Theor Appl Genet (1985) 71:375–384

The nature of inbreeding in a seed orchard of Douglas fir as shown by an efficient multilocus model

K. Ritland [1, *] and Y. A. El-Kassaby [2]

[1] Department of Botany, University of British Columbia, Vancouver, BC V6T 1W5, Canada
[2] CIP Inc. Tahsis Pacific Region, 8067 East Saanich Road, R.R. 1, Saanichton, B.C. V0S 1M0, Canada and Faculty of Forestry, University of British Columbia, Vancouver, B.C. V6T 1W5, Canada

Received June 29, 1985; Accepted July 31, 1985
Communicated by P. M. A. Tigerstedt

Summary. The amounts of self-fertilization versus consanguineous matings (as measured by effective selfing) was estimated in a seed orchard of Douglas-fir, using progeny array data at six allozyme loci. The orchard is family structured, consisting several grafts (clones) and/or open-pollinated (o-p) progeny from each of several 'plus-trees'. Population-wide selfing rates were found to be 7% for the o-p trees and 2% for the cloned trees. Estimates of mating system parameters for individual trees showed this difference for average outcrossing rate t (1) still largely remained when outcrossing-pollen gene frequency p was not allowed to vary among trees and (2) disappeared when p was allowed to vary among trees. Under this joint t and p estimation, o-p trees showed both significant variation of t (based upon a one-way ANOVA grouped by common plus-tree) and significant regressions of p on ovule genotype (indicative of consanguineous matings); cloned trees showed neither. This higher rate of consanguineous mating for o-p trees might be explained by the larger and more variable size of o-p families in the orchard. Estimates of outcrossing rate t and outcrossing-pollen gene frequency p were based upon a multilocus model which makes full use of the information in the data. The increased information it gives over 'observed outcross' models is equivalent to adding 30–50% more loci, and it gives enough degrees of freedom to jointly estimate t and p for individual trees (individual progeny arrays) under certain conditions. In addition, inclusion of megagametophyte data nearly doubles the information about the mating system of individual trees.

Key words: Outcrossing – Douglas fir – Seed orchard – Inbreeding

Introduction

Conifers are among the most genetically heterozygous of plants, at least at electrophoretic loci (Mitton 1983). The maintenance and utilization of this high level of diversity for reforestation is a major criterion for breeding programs. As seed orchards become a predominant source for the production of conifer seed, for ease and frequency of harvest of genetically improved seed, the breeding systems of conifers in seed orchards merit increased scientific and managerial attention (Weir and Zobel 1975; Adams and Joly 1983).

Several measures can control the mating system of seed orchard trees. Water spray cooling can delay reproductive development and reduce pollen contamination from outside sources (Fashler and Devitt 1980), isolation zones can limit pollen contamination (Denison and Franklin 1975), certain planting arrangements can limit consanguineous matings (Bell and Fletcher 1975; Hadders and Koski 1975), and monitoring of pollination by phenological observations (Fashler and Sziklai 1980; O'Reilly et al. 1983; El-Kassaby et al. 1984) and application of supplemental mass pollination (Denison and Franklin 1975) can increase outcrossing and fertility.

The ability of these measures to control the pattern of genetic relatedness between mates can be estimated through either 'indirect' or 'direct' methods. Indirect methods include using pollen traps to measure pollen contamination (Fashler and Devitt 1980), monitoring reproductive phenology to assess the potential for cross-breeding (El-Kassaby et al. 1984), and examining proportions of aborted or empty seed to assess selfing (McKinley and Cunningham 1983). Direct methods use Mendelian gene markers such as terpenes (Squillace and Goddard 1982) or, more commonly, allozymes (Adams 1983). The pattern of genetic segregation of these biochemical markers in progenies derived from common mothers ('progeny

* Current address: Department of Botany, University of Toronto, Toronto, Ontario, M5S 1A1, Canada

arrays') provides information about the percentage of zygotes derived from self-fertilization (Brown and Allard 1970; Clegg 1980).

In conifers, allozymes are particularly useful because of their high polymorphism and the presence of two types of seed tissue: the haploid megagametophyte tissue (maternally derived), and the diploid embryo tissue. Census of megagametophytes allows easy inference of the maternal genotype of each progeny array (Tigerstedt 1973) and of the pollen-derived allele in each zygote (Müller 1976). Several workers have used allozyme segregation patterns in progeny arrays to estimate the mating system in seed orchards (Rudin and Lindgren 1977; Adams and Joly 1980; Moran et al. 1980; Shen et al. 1981; Müller-Stark 1982; Shaw and Allard 1982a). In orchards with plantings of grafted plus-trees and/or seed progeny descended from the same plus-tree, consanguineous matings can occur and cause some 'effective selfing' as defined by Ritland (1984). The low rates of selfing found in seed orchards and the lack of difference between single-locus and multilocus estimates of selfing (Shaw and Allard 1982a) have suggested that orchard designs succeed in preventing consanguineous matings.

This paper assesses the extent of self-fertilization and consanguineous matings in a seed orchard of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) on Vancouver Island, British Columbia. Single- and multilocus analyses of progeny array data allow comparison of these forms of inbreeding for trees propagated by grafting of plus-trees vs. trees derived from open-pollinated seeds of plus-trees. The multilocus analysis, given in the appendix, allows estimates of mating system parameters for individual trees.

Materials and methods

The 3.4-ha CIP Inc. (formerly Pacific Forest Products Ltd.) high-elevation Douglas-fir seed orchard at Saanichton, BC is a breeding population consisting of 80 families descended from 80 'plus-trees' selected from sites ranging from 450–1,000 m in Vancouver Island and SW British Columbia. Forty-three families consisted entirely of grafts (clones) of plus-trees; 19 families consisted entirely of open-pollinated (o-p) progeny, grown from seed of a plus-tree parent; the remaining 18 families were mixtures. Average family size and age of clonal trees was 6.1 and 14, respectively, and of o-p trees was 17.4 and 17, respectively; tree height did not exceed 8 m. The orchard planting design was random but incomplete: members of the same family were spread over several blocks, and 4 blocks contained o-p trees, 4 blocks contained cloned trees, and 8 blocks contained a mixture. The orchard has an overhead irrigation system that delays reproductive bud maturation, reducing pollen contamination from adjacent natural stands (Fashler nd Devitt 1980), which was estimated to be less than 6% (El-Kassaby and Ritland 1985).

