied from 71% (family 1) to 90% (family 5). Overall, 81% of the offspring were assigned to a single male parent, 6% were assigned to more than one male parent, and 13% were evidently fathered by outside pollen. The high proportion of offspring that could be linked to a unique male is largely due to the use of highly polymorphic nSSR in combination with cpSSR loci. The total paternal exclusion probability of four nSSR markers in our study was 95%, which increased to 99.1% after including five cpSSR loci. Gerber et al. (1999) showed that 4–6 highly polymorphic nSSR were sufficient to produce paternal exclusion of 99.9% in a natural population of oak. Thus, addition of one or more cpSSR and nSSR in our study could have virtually eliminated the possibility of unrelated individuals sharing the same haplotypes. Chakraborty et al. (1989) suggested that paternity determination by exclusion criteria alone would give a rather optimistic idea about the ability to assign paternity based on a given set of markers. Selection of breeding individuals using wrong paternity assignment could have significant impact on the genetic progress. Various studies (e.g. Geldermann et al. 1986; Long et al. 1990; Ericsson 1999; Dodds et al. 2005) showed the impact of errors of misidentification on the estimates of genetic parameters and breeding values.

Table 4 The number (%) of progenies assigned to a unique (1 out of 15) pollen parent and more than one pollen parent

	Unique pollen parent		More than one pollen parent		Outside pollen	
Family	Seed	Progeny test	Seed	Progeny test	Seed	Progeny test
1	74	71	13	10	13	19
3	–	74	–	7	–	19
4	–	88	–	6	–	6
5	90	83	5	3	5	14
10	82	80	1	6	17	14

Estimated frequency of contaminant pollen is also shown

It is surprising to note that the estimated frequency of pollen contamination in our study is much higher than one would expect from controlled-pollination. Although female flowers were bagged at the right time, some other factors could have contributed to these unexpected results. Since pollen is constantly flying during the collection phase, undesirable pollen could have been present on the outside of the collected microsporangiate. Other possibilities include the collection of some microsporangiates of incorrect identity along with those of the prescribed clone, and accidentally collecting OP cones while the CP cones were being harvested. Kumar and Richardson (2005) reported that about 5% of the documented relationships (e.g. full-sibs, parent–offspring, etc.) in the New Zealand radiata pine breeding population were excluded by the molecular marker data. In a detailed study using allozymes in conifer tree improvement programmes in the United States, Adams et al. (1988) found that about 30% of the controlled crosses in Douglas-fir and loblolly pine were incorrect. Mislabeling of individuals during various field operations is the likely explanation for discrepancies of this sort. It should also be noted that DNA markers, when identical in size, may represent alleles that are identical-in-state rather than identical-by-descent, which could reduce the reliability of pedigree reconstruction using molecular markers.

Only those seeds and 8-year-old offspring that were assigned to a unique pollen parent were considered for testing the hypothesis of equal contribution of different pollen parents. The expected frequencies were obtained by dividing the number of offspring by the number of non-self pollen parents. In the sample of seeds from family 10, paternal contribution differed significantly ($p=0.012$) from that expected under the null hypothesis. The null hypothesis of equal representation of all pollen parents was accepted ($p>0.05$) for samples of 8-year-old offspring in each polycross family (Table 3). However, three out of eight chi-square tests were significant at $p=0.06$, thus there is inconclusive evidence that equal contribution of pollen is not the case in some families.

Since the number of seeds or offspring sampled from each polycross family was small, we also tested for equal pollen contribution across all maternal parents. In pooled samples, the probabilities of getting the observed data under the null hypothesis were 0.031 and 0.061 for seeds and 8-year-old offspring, respectively. When putative self seeds and offspring were also included in pooled samples, the paternal contribution differed non-significantly (p=0.143 and 0.077, respectively). However, further studies would be required to test whether the absence of evidence for unequal paternal contribution in this study was due to lack of statistical power. Assuming that there were five parents in the pollen-mix, and one parent contributed twice and one had no representation—it would require genotyping about 25 offspring to reject the null hypothesis of equal contribution at p=0.05. For a similar reproductive success scenario in a 10, 15 and 20 pollen–parent mix, we would need about 90, 180 and 320 offspring, respectively, to declare it significantly (p=0.05) different from equal contribution. Since there could be numerous combinations of unequal contribution, it would require developing an algorithm to calculate the power of test for various sample sizes.

