

Estimation of mating system parameters when outcrossing events are correlated

(mixed-mating model/mating type/outcrossing rate/allozyme/multiple paternity)

DANIEL J. SCHOEN* AND MICHAEL T. CLEGG†

Departments of Botany and Molecular and Population Genetics, University of Georgia, Athens, GA 30602

Communicated by R. W. Allard, April 6, 1984

ABSTRACT Many plant species have mating systems characterized by a mixture of self-fertilization and outcrossing. Statistical estimation of the outcrossing rate has relied on a model of the mating process that assumes that successive outcross events within a family arise from independent draws of pollen from the total population of male plants. Although this assumption is likely to be most appropriate for wind-pollinated plants, it is not appropriate in certain insect-pollinated plants. An alternative model is developed that assumes that successive outcross events within a family involve pollen drawn from a single male parent. The estimation of the parameters that index this model is outlined and a procedure for calculating the variances of the parameter estimates is presented. Monte Carlo simulations of the sampling processes assumed by each model are also presented. The simulations show that application of the incorrect estimation model to data can lead to a large bias in parameter estimates.

The mating system of a population determines how genetic information is transferred from one generation to the next (1). Moreover, the recombinational potential of the population may be strongly influenced by the mating system (2). Accurate information on the mating system in plant populations cannot be obtained through direct observation. This is due, in part, to difficulties inherent in tracing the flow of pollen among and within individuals. Plant population biologists have instead relied on methods to infer the mating system retrospectively from progeny genotypes. These methods use marker loci and statistical estimation (3–18). A common practice in mating system studies that use allozymic markers is to assay several progeny from single maternal (seed) parents and estimate the maternal parent frequencies from the progeny genotype distributions using, as a basis, some underlying statistical model of the mating process. This logistically simple method of sampling the population avoids problems associated with age- and/or tissue-specific expression of allozyme phenotypes.

In plant population genetics, one model of mating, the mixed-mating model, has held a central position in both theoretical and experimental investigations (7, 12). There are several reasons for choosing the mixed-mating model rather than assuming random mating to describe the pattern of gene transmission in plant populations. First, inbreeding is common in many plant species and the random-mating model is, therefore, inappropriate. Second, the mixed-mating model is simple and requires the estimation of only one additional parameter, the outcrossing rate. Third, the assumptions of the mixed-mating model are reasonable descriptions of the reproductive biology of many species. Despite the wide utility of the mixed-mating model, it may fail to properly describe the reproductive biology of certain insect-pollinated plants. Below we document this failure and then propose an alterna-

tive to the mixed-mating model that is appropriate when successive outcrossing events involve pollen drawn from a single paternal parent.

The Mixed-Mating Model

The mixed-mating model makes several basic assumptions (7, 19). First, it assumes that mating events are of two types—random outcrossing with probability t and self-fertilization with probability $s = 1 - t$. It further assumes that the outcrossing rate is uniform for all maternal parents and that pollen genotypes are distributed uniformly over all eggs. Violation of the assumption of a uniform pollen genotype distribution may occur for a variety of reasons. For example, there may be temporal or spatial heterogeneity in the distribution of genotypes in the population (20). Alternatively, there may be correlation among the outcrossed pollen types received by individual maternal parents (i.e., pollen is not drawn randomly from the entire pollen pool). In probabilistic terms, the assumption of random pollen sampling may be stated as $I_p = \text{Prob}(G_{ij}^{k+1} | G_{ij}^k) = tf(A_i A_j)$, where G_{ij}^k and G_{ij}^{k+1} are the events that the k th and $(k + 1)$ st eggs of a single maternal parent are fertilized by pollen from a male of genotype $A_i A_j$ and where $f(A_i A_j)$ denotes the relative frequency of males of genotype $A_i A_j$. When mating events are correlated, I_p will be greater than indicated by this equation. With complete correlation among mating events, $I_p = t$. In insect-pollinated plants, correlation is likely to be a factor when the families assayed for marker genotype are collected from one, or a limited number of, multiseeded fruit. Specifically, if the pollen deposited on a stigma derives from one or a few previously visited plants, many of the progeny produced as the result of the outcrossed mating events will be full sibs. Spatial structuring of genotypes in the population may also contribute to correlation among mating events if it modifies the fertilization probabilities by the various pollen types (21, 22).

