Theor. Appl. Genet. 62, 113–120 (1982)

THEORETICAL AND APPLIED GENETICS

Estimation of Outcrossing Rates in Douglas-fir Using Isozyme Markers

D. V. Shaw and R. W. Allard
Department of Genetics, University of California, Davis, Calif. (USA)

Summary. Seeds produced under open-pollination were collected from eight natural stands and a plus-tree seed orchard of Douglas-fir. These seeds were germinated and both diploid embryos and haploid gametophytes were analyzed by starch-gel electrophoresis. Eleven variable loci were resolved for both kinds of tissue and used as genetic markers for estimating outcrossing rates. Estimates made with single-locus and multilocus methods both indicated that the proportion of viable embryos resulting from outcrossing is about 0.90 for the natural stands, and for the seed orchard. Comparison of single-locus and multilocus estimates of outcrossing rates indicated that little or no inbreeding other than selfing occurred. Estimated outcrossing rates were higher for seeds from the upper portion of the crown than for seeds from the lower crown. It was also found that some trees selfed at a much higher rate than other trees.

Key words: Douglas-fir – Outcrossing rates – Isozymes – Inbreeding – Open-pollination

Introduction

In forest tree breeding open-pollinated seeds are often used to produce seedlings both for regeneration of commercial stands and as materials in studies designed to provide information used in predicting genetic gain under selection. The mating system determines the amount of assortative or disassortative mating that takes place during the formation of open-pollinated progenies and thus the degree of relatedness among the offspring within such progenies. In this paper we report the results of studies, based on electrophoretically revealed marker loci, of the mating system of Douglas-fir (*Pseudotsuga menziesii* var. 'menziesii') in natural stands and in a clonal plus-tree seed orchard.

The closest form of inbreeding possible in plants is self-fertilization and the proportion of progeny resulting from self-fertilization (s) and from outcrossing ($t = 1 - s$) each generation are the parameters most commonly used to describe the mating system. However in natural populations of forest trees most of the seeds shed from a single tree travel short distances. This tendency for the offspring of a single maternal parent to fall in close proximity to the maternal tree can lead to clusters of relatives (family structure) with natural stands. Subsequent matings among these relatives (e.g. between parent and offspring, full-sibs, half-sibs) may result in significant inbreeding in natural stands of forest trees (Libby 1976) and this inbreeding is also a component of the mating system.

Among the many possible effects of departures from random mating two are of particular importance to the forest geneticist using open-pollinated progenies. First, inbreeding depression for economically or adaptively important characters, such as height, volume, survival, and seed set, is common in forest tree species (earlier studies reviewed by Franklin 1970). Sorensen and Miles (1974) have presented evidence that the average first-year height growth for selfed Douglas-fir seedlings was 18 percent lower than for cross-pollinated seedlings and Rehfeldt (1978) indicated that such depression may be as great as 30 percent in subsequent years. Second, in experiments designed to estimate components of variance it is commonly assumed that the individuals within open-pollinated progenies are half-sibs. However, when inbreeding occurs, the degree of relatedness among individuals within such progenies will be greater than that of half-sibs (Squillace 1974), thus resulting in biased estimates of variance components.

Observations on pollination biology and on self-fertility lead to conflicting predictions concerning the degree of inbreeding that occurs in Douglas-fir. Although Douglas-fir pollen has been observed to travel substantial distances (Silen 1962), Sorensen (1980) estimated the proportion of self pollen available to individual maternal parents to be 50–60 percent of the total available pollen. Self-fertility is highly variable among individual trees and several studies of this character in Douglas-fir indicate relative self-fertilities ranging from nearly 0 to 94 percent (Orr-Ewing 1954; Sorensen 1969; Piesch and Stettler 1972). Average self-fertility is low and average seed set on selfing is about 10 percent of the obtained with cross pollination. However, an occasional tree is highly self-fertile.