Our results do not entirely agree with an earlier report (Moran and Griffin 1985) of unequal pollen contribution in polycross families of radiata pine. Since there were only three or four pollen parents represented in their pollen-mix, it is likely that results presented by Moran and Griffin (1985) were affected by the unusual competitive advantage of one male parent. In fact, a particular male parent was reported overly successful in three out of the four polycross families and thus Moran and Griffin (1985) suggested that a larger sample of pollen donors should be used to validate their results. Using a pollen-mix of three parents, Skrøppa and Lindgren (1994) reported that the number of seeds sired by each pollen parent differed significantly in majority of the polycross families of Norway spruce. However, using a pollen-mix of 16 parents, Schoen and Cheliak (1987) reported that in majority of the polycross families of Norway spruce, the male gamete frequencies in the seeds did not differ significantly from expectation under the assumption of equal male fertility. In loblolly pine, with a nine-parent pollen-mix applied to four females, there was no significant departure from expectation in male parental representation (Wiselogel and van Buijtenen 1988). These results suggest that the number of pollen donors used in the pollen-mix have a significant impact on their relative reproductive success in polycross families. However, unequal male contribution could also occur if there is genetically determined differential male reproductive success in fertilization.

Even if the hypothesis of equal representation is not rejected in some families, the progenies sired by different pollen parents in different polycross families may not be homogenous. The G-test showed significant (p=0.0007) differences among polycross families in ratios of seeds from each pollen parent. For example, 7%, 10% and 18% seeds were credited to male parent 14 in polycross families 1, 5 and 10 respectively, and 5%, 0% and 15%, respectively, to male parent 7 (results not tabulated). These results provide some indication that no particular male parent had an advantage in every polycross family, and female parent genotype appeared to influence the selection between pollen gametes. Similar results were reported by Rogers and Boyle (1991) in polycross seedlots of black spruce. Apsit et al. (1989) suggested that differential reproductive success in Douglas-fir could be due to the abortion of some embryos sired by certain pollen parent. Potential causes of such male–female complementarity include inbreeding or outbreeding depression (Charlesworth et al. 1987). In the sample of 8-year-old offspring in our study, the G-test for heterogeneity was significant (p=0.026), indicating that offspring sired by different pollen parents in different polycross families were not homogenous. When putative self seeds and offspring were also included in the analyses, the G-test was still significant ($p \leq 0.05$)

Comparison of GCA estimates

Analysis of variance for diameter growth (DBH) at age 8 years revealed that the effects of mating type and family-by-mating type interaction were not significant at both sites (Table 5). Correlation between family means, obtained from polycross and female-tester designs, was 0.60 and 0.70 for DBH at Kinleith and Kaingaroa, respectively. These correlations were somewhat higher (0.71 and 0.87, respectively) for a relatively high heritability trait such as branch cluster frequency. As mentioned earlier, the parents whose GCA estimates were obtained were used as female parents in the polycross design and as male parents in the tester design. With the exception of seven common 'half-sib' parents, the set of female parents used in the female-tester designs was different from the set of pollen parents used in the polycross design. Even after these complications, a high

Table 5 Probabilities of getting observed data under the null hypothesis of no different family-by-mating and mating type effects for diameter growth at age 8-years at the two sites (Kinleith and Kaingaroa)

Source	*df*	Kinleith	Kaingaroa
Replicates	29	0.278	0.132
Mating type	1	0.204	0.164
Families	6	0.104	0.052
Families-by-mating type	6	0.208	0.517
Residual	241		

magnitude of parental–mean correlations suggested that GCA ranking obtained from polycross designs, such as in this study, would be satisfactory. Relative efficiencies of various mating designs for the reselection of parents (backwards selection) were reported as generally similar (Pepper and Namkoong 1978; Burdon and van Buijtenen 1990).