For allozyme assay, six blocks were selected (2 with exclusively o-p trees, 2 with cloned trees, and 2 mixed) and cones collected from every fertile tree. 302 trees were sampled; 226 were o-p offspring and 76 were clones. The 226 o-p offspring were derived from 37 plus-trees, and the 76 clones were derived from 41 plus-trees. Cones were air dried at room temperature, and seeds extracted, dewinged, cleaned by hand, and stored at 3 °C until use. Thirty-two seeds from each tree were electrophoretically assayed, each for its zygote genotype (diploid) and its megagametophyte gene (haploid) at 6 enzyme loci. Procedures and staining recipes are given in El-Kassaby et al. (1982a). Enzyme systems assayed were esterase (EST) E.C. 3.1.1.1; phosphogluco-isomerase (PGI) E.C. 5.3.1.9; glucose-6-phosphate dehydrogenase (G6PD) E.C. 1.1.1.49; 6-phosphoglucose dehydrogenase (6PGD) E.C. 1.1.1.44; phosphoglucomutase (PGM) E.C. 2.7.5.1; and isocitrate dehydrogenase (IDH) E.C. 1.1.1.42; their Mendelian inheritance and lack of linkage was shown in El-Kassaby et al. (1982a, b).

The genotype of each sampled tree was inferred from its megagametophyte progeny array. Single- and multilocus population estimates of outcrossing rate t and pollen gene frequencies p were calculated for the o-p tree population and for the clonal tree population, using the maximum likelihood procedure given in the appendix. Confidence intervals were found by inversion of the information matrix. This multilocus procedure also gave estimates of female outcrossing rate and outcrossing-pollen gene frequencies for each tree in the clonal sample (76 trees) and in the o-p sample (226 trees), for (1) pollen gene frequencies set constant, equal to the population frequency, and (2) pollen gene frequencies allowed to vary among families, to their maximum likelihood value. However, because the progeny array size, 32, is small for obtaining mating system estimates, and perhaps because p was constrained to the (0, 1) interval (see "Discussion"), 15–25% of the estimates did not converge ($\hat{t}$ was greater than 2) and were not used (sample sizes, Table 3). It should be noted that, at the individual level, the female outcrossing rate probably does not equal the male outcrossing rate (Horovitz and Harding 1972; Ross 1977; Ross and Gregorius 1983).

First, to detect among-tree variation of female selfing rate, for (1) cloned trees as a group and for (2) o-p trees as a group, a one-way analysis of variance was performed, with estimates grouped by common plus-tree. Second, a direct measure of effective selfing caused by consanguineous matings is the regression of outcrossing-pollen allele frequency on the additive value of the ovule genotype (Ritland 1985). It was estimated by assigning the additive values (0, 0.5, 1.0) to the maternal genotypes (*aa, Aa, AA*), where *'A'* is the most common allele and *'a'* is the class of other alleles at that locus. Regressions were found for each locus (the inferred outcrossing-pollen allele frequency for each tree was the dependent variable) and a minimum variance average over loci was found.

Results

Estimates of allele frequencies for maternal parents (the ovule pool) and for outcrossing pollen parents (the pollen pool) are listed in Table 1, for o-p trees as a group and for cloned trees as a group. Sufficient polymorphism for estimation of inbreeding parameters is apparent. Ovule pool allele frequencies did differ significantly at the 95% level from pollen pool allele frequencies in only two of the possible 24 comparisons within the two groups: at EST-1 (allele 2) and at G6PD (allele 1) within o-p trees (Table 1).

Single-locus population estimates of outcrossing rate (Table 2) show a significant departure from complete outcrossing ($t = 1.0$) at all loci for o-p trees and at 2 of 6 loci (EST-1 and 6PGD-1) for clonal trees. A

Table 1. Allele frequencies for ovule gene pool and outcrossing pollen gene pools, for clonal and open-pollinated trees. The two most common alleles for each locus, with 95% confidence intervals, are given (all loci are triallelic)

Locus/allele		Clonal trees		Open-pollinated trees	
		Ovule	Pollen	Ovule	Pollen
EST-1[a]	1	0.553±0.079	0.509±0.033	0.544±0.046	0.462±0.016
	2	0.283±0.072	0.199±0.026	0.223±0.038	0.242±0.013
PGI-2	1	0.895±0.049	0.925±0.013	0.940±0.022	0.931±0.006
	2	0.105±0.049	0.075±0.013	0.058±0.022	0.068±0.006
G6PD	1	0.507±0.079	0.531±0.020	0.460±0.046	0.513±0.011
	2	0.421±0.078	0.404±0.019	0.511±0.046	0.446±0.011
6PGD-1	1	0.908±0.046	0.883±0.013	0.894±0.028	0.884±0.007
	2	0.059±0.037	0.047±0.010	0.046±0.019	0.048±0.005
PGM	1	0.895±0.049	0.883±0.014	0.889±0.029	0.882±0.007
	2	0.099±0.047	0.076±0.012	0.053±0.021	0.045±0.005
IDH	1	0.809±0.062	0.821±0.017	0.852±0.033	0.810±0.009
	2	0.125±0.053	0.074±0.013	0.069±0.023	0.087±0.006
n=		76	1,774–1,789[b]	226	7,001–7,056[b]

[a] Based upon single-locus estimate; others based upon 5-locus estimate
[b] Range of n among loci

Table 2. Single-locus and multilocus population estimates of outcrossing for clonal and open-pollinated trees (95% confidence intervals)

	Clonal trees	Open-pollinated trees
Locus		
EST-1	0.680±0.045	0.690±0.022
PGI-2	0.932±0.072	0.899±0.036
G6PD	1.012±0.046	0.925±0.026
6PGD-1	0.934±0.057	0.931±0.028
PGM	0.990±0.049	0.899±0.030
IDH	0.991±0.046	0.975±0.024
Mean single-locus[a]		
$\hat{t}$[b]	0.972±0.025	0.925±0.013
$\hat{t}$[c]	0.981±0.023	0.932±0.013
Multilocus[a]		
$\hat{t}$	0.976±0.019	0.932±0.011
n	1,974	7,051

[a] Excluding EST-1 locus
[b] Arithmetic mean
[c] Minimum variance mean

significantly higher outcrossing rate at G6PD and PGM in clonal trees relative to o-p trees indicates that o-p trees practice more inbreeding.

There was significant heterogeneity of estimates among loci caused primarily by the esterase locus, which gave an abnormally low estimate of outcrossing in both groups of trees (EST, Table 2). After exclusion of this locus, both the arithmetic and minimum variance means of outcrossing over loci showed significant levels of selfing for both groups of trees (7% for o-p trees, 2–3% for clonal trees); the single-locus selfing rate for o-p trees was 5% greater than that of clonal trees (Table 2). The multilocus outcrossing model gave population estimates of t nearly identical to mean single-locus estimates calculated by either averaging procedure (Table 2).