Failure of the Mixed-Mating Model

In the common and ivy-leaved morning glories, *Ipomoea purpurea* and *Ipomoea hederacea*, two species of insect-pollinated plants bearing one to many multiseeded fruits, the mixed-mating model has proven to be inappropriate as the basis for estimating mating system parameters (23). The difficulty arises in the stage of the estimation procedure where maternal genotypes are inferred on the basis of the distribution of family genotype arrays (12). For example, Ennos (23) reports significant excesses of heterozygous progeny in the families of inferred heterozygous maternal parents [where expected numbers of such progeny are given by the transition probabilities associated with the mixed-mating model

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, Quebec H3A 1B1, Canada.

†Present address: Department of Botany and Plant Sciences, University of California, Riverside, CA 92521.

(12)]. In this study, the progeny were sampled from a limited number of capsules on each plant, a situation where correlation among mating events is expected to influence the estimation procedure (see below).

To investigate further the supposition that correlated mating events cause the mixed-mating model to yield incorrect estimates of mating system parameters, we carried out computer simulations of the mating process. Family distributions were generated via Monte Carlo methods using a Hewlett-Packard 1000 minicomputer. Two sampling processes were simulated. First, to generate data that conform to the independence assumption, maternal genotype frequencies were calculated assuming inbreeding equilibrium as

$$\begin{aligned} f(A_1A_1) &= p^2 + pqF \\ f(A_1A_2) &= 2pq(1 - F) \\ f(A_2A_2) &= q^2 + pqF, \end{aligned}$$

where p is the frequency of allele A_1 in the population, $q = 1 - p$ is the frequency of allele A_2 , and $F = (1 - t)/(1 + t)$ (24). The algorithm began by selecting a maternal parent at random. Next, five eggs were drawn at random from the maternal parent. Whether each egg was outcrossed or self-fertilized was determined at random based on the probability t or $1 - t$, respectively. If outcrossed, pollen of type A_1 or A_2 was selected at random with probabilities p and q , respectively. If selfed, pollen was selected at random from the maternal parent. Two hundred maternal parents and their progeny were sampled, yielding 1000 genotypes. Second, to generate family distributions resulting from complete correlation of outcrossed mating events, the procedure described above was modified so that, for a particular maternal parent, a single randomly chosen paternal parent served as the pollen donor for all outcrossed eggs. The families produced using the second method were otherwise generated in the same manner as in the first. We refer below to these two methods of generating family data as the "independence" and "correlation" methods.

After a data set was generated, the maximum likelihood procedure of Clegg *et al.* (12) was used to compute estimates of the maternal parent frequencies (M_j s), the frequency of allele A_1 in the pollen pool (p), and the outcrossing rate (t). The outcrossing rate ($t = 0.8$) used in the simulations was chosen to represent the moderately high values reported for many insect-pollinated plants. Simulation-estimation trials were carried out 1000 times for family data generated via both methods described above.

Analysis of the results obtained from the family data simu-

lated assuming correlated mating events reveals a significant excess of heterozygous progeny in families of heterozygous parents (Table 1). This is consistent with the morning glory data (23). The cause of this error can be attributed to matings of the type $A_1A_1 \times A_2A_2$ (and the reciprocal mating), which yield only heterozygous progeny. Under the transition probabilities of the mixed-mating model, such arrays have a high likelihood of arising from heterozygous maternal parents. Because only one-half of the progeny of heterozygous mothers are expected to be heterozygous under the mixed-mating model assumptions (12), a poor fit between the observed and expected numbers of progeny is obtained. Moreover, the estimates of mating system parameters are severely biased, have extremely large associated mean squares, and are distributed in a manner that radically departs from normality (Fig. 1 and Table 2). For example, the mean estimate of t in this case is 37% below the population value, and the variances associated with the mean estimates of p and t are between one and two orders of magnitude larger than those obtained when family data are generated via the independence method.

Estimates of mating system parameters when data are generated via the correlation method, but with different parameter values, yield results that are qualitatively similar to those summarized above, with the following two exceptions: (i) when allelic frequencies are asymmetrical, the estimate of p is significantly larger than the population value (data not shown) and (ii) when the outcrossing rate is low ($t < 0.5$) the severity of the bias and magnitude of the sample variances associated with the estimated parameters becomes less pronounced (data not shown). The results of the simulations suggest that misapplication of the traditional mixed-mating model may lead to grossly incorrect estimates of mating system parameters. This indicates the need for a modified procedure to analyze family data in populations in which correlated mating events are likely to be common. We develop one such model in the next section.