Observations on pollination biology and relative self-fertility suggest that substantial self-fertilization may oc-

0040-5752/82/0062/0113/$ 1.60

Table 1. Location, sample size, age and height data for the eight natural stands sampled. Sample size data are given for the Jefferson seed orchard

		Location		Sample size Number of		Stand age (approx.)	Stand height (m)
		Latitude	Longitude	Open-pollinated families	Total seeds assayed		
Springfield	1	44° 18′	122° 32′	24	274	35	14
	3	44° 5′	122° 43′	25	299	25	16
	4	44° 5′	122° 40′	23	232	100	30
	5	44° 17′	122° 34′	25	288	40	18
Longview	1	46° 9′	122° 32′	18	167	60	26
	3	46° 12′	122° 38′	25	320	35	12
	4	46° 13′	122° 46′	25	289	70	26
	5	46° 11′	122° 33′	20	244	60	24
Jefferson[a]				42	727		

[a] The Jefferson seed orchard contains several ramets of each of many clonal genotypes. Seventy-six ramets, representing 42 distinct genotypes, were sampled

cur in open-pollinated progenies. However such observations do not provide quantitative measures of the actual proportions of selfed and outcrossed offspring in samples of viable embryos. Two estimates of these proportions have been reported for Douglas-fir. Sorensen (1973), using data on seed set for selfed, crossed, and open-pollinated progenies, estimated the proportion of viable selfed offspring as 7 percent. Rehfeldt (1978), using an albino seedling character as a genetic marker, estimated the proportion of viable selfed offspring to be 5 percent for Rocky Mountain Douglas-fir (*P. menziesii* var. 'glauca').

Mating system parameters in our study were estimated using alleles at electrophoretically detectable loci as genetic markers. The two principal advantages of this method are (Brown and Allard 1970): 1) electrophoretic loci are usually codominant, permitting the genotypes of each individual to be identified without progeny testing; this more complete classification than with dominant markers allows use of more efficient methods of estimation, and; 2) each individual can be scored for several marker loci with relative ease, thus allowing estimates of mating system parameters based on large numbers of both families and offspring. In addition, comparison of outcrossing rates estimated with single-locus and multiple-locus methods yields information about inbreeding other than selfing.

Samples and Electrophoretic Methods

Open-pollinated progenies were collected in September, 1978, from eight stands of naturally regenerated trees, four from each of two breeding zones located near Springfield, Oregon, and Longview, Washington. Location, age, height, the number of progenies sampled, and the number of offspring assayed by electrophoresis are given for each stand in Table 1. In 1978 a high proportion of Douglas-fir trees produced cones and many trees produced heavy cone crops. Because Douglas-fir trees growing in dense stands rarely produce cones, most of the collections were made along roadsides or along natural breaks in the canopy. Four transects radiating from an origin were sampled in each stand; collections were made from at least 10 trees within a 100 meter radius of this origin and the remaining collections were made within 400 meters of the origin.

Seed collections were also made in 1978 from the Springfield low-elevation block of the Weyerhaeuser Jefferson seed orchard. This orchard contains grafts from plus-trees selected from stands similar in age, location, and altitude to those of the Springfield natural stands. As no pollen was artifically introduced into this orchard in 1978, the samples collected are open-pollinated progenies.

In Douglas-fir male strobili occur mainly in the lower part of the crown, whereas female cones are often distributed throughout the crown (Allen and Owens 1972). It follows that a higher proportion of self pollen may be available to the ovules located in the lower crown than to ovules in the upper crown. When possible, seeds were collected separately from both upper and lower parts of the crown of single trees. When the distribution of female cones in the crown did not allow this stratification, a single sample was taken from the center of the available distribution.

Eleven loci were resolved with starch-gel electrophoresis using methods provided by M. Thompson Conkle[1] (personal communication): Glutamate-oxaloacetate transaminase (Got/2 loci/2.6.1.1), glucose-6-phosphate dehydrogenase (G6PD/1 locus/1.1.1.49), glutamate dehydrogenase (Est/1 locus/1.4.1.2), leucine aminopeptidase (Lap/2 locus/3.4.1.1), phosphoglucomutase (Pgm/2 loci/2.7.5.1), phosphoglucoisomerase (Pgi/1 locus/5.3.1.9). Descriptions of the formal genetics of eight of these loci were provided by Richard Morris (unpublished data). The remaining three loci (To, Pgm-2, Pgi)

1 Dr. M. Thompson Conkle, USDA, Forest Service, Pacific Southwest Forest and Range Experiment Station. A description of these methods is currently in preparation for publication with Paul Hodgkiss

show segregation patterns consistent with single-locus inheritance. No deviations from independent assortment were observed for any pair of loci; thus none of the marker loci used in this study appear to be tightly linked.

Data and Estimation Procedures

Seeds of coniferous forest trees contain (1) a haploid gametophyte which produces female gametes of like haploid genotype and (2) a diploid embryo resulting from the fertilization of the female gamete with a male gamete from either the same tree (self-fertilization) or a different tree (outcross). The genotypes of the gametophyte – and thus the maternal gamete – and the embryo can be determined by electrophoretic assay. Consequently, the genotype of each contributed pollen gamete can also be deduced (Shaw and Allard 1981).