When multilocus estimates of t and p are obtained for each tree (Table 3; note these are female outcrossing rate estimates), the mean estimate of t over trees was generally greater than the population estimates of t as given in Table 2. There were two methods to perform these per-tree estimates: (1) per-tree estimation of t only, keeping p constant at the population estimate, and (2) per-tree joint estimation of t and p, which is only possible with multilocus data as discussed in the appendix. The mean estimate of t using the first method was less than that given by the second method (Table 3). With the first method, the difference between the mean $\hat{t}$ values of clonal trees versus o-p trees decreased, relative to the mean $\hat{t}$ values of Table 2. With the second method, the difference between the mean $\hat{t}$ values of clonal vs. o-p trees disappeared (Table 3).

There was wide among-tree fluctuation of female t estimates (Table 3), ranging from near zero to almost 200%. Most of this variation is statistical. The among-tree variances of $\hat{t}$ per individual were of comparable magnitude for clonal vs. o-p trees (Table 3). The expected per-individual variance of $\hat{t}$, for t and p jointly estimated when t has no true variation, is 2.14 as determined by inversion of the expected information matrix (discussed in the appendix); the actual variances are nearly twice this (Table 3), indicating non-statistical

Table 3. Distribution of estimates for female outcrossing rate of individual trees, for intervals with midpoints; one-way analysis of variance (ANOVA) of female $\hat{t}$ by families; both for clonal vs. open pollinated (o-p) trees. Two methods of estimation of t were used: (1) p constant and equal to population frequency (Only $\hat{t}$) and (2) p allowed to vary among trees (Joint $\hat{t}, \hat{p}$)

Midpoint		(1) Only $\hat{t}$ (%)		(2) Joint $\hat{t}, \hat{p}$ (%)	
		Clonal	O-P	Clonal	O-P
	0.1	3	0	2	0
	0.4	6	3	11	6
	0.7	17	28	20	20
$\hat{t}$	1.0	51	53	25	40
	1.3	14	10	27	17
	1.6	5	3	11	10
	1.9	3	2	5	6
Sample size		63	215	56	191
Mean $\hat{t}$		0.982	0.952	1.05	1.05
Variance $\hat{t}$[a]		2.29	2.08	3.59	4.26
ANOVA mean squares					
between		0.125	0.104	0.194	0.245
within		0.083	0.057	0.139	0.107
ANOVA F		1.52	1.82*	1.39	2.30**
(d.f.)		(38, 24)	(36, 178)	(32, 23)	(36, 154)

[a] Per individual
* $P = 0.0058$; ** $P = 0.0002$

Table 4. Regressions of pollen allele frequency on maternal genotype for each locus, and minimum variance mean over loci

	Clonal	Open-pollinated
Locus		
PGI-2	0.022	0.074*
G6PD	−0.186	−0.003
6PGD-1	0.239*	0.167**
PGM	0.083	0.197**
IDH	0.156	0.098*
Sample size	56	191
Mean	0.032	0.103**

* $P < 0.01$; ** $P < 0.001$

variation of outcrossing. The one-way analysis of variance of outcrossing, grouped by plus-tree, gave a non-significant F for variation of selfing among cloned trees, and a significant F for variation of selfing among o-p trees (Table 3).

The regressions of outcrossing-pollen allele frequency on ovule genotype varied widely among loci (Table 4), and were significantly positive for 6PGD-1 of cloned trees and for all loci except G6PD of o-p trees. The minimum variance mean over loci was significantly positive for only the o-p trees (Table 4).

Discussion

As a population, the open-pollinated (o-p) trees showed higher selfing rates than the grafted (clonal) trees, both on a single-locus basis and on a multilocus basis. The generally low rates of selfing detected in this seed orchard, on the order of 5%, are slightly less than those reported in previous studies of Douglas-fir (7% in Sorensen 1973; 10% in El-Kassaby et al. 1981; 10% in Shaw and Allard 1982a). The lack of difference between the single-locus and multilocus estimates of selfing should indicate that most or all apparent selfing in the orchard is actual self-fertilization, and is not the result of consanguineous (non-self) matings (Shaw and Allard 1982a). However, simulation studies by Ritland and Jain (1981) show that bias of estimates of self-fertilization still remain when few loci are used for estimation, suggesting that, for our case of relatively few loci, a more direct method is needed for estimating the proportion of selfing caused by consanguineous matings.

One such method might to be use joint multilocus estimates of t and p for individual trees (progeny arrays). Such estimates tend to yield (1) t estimates that exclude effective selfing caused by mating between weaker relatives and (2) p estimates that include pollen of consanguineous matings (i.e., pollen related to the mother plant, Ritland and Ganders 1985). However, only the female component of outcrossing was measured. When female t only was estimated for each tree, the difference between the average outcrossing rates of clonals vs. o-p trees largely remained (Table 3), probably because the estimates of t still included effective selfing caused by covariation of pollen allele frequency with the maternal genotype. By contrast, when female t and p were jointly estimated for individual trees, the difference between the average outcrossing rates of clonals vs. o-p trees disappeared (Table 3). This suggests that the difference between the population estimates of selfing rates of clonal versus o-p trees in Table 2 is caused by consanguineous matings.

A direct estimate of effective selfing caused by consanguineous matings is the regression of outcrossing-pollen allele frequency on the additive value of the ovule genotype (Ritland 1985). The regression line has an expected slope equal to the amount of selfing caused by consanguineous matings. Using the p estimates derived from the joint t and p estimates for individual trees, together with the inferred maternal genotype based upon its megagametophyte segregation pattern, the regression was found have a highly significant average slope of 0.103 for the o-p population, and a non-significant slope of 0.032 for the clonal population. These regressions have probably overestimated effective selfing, but do show a significant difference between o-p and clonal trees for their rate of consanguineous mating.

Evidently the multilocus estimates of this study included much of the selfing caused by consanguineous matings, probably because only 5 loci (of the potentially many) were used. Previous studies have demonstrated that mating to relatives lowers the single-locus estimate of outcrossing (Ennos and Clegg 1982; Ellstrand and Foster 1983). The same effect appears to

have occurred with the multilocus estimates of this study. However, the single family estimates of mating system may have their own statistical bias, caused by the relatively small sample sizes of 32 in this study. In addition, allowance of outcrossing-pollen gene frequency p estimates to fall outside the (0, 1) interval (not done here) might change t estimates (negative p estimates would be caused by the incomplete nature of the data). Studies of the properties of small sample estimation and the possibility of maximum likelihoods of p outside (0, 1) are needed. Regardless, the ability of the multilocus model to produce estimates on individual trees shows that the primary value of multilocus models lies in their increased degrees of freedom, and not in their ability to lower the statistical variance of outcrossing estimates. These increased degrees of freedom should allow joint estimation of the various modes of selfing caused by self-fertilization vs. consanguineous matings, with a multilocus version of the effective selfing model of Ritland (1984, 1985).