Implications for forest tree breeding

This study was designed to test for evidence of unequal contribution of pollen parents in polycross families of radiata pine. When deviations from equal mating occur, the GCA of some female parents may be over or under-estimated because the mixture of male gametes effecting fertilisation may contain an excess of superior or inferior genotypes (Wiselogel and van Buijtenen 1988). Our results provided inconclusive evidence that the number of offspring sired by different pollen parents in some polycross families could differ from that expected under the assumption of equal contribution, but it did not appear to influence GCA estimates appreciably because a good correspondence between GCA estimates obtained from polycross and female-tester design was observed. Thus, polycross designs may be used successfully as an economical system for evaluating GCA provided that the number of male parents represented in the pollen-mix is large (say, ≥15).

Conventionally, the use of polycross designs was restricted to GCA estimation only because lack of information on the male parentage has deterred breeders from any advanced-generation selection. Lambeth et al. (2001) proposed the use of molecular markers to identify the male parents of forwards selections, and thus enhancing the utility of polycross designs. Various practical scenarios for using molecular markers to identify male parentage in polycross designs were presented by Lambeth et al. (2001). However, it is not always straightforward to assign a unique male parent to each offspring. Overall, about 6% of the offspring in our study were assigned multiple male parentage (shown in Table 4). The presence of gametes indistinguishable for several males is more common, when using isozyme loci. Rogers and Boyle (1991), using seven isozyme loci, reported that the number of progenies that were attributable to unique males ranged from 25% to 52% in various black spruce polycross seedlots. It is, however, possible to reduce this number further by increasing the number of markers and/or creation of pollen-mixes that avoid parents known to share same haplotypes (Lambeth et al. 2001). Since the occurrence of multiple parentage is a possibility when using molecular markers, Dodds et al. (2005) proposed various statistical approaches to account for multiple possibilities of male pedigrees during the genetic evaluation process.

Within-family selection is an important component of advanced-generation breeding plans in forest trees. Using polycross designs, as a stand-alone system, for the advancement of breeding population could have some drawbacks (van Buijtenen and Burdon 1990; Bridgwater 1992). Any major differences in the relative reproductive success (i.e. the number of offspring sired) of male parents could result in related selections from the polycross families and thus an increase in inbreeding level in advanced generations. Avoiding related selections will result in reduced genetic gain in the short-term. The magnitude of this problem would be much higher if the overly successful male parent has relatively high GCA. The relative representation of different pollen parents in 8-year-old offspring, across all five families, in our study varied from about 2% to 12%. However, in the combined sample of seeds and 8-year-old offspring, the relative representation of pollen parents varied from 4% to 11%. A non-significant negative correlation of 0.15 was obtained between predicted breeding values for breast height diameter (obtained previously from a large number of progeny trials) and the relative reproductive success of the 15 parents. Since no particular male parent was found to be overly successful in every polycross family, and the correlation between GCA and reproductive success was insignificant, it should be possible, via pedigree reconstruction, to identify advanced-generation selections and minimise inbreeding. Also, pedigree reconstruction in polycross families will essentially convert the polycross design into a complete-pedigree design and will allow estimation of the specific combining ability (SCA) of all crosses, and variance-component information for estimating narrow-sense and broad-sense heritabilities.

Acknowledgements This project was funded by the New Zealand Foundation for Research, Science and Technology. Toby Stovold, Mark Miller and Kane Fleet are thanked for collecting the seeds and needle samples. Signagen Laboratory carried out the genotyping work. We thank Rowland Burdon and Tony Shelbourne for the helpful comments on an earlier version of this manuscript. We thank Randy Johnson, Nick Wheeler and an anonymous referee for providing constructive criticism and helpful suggestions.