Characterization of the mating system requires identification of maternal genotype and estimation of maternal genotypic frequencies, estimation of allele frequencies in the outcross pollen pool, and estimation of the proportions of offspring resulting from self-fertilization and outcrossing. Brown and Allard (1970) presented a two-step maximum likelihood estimator of mating system parameters using single electrophoretic marker loci. Clegg et al. (1978) later provided a one-step estimation procedure. In both of these methods maternal genotypes are inferred from diploid progeny arrays. The method used for single-locus estimation of outcrossing rates in our study is that of Clegg et al. (1978), modified to take advantage of the additional information available from embryo-gametophyte pairs (Appendix 1).

It is possible to derive a single-locus estimator of outcrossing rates that accommodates any number of alleles. If, however, many alleles are infrequent, large numbers of the observational classes are left empty and these empty classes cause difficulties in both estimation and tests of significance. In our study the policy adopted for single-locus estimation was to reduce all data to a form suitable for analysis with a two-allele model. Traditionally such reductions have taken the most frequent allele as one class and all remaining alleles have been pooled to form the second class (a synthetic allele). However this method does not necessarily provide minimum variance estimates of outcrossing rates and, to increase efficiency, we made estimates using all possible combinations of alleles into two classes.

The assumptions made in deriving single-locus estimators based on electrophoretic marker loci have been discussed by Brown and Allard (1970). Shaw et al. (1981) discussed possible violations of the assumptions of single-locus models in natural populations of predominantly outcrossing plants. In our study the assumption that is perhaps most likely to be invalid is that all inbreeding is due to selfing. If other kinds of inbreeding occur they will cause the amount of selfing to be overestimated.

In 1981 Shaw et al. presented an estimator of s and t based on multiple marker loci. The multilocus method distinguishes directly between selfs and outcrosses: it is less affected by invalidity of assumptions than single-locus estimates and, in particular, when inbreeding in addition to selfing occurs, the multilocus estimates of t are expected to be higher than the single-locus estimates. Comparison of single- and multilocus estimates of t thus provide a measure of the total effect of other forms of inbreeding.

Results

Several single-locus estimates of t were made for each population. The precision of a single-locus estimate of the outcrossing rate depends on the frequency of the alleles and the frequency of the maternal genotypes in the sample. Data from marker loci with low levels of allelic variability provide little information about t and estimates made with such loci have large standard errors; consequently marker loci for which the least frequent allelic class occurred at a frequency of less than 0.03 were not used in estimating outcrossing rates. A summary of embryo allele frequencies for the two breeding zones, and for the Jefferson seed orchard, is given in Appendix 2.

Single-locus estimates of outcrossing rates and of allele frequencies in the outcross pollen pool are illustrated in Table 2 with results from Springfield 1 and Springfield 3. These results are typical of those for all eight natural stands. Note that there is considerable

Table 2. Single-locus estimates of the outcrossing rate ($\hat{t}$) and the frequency of the indicated allele in the outcross pollen pool ($\hat{q}$) for the Springfield 1 and Springfield 3 samples

Locus	Allele frequency[a] represented by				
	$\hat{q}$	$\hat{t}$[b]		$\hat{q}$[b]	
	Springfield 1				
To	3	0.91	(0.08)	0.04	(0.02)
Est	1	0.66	(0.10)	0.32	(0.05)
	2	0.58	(0.08)	0.48	(0.05)
	3	0.68	(0.08)	0.20	(0.03)
Lap 1	1	0.93	(0.08)	0.53	(0.03)
	2	0.89	(0.10)	0.22	(0.03)
	3	0.88	(0.10)	0.26	(0.03)
Pgm 1	1	0.95	(0.16)	0.13	(0.03)
	2	0.81	(0.13)	0.81	(0.03)
	3	0.81	(0.10)	0.06	(0.03)
Pgm 2	2	0.73	(0.11)	0.92	(0.03)
	Springfield 3				
Got 3	1	0.93	(0.13)	0.93	(0.04)
	2	0.97	(0.13)	0.06	(0.06)
Est	1	0.62	(0.16)	0.26	(0.06)
	2	0.63	(0.07)	0.46	(0.04)
	3	0.72	(0.10)	0.29	(0.03)
Lap 1	1	1.09	(0.09)	0.49	(0.05)
	2	0.81	(0.13)	0.22	(0.04)
	3	0.94	(0.07)	0.32	(0.03)
Lap 2	1	0.70	(0.10)	0.02	(0.04)
	2	0.68	(0.10)	0.94	(0.04)
Pgm 1	1	1.00	(0.06)	0.12	(0.06)
	2	0.96	(0.06)	0.79	(0.02)
	3	0.88	(0.10)	0.07	(0.02)
Pgm 2	2	0.92	(0.11)	0.97	(0.06)
Pgi	1	0.95	(0.11)	0.05	(0.02)