The basis for selfing caused by consanguineous matings is identity of genes by descent. The amount of selfing caused by matings between relatives is given by the definition of 'effective selfing' (Ritland 1984): the effective selfing rate of individual A when mated to a relative B is the probability that an allele chosen at random from relative B is identical by descent to either allele at the same locus in individual A. For outbred populations, the effective selfing rate is twice the coefficient of consanguinity between mates. In our orchard, there are four mating types with associated female effective selfing rates E: (1) clonal × clonal ($E = 1.0$), (2) clonal (female) × o-p (male) ($E = 0.5$), (3) o-p (female) × clonal (male) ($E = 0.5$), and o-p × o-p ($E = 0.25$). The clonal × clonal mating would be identical to self-fertilization; the other matings could be detected as consanguineous.

How might have differential mating to relatives occurred? The orchard consists of grafts and open-pollinated offspring of 80 plus-trees, so matings between siblings of the same plus-tree can occur. The average size of an o-p family (descended from a plus-tree) was 15.6, in contrast to only 8.7 for clonal families. In addition, the coefficient of variation of family size was 1.26 for o-p families, in contrast to only 0.59 for clonal families. Both these of factors favor higher frequencies of o-p × o-p matings and lower frequencies of clonal × clonal matings. In addition, increased within-family mating would be caused by any inheritance of reproductive timing (as shown by significant correlations between bud burst dates for two years for family females and for family males, El-Kassaby et al. 1984). Overall, the higher rate of selfing among o-p trees due to consanguineous matings, as shown by the regression analysis, indicates the lower effective selfing rate of the o-p × o-p mating was offset by a much higher frequency of this mating type.

Other factors may have also contributed to the differential selfing rates observed in the orchard. The higher mortality of grafted trees in the orchard, caused by grafting incompatibility, resulted in wider spacing between cloned trees. The cloned trees also developed less robust crowns relative to o-p trees because of developmental aberrations. These two factors should allow freer air movement and mixing of pollen in those orchard blocks consisting entirely of cloned trees, promoting outbreeding among cloned trees. Ellstrand et al. (1978) found a positive association between density and rates of apparent selfing in the self-incompatible, entomophilous *Helianthus,* attributed to higher consanguinity between mates in more dense patches. In this orchard, the same relation might hold because closer spacing of trees would restrict wind-borne pollen movement for the wind-pollinated Douglas-fir, promoting self-pollination.

The estimates of female t varied greatly among trees, being as great as 200%. Estimates of t greater than 100% are likely caused by sampling error rather than negative assortative mating and thus may not be 'biologically reasonable' (Cheliak et al. 1983), but were necessary to provide an unbiased estimate of average $\hat{t}$ over trees. The variation of outcrossing rate among trees was found to exceed the 'null' statistical variance, as determined by information matrix calculations, by two-fold, for both clonal and o-p trees. However, calculation of this null statistical variance may not be valid for the relatively small progeny array size of 32, since its calculation relies on asymptotic theory.

A more robust means of detecting variation of female outcrossing rates was the analysis of variance for outcrossing rates, grouped by common plus-tree. It did detect significant variation of outcrossing for the o-p trees, and not for the clonal trees, for both (1) t only estimated and (2) t and p jointly estimated (Table 3). Such variation may have been caused by a tendency for larger families to practice more consanguineous matings because of their higher frequency. Alternatively, genetic variation among o-p trees for self-fertilization rate may exist, but this seems less likely since clonal trees, being derived from the same plus-trees, should also show such genetic selfing variation.

The esterase locus displayed abnormally high selfing rates relative to the other loci. This excess inbreeding was reported in other studies of Douglas-fir (Shaw and Allard 1982b; Yeh, personal communication) and in loblolly pine (*Pinus taeda* L.) by Roberds and Conkle (1984), who in ruling out inbreeding and genetic drift, suggested that natural selection may have acted upon the esterase locus or indirectly at a linked locus. Further examination of our data for this locus revealed that megagametophytes were being produced at the expected 1 : 1 ratio from heterozygotes, but that there was an excess of heterozygous progeny from heterozygous parents. This means

that the homozygote selfs more than the heterozygote, which is exactly the pattern observed when effective selfing is caused by natural population structures (Ritland and Ganders 1985).

Significant among-locus variation of outcrossing has been reported in Douglas-fir (El-Kassaby et al. 1981; Shaw and Allard 1982a), *Pinus radiata* (Moran et al. 1980), *Pinus ponderosa* (Mitton et al. 1981), *Pinus contorta* (Yeh and Layton 1979, but see Epperson and Allard 1984), and several *Eucalyptus* species (Brown et al. 1975; Phillips and Brown 1977; Moran and Brown 1980; Yeh et al. 1983). It is not known whether differential selfing rates of homozygotes vs. heterozygotes were observed in these instances.

Although the significant differences between pollen and ovule gene frequencies at EST-1 and G6PD for the o-p trees as indicated by confidence intervals (Table 1) could be due to chance, the chi-square contingency test will often reveal heterogeneity of gene frequencies (Adams and Joly 1980; Mitton et al. 1981; El-Kassaby and Ritland 1985) that is perhaps indicative of the pollen and ovule gene populations being finite samples of some parent population. This test (not given in Table 1) did not indicate any significant heterogeneity. Differences of allele frequency between progeny and maternal parent trees were observed in a Douglas-fir natural stand (El-Kassaby et al. 1981), in *Pinus radiata* (Moran et al. 1980), and in *Eucalyptus* spp. (Moran and Brown 1980; Yeh et al. 1983), using this test.

In conclusion, inbreeding depression caused by selfing is generally regarded to be high in Douglas-fir such that most selfed zygotes either do not germinate or die as seedlings (Orr-Ewing 1954, 1957, 1965; Sorensen 1979). The inbreeding depression caused by mating to non-self relatives is not as severe as self-fertilization, but its consequences may be more critical for reforestation programs which rely on non-competitive plantings. Several instances of selection for heterozygotes have been demonstrated in Douglas-fir (Orr-Ewing 1954; Piesch and Stettler 1972; Sorensen and Miles 1974; Rehfeldt 1978; Sorensen 1979). The mildly inbred seedlings produced by consanguineous matings may survive past the seedling stage, but would still suffer from a loss of heterozygosity, which leads to inferior growth and productivity as mature trees. Thus the detection and control of consanguineous matings is of vital importance for seed orchards.