References

Adams WT, Neale DB, Loopstra CA (1988) Verifying controlled crosses in conifer tree improvement programmes. 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

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

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

Cato SA, Richardson TE (1996) Inter- and intraspecific polymorphism at chloroplast SSR loci and the inheritance of plastids in *Pinus radiata* D. Don. Theor Appl Genet 93:587–592

Chakraborty R, Meaghert TR, Smouse PE (1989) Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. Genetics 118:528–536

Charlesworth D, Schemske DW, Sork VL (1987) The evolution of plant reproductive characters: sexual versus natural selection. In: Stearns SC (ed) The evolution of sex and its consequences. Birkhäuser, Basel, pp 317–335

Dodds KG, Tate ML, Sise JA (2005) Genetic evaluation using parentage information from genetic markers. J Anim Sci 83:2271–2279

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

Geldermann H, Pieper U, Weber WE (1986) Effect of misidentification on the estimation of breeding value and heritability in cattle. J Anim Sci 63:1759–1768

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 in forest trees" in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity, pp 1–7

Gerber S, Chabrier P, Kremer A (2003) FaMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. Mol Ecol Notes 3:479–481

Grattapaglia D, Ribeiro VJ, Rezende GDSP (2004) Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. Theor Appl Genet 109:192–199

Kumar S, Richardson TE (2005) Inferring relatedness and heritability using molecular markers in radiata pine. Mol Breed 15:55–64

Kumar S, Wu H (2003) Reciprocal and maternal effects on growth and form traits in radiata pine in New Zealand. Silvae Genet 52:71–74

Kumar S, Echt C, Wilcox PL, Richardson TE (2004) Testing for linkage disequilibrium in the New Zealand radiata pine breeding population. Theor Appl Genet 108:292–298

Lambeth C, Lee BC, O'Malley D, Wheeler N (2001) Polymix breeding with parental analysis of progeny: an alternative to full-sib breeding and testing. Theor Appl Genet 103:930–943

Long TE, Johnson RK, Keele JW (1990) Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine. J Anim Sci 68:4069–4078

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

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

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

Pepper WD, Namkoong G (1978) Comparing efficiency of balanced mating designs for progeny testing. Silvae Genet 27:161–169

Plomion C, LeProvost G, Pot D, Vendramin G, Gerber S, Decroocq S, Brach J, Raffin A, Pastuszka P (2001) Pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using chloroplast microsatellites. Can J For Res 31:1816–1825

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

SAS Institute Inc. (1989) SAS/STAT user's guide, release 6.03 edition. SAS Institute Inc., Cary, NC, USA

Schoen DJ, Cheliak WM (1987) Genetics of the polycross. 2. Male fertility variation in Norway Spruce, *Picea abies* (L.) Karst. Theor Appl Genet 74:554–559

Schoen DJ, Stewart SC (1986) Variation in male reproductive investment and male reproductive success in white spruce. Evolution 40:1109–1120

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

Snedecor GW, Cochran WG (1967) Statistical methods. 6th edn. Iowa State University Press, Ames, Iowa

Sokal RR, Rohlf FJ (1981) Biometry. 2nd edn. W. H. Freeman Company, New York

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

Vogl C, Karhu A, Moran G, Savolainen O (2002) High resolution analysis of mating systems: inbreeding in natural populations of *Pinus radiata*. J Evol Biol 15:433–440

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

Wheeler N, Payne P, Hipkins V, Saich R, Kenny S, Tuskan G (2006) Polymix breeding with paternity analysis in Populus: a test for differential reproductive success (DRS) among pollen donors. Tree Genetics & Genomes 2:53–60

Wilcox PL, Richardson TE, Corbett GE, Ball RD, Lee JR, Djorovic A. Carson SD (2001) Framework linkage maps of *Pinus radiata* D. Don based on pseudotestcross markers. For Genet 8:109–117

Williams DA (1976) Improved likelihood ratio tests for complete contingency tables. Biometrika 63:33–37

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

Zobel BJ, Talbert JT (1984) Applied forest tree improvement. Wiley, New York