[a] The estimated frequency of the allele listed in this column is q. All other alleles present at this locus have been combined to form a synthetic allele (see text)

[b] Standard errors in parentheses

Table 3. Chi-square test for heterogeneity of outcrossing rates estimated by single-locus methods

		Heterogeneity test including estimates made using the Esterase locus			Heterogeneity test excluding estimates made using the Esterase locus		
		d.f.	$\chi^2_{(d.f.)}$	P	d.f.	$\chi^2_{(d.f.)}$	P
Springfield	1	11	21.03	$0.01<P<0.05$	8	4.63	$0.70<P<0.80$
	3	13	25.10	$0.01<P<0.05$	10	13.00	$0.20<P<0.30$
	4	10	16.05	$0.05<P<0.10$	7	2.78	$0.90<P<1.00$
	5	15	43.48	$P<0.001$	12	21.51	$0.01<P<0.05$
Longview	1	14	32.29	$P<0.01$	11	10.09	$0.50<P<0.70$
	3	10	13.20	$0.20<P<0.30$	7	1.03	$0.90<P<1.00$
	4	13	50.98	$P<0.001$	10	8.20	$0.50<P<0.70$
	5	14	15.7	$0.30<P<0.50$	11	11.80	$0.30<P<0.50$
Jefferson		17	30.1	$0.01<P<0.05$	14	13.73	$0.30<P<0.50$

variation both among estimates made with different marker loci, and among estimates made with different synthetic allele combinations at a single locus. Also note that estimates of t made with the esterase marker are consistently lower than other estimates in both sample populations; this was the case in most other samples as well. Single-locus estimates of t for all samples ranged from 0.33 to 1.12 when estimates based on the esterase marker are included; when this marker locus was excluded the minimum estimate of t is 0.67. Chi-square tests for heterogeneity of single-locus estimates of t within each population (Elandt-Johnson 1971) are given in Table 3. Estimates of t are significantly heterogeneous ($P<0.05$) for six of the nine sample populations. When estimates made with the esterase marker are excluded from this test only one of the samples yields significant heterogeneity. Estimates of t made with the esterase marker are apparently affected by factors other than the mating system and they have consequently been excluded from subsequent analyses of single-locus estimates.

Table 4. Mean estimates of outcrossing rates ($\hat{t}$) calculated by weighting the individual single-locus estimates by the inverse of estimated maximum-likelihood variances, and multilocus estimates of outcrossing rates ($\hat{t}_m$)

		$\hat{t}$ [a]		$\hat{t}_m$ [b]	
Springfield	1	0.88	(0.03)	0.92	(0.04)
	3	0.90	(0.03)	0.89	(0.03)
	4	0.90	(0.04)	0.86	(0.03)
	5	0.93	(0.03)	0.91	(0.03)
Longview	1	0.90	(0.03)	0.87	(0.04)
	3	0.91	(0.03)	0.92	(0.03)
	4	0.90	(0.03)	0.87	(0.03)
	5	0.92	(0.02)	0.96	(0.03)
Natural stand Average		0.91	(0.03)	0.90	(0.03)
Jefferson		0.94	(0.01)	0.91	(0.02)

[a] Values in parentheses are standard errors for the mean of single-locus estimates of t

[b] Standard errors in parentheses

Mean single-locus estimates of t, and their standard errors, are given for each population in Table 4. The means range from 0.88 to 0.93 for natural stands, with a grand mean of 0.91. The mean single-locus estimate of t for the Jefferson seed orchard sample is 0.94. To compensate for the differences in precision among estimates made with different marker loci or different allelic combinations, means were calculated by weighting individual estimates of t by the inverse of their estimated maximum likelihood variances.