Acknowledgments. This project was supported by the BC Ministry of Forests, Research Branch, Sect 88 nos. JS-02-016 and KS-02-012; by Pacific Forest Products Ltd, Victoria, B.C.; by a National Sci Eng Res Council (Canada) grant to Fred R. Ganders (UBC); and by a Nat Institute of Health (USA) grant to KR. The authors are indebted to J. Parkinson for technical assistance and to Norm Ellstrand and Outi Muona for critical comments.

References

Adams WT (1983) Application of isozymes in tree breeding. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part A. Elsevier, Amsterdam, pp 381–400

Adams WT, Joly RJ (1980) Allozyme studies in loblolly pine seed orchards: clonal variation and frequency of progeny due to self-fertilization. Silvae Genet 29:1–4

Bell GD, Fletcher AM (1978) Computer organized orchard layouts (COOL) based on the permutated neighborhood design concept. Silvae Genet 27:223–225

Brown AHD, Allard RW (1970) Estimation of the mating system in open pollinated maize populations using isozyme polymorphisms. Genetics 66:133–145

Brown ADH, Matheson AC, Eldridge KG (1975) Estimation of the mating system of *Eucalyptus obliqua* by using allozyme polymorphisms. Aust J Bot 23:931–949

Ceppellini R, Siniscalco M, Smith CAB (1955) The estimation of gene frequencies in a randomly mating population. Ann Hum Genet 20:97–115

Cheliak WM, Morgan K, Strobeck C, Yeh FCH, Dancik BP (1983) Estimation of mating system parameters in plant populations using the EM algorithm. Theor Appl Genet 65:157–161

Clegg MT (1980) Measuring plant mating systems. Bioscience 30:814–848

Clegg MT, Kahler AL, Allard RW (1978) Estimation of life cycle components of selection in an experimental plant population. Genetics 66:133–145

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithim. J R Stat Soc, Ser B 39:1–38

Denison NP, Franklin EC (1975) Pollen management. In: Faulkner (ed) Seed orchards. For Comm London Bull 54: 92–100

El-Kassaby YA, Yeh FC, Sziklai O (1981) Estimation of the outcrossing rate of Douglas-fir (*Pseudotsuga menziesii* var. 'menziesii') using allozyme polymorphisms. Silvae Genet 30:182–184

El-Kassaby YA, Yeh FC, Sziklai O (1982a) Inheritance of allozyme variants in coastal Douglas-fir (*Pseudotsuga menziesii* var. 'menziesii'). Can J Genet Cytol 24:325–335

El-Kassaby YA, Sziklai O, Yeh FC (1982b) Linkage relationships among 19 polymorphic allozyme loci in coastal Douglas-fir (*Pseudotsuga menziesii* var. 'menziesii'). Can J Genet Cytol 24:101–108

El-Kassaby YA, Fashler AMK, Sziklai O (1984) Reproductive phenology and its impact on genetically improved seed production in a Douglas-fir seed orchard. Silvae Genet 33:120–125

El-Kassaby YA, Ritland K (1985) Low levels of pollen contamination in a Douglas-fir seed orchard as by allozyme markers. Silvae Genet (in press)

Ellstrand NC, Torres AM, Levin DA (1978) Density and the rate of apparent outcrossing in *Helianthus annuus* (Asteraceae). Syst Bot 3:403–407

Ellstrand NC, Foster KW (1983) Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). Theor Appl Genet 66:323–327

Ennos RA, Clegg MT (1982) Effect of population substructuring on estimates of outcrossing rate in plant populations. Heredity 48:283–292

Epperson BK, Allard RW (1984) Allozyme analysis of the mating system in lodgepole pine populations. Heredity 48:283–292

Fashler A, Devitt WB (1980) A practical solution to Douglas-fir seed orchard pollen contamination. For Chron 56: 237–241

Fashler A, Sziklai O (1980) The importance of flower phenology in seed orchard designs. For Chron 56:241–242

Green AG, Brown AHD, Oram RN (1980) Determination of outcrossing in a breeding population of *Lupinus albus* L. Z Pflanzenzücht 84:181–191

Hadders G, Koski V (1975) Probability of inbreeding in seed orchards. In: Faulkner (ed) Seed orchards. For Comm London Bull 54:108–116
Horovitz A, Harding J (1972) The concept of male outcrossing in hermaphrodite higher plants. Heredity 29:223–236
Kendall M, Stuart S (1979) The advanced theory of statistics, vol 2. Inference and relationship. Charles Griffin and Co, London
McKinley CR, Cunningham MW (1983) Theoretical effects of seed orchard roguing on selfing and percent filled seeds. Can J For Res 13:187–190
Mitton JB (1983) Conifers. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetic and plant breeding, part B. Elsevier, Amsterdam, pp 443–472
Mitton JB, Linhart YB, Davis ML, Sturgeon KB (1981) Estimation of outcrossing in ponderosa pine, *Pinus ponderosa* Laws. from patterns of segregation of protein polymorphisms and from frequencies of albino seedlings. Silvae Genet 30:117–121
Moran GF, Bell JC, Matheson AC (1980) The genetic structure and levels of inbreeding in a *Pinus radiata* D. Don seed orchard. Silvae Genet 29:190–193
Moran GF, Brown AHD (1980) Temporal heterogeneity of outcrossing rates in Alpine ash (*Eucalyptus delegatensis* RT Bak). Theor Appl Genet 57:101–105
Müller G (1976) A simple method of estimating rates of self-fertilization by analyzing isozymes in tree seeds. Silvae Genet 25:15–17
Müller-Starck G (1982) Reproductive systems in conifer seed orchards. I. Mating probabilities in a seed orchard of *Pinus sylvestris* L. Silvae Genet 31:188–197
Orr-Ewing AL (1954) Inbreeding experiments with Douglas-fir. For Chron 30:7–16
Orr-Ewing AL (1957) Further inbreeding studies with Douglas-fir. For Chron 33:318–332
Orr-Ewing AL (1965) Inbreeding and single crossing in Douglas-fir. For Sci 11:289–290
O'Reilly C, Parker WH, Barker JE (1983) Effect of pollination period and strobili number on random mating in a clonal seed orchard of *Picea mariana*. Silvae Genet 31:90–94
Phillips MA, Brown AHD (1977) Mating system and hybridity in *Eucalyptus pauciflora*. Aust J Biol Sci 30:337–344
Piesch RF, Stettler RF (1971) The detection of good selfers for haploid induction in Douglas-fir. Silvae Genet 20:144–148
Rao CR (1952) Advanced Statistical Methods in Biometric Research. Elsevier, New York
Rehfeldt GE (1978) The genetic structure of populations of Douglas-fir (*Pseudotsuga menziesii* var. 'glauca') as reflected by its wind pollinated progenies. Silvae Genet 27:49–52
Ritland K (1984) The effective proportion of self-fertilization with consanguineous matings in inbred populations. Genetics 106:139–152
Ritland K (1985) The genetic mating structure of subdivided populations. 1. Open-mating model. Theor Popul Biol 27:51–74
Ritland K (1986) Joint maximum likelihood estimation of genetic and mating structure using open-pollinated progenies. Biometrics (in press)
Ritland K, Jain SK (1981) A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. Heredity 47:35–52
Ritland K, Ganders FR (1985) Variation in the mating system of *Bidens menziesii* in relation to population substructure. Heredity (in press)
Roberds JH, Conkle MT (1984) Genetic structure in loblolly pine stands: allozyme variation in parents and progeny. For Sci 30:319–327
Ross MD (1977) Frequency-dependent fitness and differential outcrossing in hermaphrodite populations. Am Nat 111:200–202
Ross MD, Gregorius H-R (1983) Outcrossing and sex function in hermaphrodites: a resource-allocation model. Am Nat 121:204–222
Rudin D, Lindgren D (1977) Isozyme studies in seed orchards. Stud For Suec 139, 23 pp
Shaw DV, Allard RW (1982a) Estimation of outcrossing rates in Douglas-fir using isozyme markers. Theor Appl Genet 62:113–120
Shaw DV, Allard RW (1982b) Isozyme heterozygosity in adult and open-pollinated embryo samples of Douglas-fir. Silvae Fenn 16:115–121
Shen H-H, Rudin D, Lindgren D (1981) Study of the pollination pattern in a Scots pine seed orchard by means of isozyme analysis. Silvae Genet 30:7–15
Smith CAB (1957) Counting methods in genetical statistics. Ann Hum Genet 21:254–276
Sorensen FG (1973) Frequency of seedlings from natural self-fertilization in coastal Douglas-fir. Silvae Genet 22:20–24
Sorensen FG (1979) Estimates of self-fertility in coastal Douglas-fir from inbreeding studies. Am Nat 103:389–398
Sorensen FG, Miles RS (1974) Self-pollination effects on Douglas-fir and ponderosa pine seedlings. Silvae Genet 23:135–138
Squillace AC, Goddard RE (1982) Selfing in clonal seed orchards of slash pine. For Sci 28:71–83
Tigerstedt PMA (1973) Studies on isozyme variation in marginal and central populations of *Picea abies*. Hereditas 75:47–60
Weir RJ, Zoble BJ (1975) Advanced-generation seed orchards. In: Faulkner R (ed) Seed orchards. For Comm London Bull 54:118–127
Yeh FC, Layton C (1979) The organization of genetic variability in central and marginal populations of lodgepole pine (*Pinus contorta* ssp. *latifolia*). Can J Genet Cytol 21:487–503
Yeh FC, Brune A, Cheliak WM, Chipman DC (1983) Mating system of *Eucalyptus citriodora* in a seed-production area. Can J For Res 13:1051–1055

Appendix

This appendix describes maximum likelihood estimation of outcrossing rate and pollen gene frequencies, using progenies of common mothers assayed for their genotype at many loci and possibly also assayed for the megagametophyte (allele contributed by the mother to the zygote), which conifer seeds allow. The procedure is based upon a multilocus model that describes all segregation events, including all progeny of heterozygous parents (which 'observed outcross' models do not). This multilocus model also gives a joint estimate of outcrossing-pollen allele frequency superior to that of single-

locus estimates, because the more accurate multilocus identification of outcrosses also more accurately identifies outcrossing pollen alleles.

The data can consist of genotypic scores for either (1) both the maternally derived gamete (megagametophyte) and zygote genotye (as for conifers), (2) the zygote genotype only (as for most angiosperms), or (3) a combination of the two (due to missing megagametophyte data). The model can be considered as a multilocus extension of the single-locus 'conifer' models of Shaw and Allard (1982a) and Cheliak et al. (1983), and/or as a conifer extension of the multilocus 'angiosperm' model of Ritland and Jain (1981). Strictly, these models measure the 'female outcrossing rate' (Horovitz and Harding 1972; Ross 1977; Ross and Gregorius 1983) because only the female parent genotype is identified. The population average outcrossing rate is identical for males and females (Ross and Gregorius 1983); the above models seek this rate. Individual tree estimates do require the strict definition of being female estimates.

To describe the multilocus estimation model, first consider a single locus. The mixed mating model that optionally includes megagametophytic segregation of maternal parents is given by the table of progeny probabilities given parent genotype

Maternal parent					
Genotype:		A_iA_i	A_iA_j		
Gamete:		A_i	A_i	A_j	Unknown
	A_iA_i	$s+p_it$	$\frac{1}{2}s+p_it$	0	$\frac{1}{4}s+\frac{1}{2}p_it$
Zygote	A_iA_j	p_jt	$\frac{1}{2}s+p_jt$	$\frac{1}{2}s+p_it$	$\frac{1}{2}s+\frac{1}{2}(p_i+p_j)t$
Genotype	A_jA_j	0	0	$\frac{1}{2}s+p_jt$	$\frac{1}{4}s+\frac{1}{2}p_jt$
	A_iA_k	p_kt	p_kt	0	$\frac{1}{2}p_kt$
	A_jA_k	0	0	p_kt	$\frac{1}{2}p_kt$

for alleles A_i, A_j and A_k with respective frequencies p_i, p_j, and p_k in the pollen gene pool; $t = 1 - s$ is the outcrossing rate. The 'unknown' column is for when the megagametophyte is not assayed. It is $\frac{1}{2}$ the second plus $\frac{1}{2}$ the third column: alternative gametes are assumed to be derived from heterozygous parents with equal probability.

This model assumes there is no mating to relatives, although it can be modified to describe the effect of consanguineous matings at a single-locus (Ritland 1984). Gametophyte selection is allowed if megagametophytes are assayed; otherwise, no selection is assumed. With more than one locus under certain conditions (discussed below), single-family estimation obviates the need to assume homogeneity of pollen gene frequencies and outcrossing rate among families; otherwise, homogeneity must be assumed.

Assay and inference of maternal parentage

Conifers

Progeny (seeds or seedlings) are electrophoretically assayed for both the megagametophyte allele and the zygote genotype; progeny are grouped by common female parent into families. If all megagametophytes of a family are allele A_i, the maternal parent is inferred as genotype A_iA_i; a mixture of megagametophytes A_i and A_j infers A_iA_j. For family size M, the probability of inference of homozygous parentage, given the parent is actually heterozygous, is $(\frac{1}{2})^{M-1}$. A family size of 8 thus seems adequate.