Multilocus estimates of t are given in Table 4. These estimates range from 0.86 to 0.96 for natural stands with a grand mean of 0.90. The multilocus estimate of t for the Jefferson seed orchard sample is 0.91.

Inferences about inbreeding other than selfing can be made from two comparisons. First, single-locus estimates of t are expected to be biased downward by any inbreeding in addition to selfing; thus the mean of such single-locus estimates is expected to be lower than the multilocus estimate of t when mating among relatives occurs. Comparison of values in Table 4 indicates that this expectation is generally not true. The grand mean for single-locus estimates of t (0.91) is, in fact, slightly higher than the grand mean for multilocus estimates of t (0.90), although not significantly so. Second, because the Jefferson orchard consists of grafted clones planted in randomly chosen locations within the orchard, it is expected to be free of family structure. Thus single-locus estimates of t made using seed orchard offspring will reflect only inbreeding due to selfing while estimates made using natural stand offspring may include the effects of other kinds of inbreeding as well. The mean single-locus estimate of t for the Jefferson sample is 0.94.

This value is outside the range of means estimated for Springfield natural stands (0.88–0.93), a result which suggests that some inbreeding in addition to selfing occurs in natural stands.

Both of the above comparisons suffer from the large standard errors associated with single- and multilocus estimates of t. The precision of estimation can be improved by increasing the number of loci and/or increasing the number and the size of families assayed. Although the present data do not exclude the possibility that inbreeding other than selfing occurs in natural stands of Douglas-fir, they indicate that most of the inbreeding observed in samples from the natural stands is due to self-fertilization.

Variability in outcrossing rates among crown levels was investigated in six of the natural stands studied, and in the Jefferson seed orchard. Multilocus estimates of t made using seeds collected from the upper third and from the lower third of the crown are given in Table 5. Estimates of t were larger for the upper-crown samples in five of the six natural stands. The mean estimate of t for upper-crown samples is 0.93 while that for lower-crown samples is 0.86. A paired Student's t test indicated significant differences ($P < 0.05$) between estimated outcrossing rates made with upper-and lower-crown samples. Estimates of outcrossing rates were also larger for the upper-crown sample of the Jefferson seed orchard, although the difference was small.

Variability in outcrossing rates among families was investigated in Jefferson seed orchard. An initial screening of individual outcrossing rates (based on multilocus estimates of t) indicated considerable variability among families, even though sample sizes for the estimates were small. The results of this original survey were used to identify the five families with the highest apparent outcrossing rate and the five families with the lowest apparent outcrossing rate; additional embryo-gametophyte pairs were then assayed in each of these families. (One of the highly outcrossed families was later excluded due to a lack of additional seed.) Multilocus estimates of t based on the larger sample sizes are given in Table 6. None of the estimates of t made for "highly outcrossed" families differ significantly from $t = 1.0$. Conversely, all "highly selfed" families yielded estimates of t that were significantly smaller ($P < 0.05$) than $t = 1.0$. Similarly, the outcrossing rate estimated over all "highly outcrossed" families, $t = 1.03$, does not differ significantly from $t = 1.0$ whereas the estimated outcrossing rate over all "highly selfed" families, $t = 0.69$, is significantly smaller ($P < 0.01$) than $t = 1.0$.

Table 5. Multilocus estimates of outcrossing ($\hat{t}_m$) for samples obtained from upper and lower portion of the crown

		$\hat{t}_m$[a]				Number of families sampled
		Upper crown		Lower crown		
Springfield	1	0.89	(0.06)	0.86	(0.06)	10
	3	0.84	(0.06)	0.89	(0.07)	10
	5	1.02	(0.05)	0.89	(0.06)	10
Longview	3	0.95	(0.05)	0.86	(0.06)	11
	4	0.99	(0.06)	0.85	(0.07)	7
	5	0.89	(0.06)	0.80	(0.07)	8
Total for Natural Stands		0.93	(0.02)	0.86	(0.03)	56
Jefferson		0.92	(0.03)	0.89	(0.04)	39

[a] Standard errors in parentheses

Table 6. Multilocus estimates of outcrossing ($\hat{t}_m$) for four highly outcrossed and five highly selfed families from the Jefferson seed orchard

Highly outcrossed families			Highly selfed families		
Clone number	$\hat{t}_m$[a]		Clone number	$\hat{t}_m$[a]	
481	1.13	(0.11)	435	0.60	(0.12)
483	0.95	(0.04)	441	0.56	(0.08)
500	1.02	(0.08)	464	0.72	(0.13)
513	1.03	(0.07)	480	0.79	(0.09)
			487	0.58	(0.11)
Over all families	1.03	(0.04)		0.69	(0.04)

[a] Standard errors in parentheses

Discussion

The eight natural stand studied differed from each other in topography, age and size of trees, and stand density, but were relatively homogenous within stands for these factors. Despite these differences, the estimated rate of outcrossing for all stands was consistently about 90 percent. The estimated rate of outcrossing for the Jefferson seed orchard was also about 90 percent. Such results suggest that 90 percent outcrossing and 10 percent selfing may be typical of Douglas-fir throughout its range in the Pacific Northwest.