Angiosperms

If zygotes only are assayed, the zygotic segregation pattern of a family at a given locus can allow inference of the most likely maternal parent genotype (Brown and Allard 1970); outbreeding populations require a family size of at least 10 while inbreeding populations [$t < 0.1$] require a family size of at least 2–3. Although the multilocus segregation pattern can allow inference of maternal parentage with greater certainty relative to consideration of single loci only (Ritland and Jain 1981), the small gain of information is not worth the extra programming effort. Smaller progenies may be used if a more complicated estimation procedure that considers parentage in probability is used (e.g. Clegg et al. 1978); for the single-locus model, optimal progeny size is from 4 to 8, and progenies as small as 2 give information (Ritland 1986).

Codification of data

For each observed zygote, we need to specify possible pollen-derived alleles (obtained by either selfing or outcrossing), and the probabilitie(s) of obtaining the pollen allele(s), given the parent has selfed. Thus each observed zygote generates a data pair or a data triplet. For each locus and individual, this encoding of data is given by the following table,

Maternal parent					
Genotype:		A_iA_i	A_iA_j		
Megagametophyte:		A_i	A_i	A_j	Unknown
	A_iA_i	$(i, 1)$	$(i, \frac{1}{2})$	–	$(i, 0, \frac{1}{4})$
Zygote	A_iA_j	$(j, 0)$	$(j, \frac{1}{2})$	$(i, \frac{1}{2})$	$(j, i, \frac{1}{2})$
Genotype	A_jA_j	–	–	$(j, \frac{1}{2})$	$(0, j, \frac{1}{4})$
	A_iA_k	$(k, 0)$	$(k, 0)$	–	$(k, 0, 0)$
	A_jA_k	–	–	$(k, 0)$	$(0, k, 0)$
Data symbolized as		(Y, Z)	(Y, Z)	(Y, Z)	(X, Y, Z)

The first number of the data pairs is the pollen allele, denoted Y, and the second number of each pair is the probability of obtaining Y by selfing, denoted Z. For example, for Y at a diallelic locus, $i = 1$, $j = 2$, and k is omitted. If the megagametophyte is not assayed and the parent is heterozygous ('unknown' column), a data triplet is needed, to specify both possible pollen alleles. For these (X, Y, Z) triplets, X is the pollen allele given A_i was the megagametophyte, Y is the pollen allele given A_j was the megagametophyte, and Z is the probability of obtaining a randomly chosen X or Y by selfing (allele zero denotes an 'impossible' allele of zero frequency).

The multilocus likelihood equation

From these single locus considerations, it is straightforward to obtain the multilocus likelihood of data. For the j-th locus in

individual i the encoded data is denoted as (Y_{ij}, Z_{ij}) for megagametophyte data or (X_{ij}, Y_{ij}, Z_{ij}) otherwise. Denote p_{jk} as the frequency in the outcrossing pollen pool of allele k at locus j with n_j alleles. Since p_{jk} is the probability of paternity by allele k at locus j given an outcrossing event, and since, under the assumption of mixed mating, outcrossing occurs at all loci, the multilocus likelihood of progeny i given it was outcrossed is the product over independent loci

$$T_i = \prod_{j=1}^{n} p_{jk}, \quad k = Y_{ij} \tag{1}$$

for megagametophyte data. Otherwise, if the data is a triplet (X_{ij}, Y_{ij}, Z_{ij}) at locus j, then in eq. (1),

$$p_{jk} = \tfrac{1}{2} p_{jl} + \tfrac{1}{2} p_{jm}, \quad l = X_{ij} \text{ and } m = Y_{ij}.$$

Likewise, since selfing occurs at all loci with mixed mating, the multilocus likelihood of progeny i given it was selfed is

$$S_i = \prod_{j=1}^{n} Z_{ij}. \tag{2}$$

The multilocus likelihood of progeny i, given the population-wide outcrossing rate is t, is

$$L_i = t\,T_i + (1-t)\,S_i. \tag{3}$$

Assuming independent sampling of families, the likelihood of the sample of N individuals is

$$\prod_{i=1}^{N} L_i. \tag{4}$$

Maximum likelihood estimation

Maximum likelihood estimates of t and p_{jk} parameters are the values of t and the p_{jk} that maximize (4). Two recursion methods for finding these values are given here. Each method has advantages and is expected to find estimates that maximize (4), within their respective constraints upon estimates. For either method, first find the scores (Rao 1952) for each individual, which requires prior values of t and p_{jk} (guesses on the first iteration).

For individual i, the score for t is $d \ln L_i/dt$.

$$S_i^t = (T_i - S_i)/L_i. \tag{5}$$

For individual i, there is a score S_{ijk}^p for most p_{jk}. First, zero these scores S_{ijk}^p for $k = 1, n_j - 1$ and $j = 1, n$ (there is no score for the last allele at each locus). Then for each locus j in individual i add or subtract to scores as follows:

(1) If the data is a pair (Y_{ij}, Z_{ij}),

if $Y_{ij} \neq n_j$, then for $k = Y_{ij}$, $S_{ijk}^p = S_{ijk}^p + t\,T_i/(p_{jk}\,L_i)$, (6a)

else if $Y_{ij} = n_j$, then for $k = 1, n_j - 1$,

$$S_{ijk}^p = S_{ijk}^p - t\,T_i/(p_{jk}\,L_i). \tag{6b}$$

(2) If the data is a triplet (X_{ij}, Y_{ij}, Z_{ij}), then for $l = Y_{ij}$,

if $X_{ij} \neq n_j$, then for $k = X_{ij}$,

$$S_{ijk}^p = S_{ijk}^p + t\,T_i/((p_{jk} + p_{jl})\,L_i), \tag{7a}$$

else if $X_{ij} = n_j$, then for $k = 1, n_j - 1$,

$$S_{ijk}^p = S_{ijk}^p - t\,T_i/((p_{jk} + p_{jl})\,L_i), \tag{7b}$$

then do (7a) and (7b) again with Y_{ij} substituted for X_{ij} and $l = X_{ij}$.