In addition, our results indicate that outcrossing rates differ between upper and lower crown levels and also from tree to tree. The former result is consistent with the observed distribution of male and female strobili in the crown of Douglas-fir. The latter result is consistent with the high individual variability in self-fertility observed in Douglas-fir. All maternal trees may be exposed to high levels of self pollen but, because of variable self-fertility, it is likely that only a few set appreciable amounts of viable selfed seed. This hypothesis raises

the interesting possibility that most of the selfed offspring originate from a few highly self-fertile maternal trees. If this is the case then self fertility will affect family performance in open-pollinated genetic tests.

Single-locus outcrossing estimates varied widely from locus-to-locus even within single populations. Such variability is common for estimates of t made for many predominantly outcrossing plants (Brown and Allard 1970; Brown et al. 1975). In our study all estimates for a single population were made from the same set of embryos so that variation in the actual proportions of selfed and outcrossed offspring did not contribute to the observed within-population variability in single-locus estimates of t. Thus the variability must be due either to random variation and/or to violations of the assumptions intrinsic to the estimation procedure. Shaw and Brown (1981) have shown that the optimum number of loci for precise estimates of mating system parameters is large in predominantly outcrossed plants.

The present data set indicates that most of the inbreeding observed in Douglas-fir results from selfing, but it does not exclude the possibility that other types of consanguineous mating may occur. Inbreeding in addition to selfing may be more important in artifically regenerated stands than in natural populations, e.g. stands composed of the progeny resulting from seed-tree regeneration programs are expected to contain higher proportions of half- and full-sibs than natural populations.

Populations in which 10 percent of the offspring result from selfing are expected – due to inbreeding depression – to have lower yields than outcrossed populations. In addition, average genetic correlations will be greater within progenies containing some selfed off spring than in outcrossed progenies. Whether these potential consequences of inbreeding are actually realized in genetic tests of open-pollinated progenies and in breeding operations depends on the survival of the selfs under operational and test conditions. Our experimental results indicate the need for experiments designed to estimate the proportions of selfed offspring surviving to various ages under such conditions.

Acknowledgment

This work was supported in part by National Science Foundation Grant DEB 78-02046. We also thank Dr. Rex McCullough and the Weyerhaeuser Company for their cooperation and support during several phases of this project.

Appendix I

Estimation of Mating System Parameters Using Single Marker Loci

Clegg et al. (1978) presented a maximum likelihood method for estimating mating system parameters based on marker locus data from diploid family arrays. The method used in our study is similar to that of Clegg et al. but has been modified to accommodate the additional information from analysis of embryo-gametophyte pairs.

Each diploid member of a family will receive at least one gamete from the maternal parent of that family (the maternal ovule). The second gamete may also originate from the maternal parent as the result of a self-fertilization, or it may originate from a second parent as the result of an outcross. The distribution of embryo genotypes within a given sample of families thus depends on the frequency of maternal genotypes in the sample, the frequencies of alleles in the outcross pollen pool, and the proportion of embryos resulting from self-fertilization and outcrossing. A joint maximum likelihood model was formulated for the simultaneous estimation of the parameters necessary for describing this mixed mating model.

The general form of the likelihood equation (after Elandt-Johnson 1971) is

$$L(\theta) = \frac{N!}{0_1! \ldots 0_k!} \prod_{i=1}^{k} (m_i)^{0_i},$$

in which θ represents the vector of parameters (θ_i) that require estimation, m_i, is the expected proportion of progeny in the i^{th} observational class, O_i is the observed number of progeny in the i^{th} observational class, and N is the sum of the O_i (or the total number of progeny). Table A 1.1 gives the necessary information for estimation of mating system parameters using a single marker locus with two alleles, when information on both embryos and gametophytes is available. The expected proportions of progeny for the eight observational classes in Table A1.1 are formed using a vector (θ) of 8 parameters. P_1–P_4 represent the inferred frequencies for the four possible maternal genotype-maternal gametic contribution combinations, s is the proportion of embryos resulting from selfing, t (= 1 – s) is the outcrossing rate, and p and q are the frequencies of the A_1 and A_2 alleles respectively in the outcross pollen pool. Because t = 1 – s, q = 1 – p and $P_4 = 1 - P_1 - P_2 - P_3$ only 5 parameters require estimation.

Maximum-likelihood estimates of the 5 independent mating system parameters (t, q and P_1–P_3) are obtained by finding the maximum of the log-likelihood function. This was

Table A 1. Model for estimating mating system parameters using a single marker locus with two alleles (A_1, A_2). $P_1 - P_4$ are the frequencies of the four possible maternal genotype-maternal gametic contribution combinations, respectively, and p and q are the frequencies of the A_1 and A_2 alleles in the outcross pollen pool. $0_1 - 0_8$ are the observed numbers of embryos in the eight possible maternal genotype-maternal gamete-pollen gamete classes

Maternal parent	Maternal gamete	Pollen	Conditional expectations (m_i)	Observations (0_i)
A_1A_1	A_1	A_1	$P_1(s+tp)$	0_1
		A_2	P_1 (tq)	0_2
A_1A_2	A_1	A_1	P_2 (1/2 s + tp)	0_3
		A_2	P_2 (1/2 s + tq)	0_4
	A_2	A_2	P_3 (1/2 s + tq)	0_5
		A_1	P_3 (1/2 s + tp)	0_6
A_2A_2	A_2	A_2	P_4 (s + tq)	0_7
		A_1	P_4 (tp)	0_8

Appendix II

Table A 2. Allele frequencies for eleven enzyme loci estimated with progeny samples from two breeding zones

	LOCUS									
	Got 1			Got 3			G6pd			
	0	1	2	1	2	3	1	4		
Springfield	[a]	0.99	0.01	0.04	0.95	0.01	0.97[b]	0.03		
Longview	0.01	0.99	–	0.10	0.89	0.01	0.96[b]	0.04		
Jefferson		1.0		0.05	0.94	0.01	0.96[b]	0.04		
	Adh		To			Est				
	1	2	2	3	4	1	2	3	4	
Springfield	0.99	0.01	0.96	0.04	–	0.24	0.50	0.23	0.03	
Longview	1.00	[a]	0.92	0.03	0.05	0.21	0.49	0.29	0.01	
Jefferson	0.99	0.01	0.96	0.04	[a]	0.14	0.58	0.27	0.01	
	Lap 1				Lap 2			Pgm 1		
	1	2	3	N[c]	1	2	3	1	2	3
Springfield	0.51	0.25	0.23	0.01	0.04	0.93	0.03	0.13	0.81	0.06
Longview	0.47	0.25	0.27	0.01	0.03	0.91	0.06	0.11	0.83	0.06
Jefferson	0.43	0.27	0.29	0.01	0.09	0.88	0.03	0.10	0.82	0.08
	Pgm 2					Pgi				
	0	1	2	3	4	1	2	3		
Springfield	[a]	0.04	0.93	0.03	[a]	0.08	0.91	0.01		
Longview	–	0.03	0.94	0.02	0.01	0.05	0.93	0.02		
Jefferson	[a]	0.06	0.92	0.02	–	0.10	0.90	[a]		

[a] Indicates that alleles of this type were found in the sample but that rounding to two decimal places has resulted in a frequency of < 0.01

[b] Allele 1 at the G6pd locus represents the pooled frequency for three alleles discernible with analysis of gametophytic tissue, but not consistently discernible in embryo analysis

[c] Null alleles are not discernible in diploid embryo tissue; these values were estimated from analysis of maternal gametophytic tissue

accomplished by solving the first derivative of the log-likelihood function to an approximately zero solution using numerical methods (Elandt-Johnson 1971). A minimum estimate of the sampling variance for each parameter was also obtained using the Cramèr-Rao inequality (Elandt-Johnson 1971).