After finding the scores for each individual, new t and p_{jk} are found through the recursions

$$t' = t + \sum_{i=1}^{N} S_i^t / I^t \tag{R1}$$

$$p'_{jk} = p_{jk} + \sum_{i=1}^{N} S_{ijk}^p / I_{jk}^p - \sum_{i=1}^{N} \sum_{m=1}^{n_j - 1} S_{ijk}^p / I_{jkm}^p,$$

for $k = 1, n_j - 1$ and $j = 1, n$, $\quad (m \neq k)$, (R2)

where the 'information indices' I are computed as follows. If the 'counting' method of Ceppellini et al. (1955) and Smith (1957) (equivalent to the E–M method) is to be used,

$$\begin{aligned} I^t &= N/[t\,(1-t)] \\ I_{jk}^p &= N\,t'/[p_{jk}\,(1-p_{jk})] \\ I_{jkm}^p &= N\,t'/[p_{jk}\,p_{jm}] \quad (m \neq k) \end{aligned} \tag{8}$$

(note the t prime, obtained from (R1); N is number of individuals assayed). If the single-variable Newton-Raphson (Fisher scoring) method is to be used,

$$\begin{aligned} I^t &= \sum_{i=1}^{N} (S_i^t)^2 \\ I_{jk}^p &= \sum_{i=1}^{N} (S_{ijk}^p)^2 \\ I_{jkm}^p &= \sum_{i=1}^{N} (S_{ijk}^p\,S_{ijm}^p) \quad (m \neq k). \end{aligned} \tag{9}$$

Maximum likelihood estimates are obtained by iterating the above, using t' and the p'_{jk} as new prior values, until (R1) and (R2) converge. Note that for a diallelic locus, the recursion for p_j does not use I_{jkm}^p. The frequency of the last allele at a locus is always 1 minus the frequency of all other alleles.

The algebraic equality of form (8) to counting equations was shown by Ritland (1986) for the single-locus case. The counting method is also equivalent to the 'expectation-maximization' (EM) method of Dempster et al. (1977) and Cheliak et al. (1983). Equation (8) is the expected information per observation when the underlying variables (selfs, pollen alleles) are directly observed, whereas, Eq. (9) is the expected information per observation when some data are not directly observable (incomplete), as is generally the case with mating system data. Since the information values of (9) are less than those of (8), the counting-EM method (8) a gives slower change per iteration than the Newton-Raphson method (9) and hence converges at a slower rate in expectation; the tradeoff is that counting-EM recursions are generally more stable as they never jump outside of the (0, 1) interval. However, this restriction of $\hat{t} < 1.0$ is not statistically valid for outbred populations or small data sets, which can give $\hat{t} > 1.0$ because of incomplete data (or negative assortative mating). Restricting $\hat{t} < 1.0$ can underestimate both mean and variance of $\hat{t}$ among families. Thus, optimally, one might use counting methods (8) for gene frequencies and the Newton-Raphson method (9) for outcrossing rate.

Information for multilocus population estimates

To compare the statistical efficiency of alternative models, computer programs found the information (expected values of second derivatives of the likelihood function, Kendall and Stuart 1979) for t and the p_{jk} under both conifer and angiosperm models. Finding these expected values required enumeration of all possible multilocus genotypes and calculation of their frequency assuming multilocus inbreeding equilibrium. If the number of possible multilocus parent genotypes was greater than 250, a specified number of parent-progeny pairs were randomly picked, because the number of possible parent-progeny combinations became impossibly large. The added information provided by megagametophytic

segregation was evaluated by comparing the conifer model information to the angiosperm model, as given by Ritland and Jain (1981), and to the information given by the models of Green et al. (1980) and Shaw and Allard (1982a), methods that estimate a multilocus t by counting observed outcrosses (and do not include megagametophytic information). An observed outcross occurs when a non-maternal allele is observed in a progeny; its probability P was computed, from which information was found as $P/t^2(1-P)$.

The information per individual for outcrossing rate t or pollen gene frequency p, in relation to number of loci used in the estimate, is for $t=0.9$ and $p=0.25$ at all diallelic loci,

Model	No. of loci					
	1	2	3	5	7	10
Conifer t	0.50	1.12	1.82	3.38	5.0	7.2
Angiosp. t	0.45	0.99	1.60	3.00	4.5	6.4
Observed t	0.27	0.58	0.94	1.82	2.9	4.9
Conifer p	4.37	4.40	4.44	4.51	4.57	4.65
Angiosp. p	3.65	3.67	3.69	3.74	3.78	3.84

Random parent-progeny pairs were used for the 7 and 10 locus calculations. The theoretically attainable information with infinitely many loci is 11.11 for t and 5.93 for p, for all models. The conifer model provides about 10% more information for t and 20% more information for p, relative to the angiosperm model. (At $p=0.5$ there is no difference in information provided by the two models; however, $p=0.25$ or 0.75 is more often observed than $p=0.5$ at electrophoretic loci). The most increase of information as a function of loci is between one and two loci, emphasizing that efficient multilocus models should be used when there are few polymorphic loci. The 'observed outcross' model gives a reduced amount of information roughly equivalent to ignoring 1–2 loci; this loss of information is relatively large with 2–5 loci.

Calculations at other parameter values indicated little difference between the information given by all models when outcrossing was 10%; at this low rate, 3 polymorphic loci gave nearly the information expected with infinite loci. For triallelic loci and high outcrossing rates, there was a rough doubling in the information provided by all models. However, overall, the conifer model did not give any impressive increase of information for *population* estimates of t or p_{jk}, relative to the angiosperm model.

Information for multilocus family estimates

The greatest value of multilocus models lies in their increased degrees of freedom. Whereas t and p for each family are not jointly estimable with single-locus data, with two or more loci, female t and p are estimable for each family (except if at all but one diallelic locus, $p=0.5$ with heterozygous parentage). By contrast, the observed outcross model cannot estimate a family p using its likelihood equation because its reduction of data is too great.

The above procedure can be used to estimate t and p for each family by simply performing recursions (R1) and (R2) for each family. However, maternal genotype should be known, family size should be relatively large (>30), and estimates do not always converge if parents are heterozygous at several loci. The female outcrossing rate is estimated.

The among-family variance of t under the null hypothesis of no true among-family variation for t is found by inverting the information matrix for each parent genotype, and taking the expectation of these family variance-covariance matrices by weighting each with the frequencies of the corresponding parents expected under multilocus inbreeding equilibrium. The inverse of its diagonal elements gave information values here. If $t=0.9$ with no among-family variance, and $p=0.25$ at all diallelic loci, the information for t and p per individual, as a function of number of loci, is

Model	No. of loci			
	2	3	4	5
Conifer family t	0.130	0.422	0.885	1.502
Angiosp. family t	0.065	0.233	0.524	0.943
Conifer family p	1.211	2.513	3.356	3.846
Angiosp. family p	0.671	1.597	2.278	2.695

There is no information with one locus. Apparently 4–5 loci are needed for good family estimation of t and p, as information is quite low for 2–3 loci. The decreased information for t calculated on a family basis (compare to previous table) is caused by a heterogeneity of information values among families. The conifer model gives twice as much information about these family t's and p's relative to the angiosperm model, suggesting that megagametophytic data is appropriate for more elaborate estimation procedures such as family estimation of t and p.