Two important differences exist between this model and models that use only information from diploid embryos. First, the distribution of diploid embryo genotypes within a family depends on both the outcrossing rate and the frequencies of alleles in the outcross pollen pool (in addition to the genotype of the maternal parent). Thus the inference of maternal genotypes from diploid embryo arrays requires conditional probabilities dependent on both t and q. When data on embryo-gametophyte pairs are available inference of maternal genotype is based only on the array of haploid gametophytes obtained from each maternal parent (Morris and Spieth 1977). Estimates of maternal genotypic frequencies made with this model are statistically independent of t and q. Secondly, note in Table A1 that the observation of female gametic contribution allows the distribution of progeny from heterozygous mothers into four observation classes. When data on gametophytes are not available the maternal gametic contribution is unknown and only three classes are discernible. This improvement in resolution removes the assumption of equal segregation of alleles in heterozygotes and renders the model more robust to any natural forces that might cause segregation distortion at the marker locus (association with selected loci, by chance or design).

Literature

Allen, G.S.; Owens, J.N. (1972): The life history of Douglas-fir. Can. For. Serv. Publ.

Brown, A.H.D. (1975): Efficient designs for estimation of genetic parameters in plant populations. Biometrics **31,** 145–160

Brown, A.H.D.; Allard, R.W. (1970): Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. Genetics **66,** 135–145

Brown, A.H.D.; Matheson, A.C.; Eldridge, K.G. (1975): Estimation of the mating system in *Eucalyptus obliqua* L'Hérit. using allozyme polymorphisms. Austr. J. Bot. **23,** 931–949

Clegg, M.T.; Kahler, A.L.; Allard, R.W. (1978): Estimation of life cycle components of selection in an experimental plant population. Genetics **89,** 765–792

Elandt-Johnson, R.C. (1971): Probability models and statistical methods in genetics. New York: Wiley

Franklin, E.C. (1970): A survey of mutant forms and inbreeding depression in species of the family Pinaceae. USDA For. Serv. Res. Paper S.E. **61,** 21

Libby, W.J. (1976): Introductory remarks: In: Proc. of the IUFRO joint meeting of working parties on Population and Ecological Genetics, Breeding Theory, Biochemical Genetics and Progeny Testing, pp. 1–9. 33610-Cestes, France: I.N.R.A. Laboratoire des eonifères

Morris, R.S.; Spieth, P.T. (1978): Sampling strategies for using female gametophytes to estimate heterozygosity in conifers. Theor. Appl. Genet. **51,** 217–222

Orr-Ewing, A.L. (1954): Inbreeding experiments with Douglas-fir. For. Chron. **38,** 7–16

Piesch, R.F.; Stettler, R.F. (1971): The detection of good selfers for haploid induction in Douglas-fir. Silvae Genet. **20,** 144–148

Rehfeldt, G.E. (1978): The genetic structure of populations of Douglas-fir (*Pseudotsuga menziessii* var 'glauca') as reflected by its wind pollinated progenies. Silvae Genet. **27,** 49–52

Shaw, D.V.; Allard, R.W. (1981): Analysis of mating system parameters and population structure in Douglas-fir using single-locus and multilocus methods. In: Proceedings of symposium on Isozymes of North American Forest Trees and Insects, July 27, 1979, Berkeley, Calif. Gen. Tech. Rep. PSW-48, p. 18–22. Pacific Southwest Forest and Range Exp. Stn., Berkeley, Calif.: For. Serv. USDA

Shaw, D.V.; Kahler, A.L.; Allard, R.W. (1981): A multilocus estimator of mating system parameters in plant populations. Proc. Nat. Acad. Sci. (USA) **78,** 1298–1302

Shaw, D.V.; Brown, A.H.D. (1982): The optimum number of loci for estimating outcrossing rates in plant populations. Theor. Appl. Genet. (in press)

Silen, R.R. (1962): Pollen dispersal considerations for Douglas-fir. J. Fores. **60,** 790–795

Sorensen, F.G. (1969): Estimates of self-fertility in coastal Douglas-fir from inbreeding studies. Am. Nat. **103,** 389–398

Sorensen, F.G. (1973): Frequency of seedlings from natural self-fertilization in coastal Douglas-fir. Silvae Genet. **22,** 20–24

Sorensen, F.G. (1980): The role of polyembryony and embryo viability in the genetic systems of conifers. (in prep.)

Sorensen, F.G.; Miles, R.S. (1974): Self pollination effects on Douglas-fir and Ponderosa Pine seedlings. Silvae Genet. **23,** 135–138

Squillace, A.E. (1974): Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genet. **23,** 149–156

Received December 12, 1981
Communicated by W.J. Libby

Dr. D.V. Shaw
Dr. R.W. Allard
Department of Genetics
University of California
Davis Calif. 95616 (USA)