Estimation of Mating System Parameters and Adult Genotypic Frequencies When Mating Events Are Correlated—The One Pollen Parent Model

The method that we outline here is a modification of the procedure given in appendix A of Clegg *et al.* (12). The sampling scheme involves collecting seeds from single maternal parents and assaying the seeds (or seedlings) for their genotype at the marker locus. We assume a diallelic marker with codominant expression of alleles. Each seedling is classified by family and genotype, resulting in a vector $N_y = [n_{1y}, n_{2y}, n_{3y}]$

Table 1. Expected and observed numbers of progeny genotype i derived from maternal genotype j for simulated data sets produced under the correlation assumption and analyzed by the mixed-mating model

Maternal genotype	Progeny genotype		
	A_1A_1 ($i = 1$)	A_1A_2 ($i = 2$)	A_2A_2 ($i = 3$)
A_1A_1 ($j = 1$)			
Expected	166.67	111.11	0
Observed	176.65 (42.15)	80.90 (62.34)	0
A_1A_2 ($j = 2$)			
Expected	111.11	222.22	111.11
Observed	100.69 (40.15)	284.62 (54.25)	109.51 (46.54)
A_2A_2 ($j = 3$)			
Expected	0	111.11	166.67
Observed	0	78.77 (71.38)	168.87 (42.30)

Expected numbers of progeny in each cell (E_{ij}) were calculated as $E_{ij} = T \cdot M_j \cdot \Theta_{ij}$, where T = total number of progeny sampled, M_j = frequency of maternal genotype j , and Θ_{ij} = probability that maternal genotype j gives rise to progeny genotype i [see appendix A of Clegg *et al.* (12)]. Results represent means and SDs (in parentheses) of observed distributions (e_{ij}) obtained in 1000 simulation-estimation trials where $p = 0.5$ and $t = 0.8$.

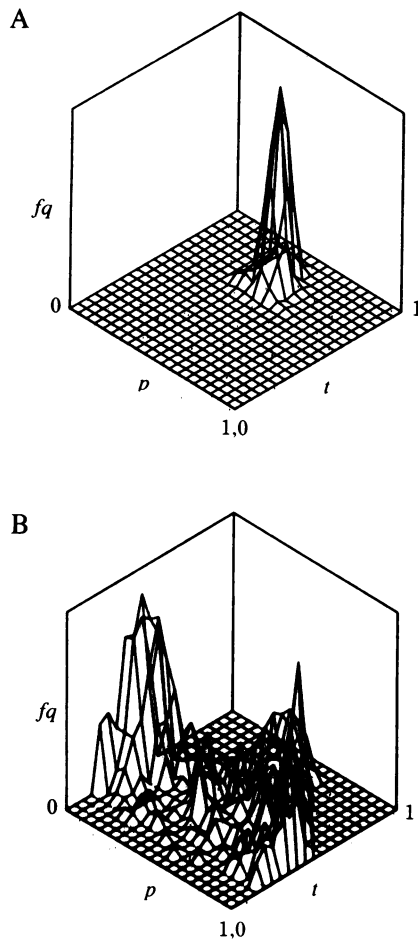


FIG. 1. Joint distribution of outcrossing rate and pollen pool allele frequency estimates (t, p) for 1000 simulation-estimation trials involving the mixed-mating model. Population values of $t = 0.8$ and $p = 0.5$ were assumed in generating family genotype distributions. Family data were generated by the independence method (A) and the correlation method (B).

of the number of progeny of each genotype in the y th family. The data set for the population sample constitutes a $3 \times m$ array of progeny genotypes observed within families, where m is the number of families sampled. For simplicity, the number of seeds sampled per family is assumed to be constant over families and represented by

$$R = \sum_{i=1}^3 n_i.$$

Table 2. Estimates of maternal frequencies and mating system parameters based on the mixed mating model: Family data were generated by the independence and correlation methods

Population value and parameter estimate	Parameter			
	$f(A_1A_1\text{♀})$	$f(A_1A_2\text{♀})$	p	t
Population value	0.444	0.278	0.500	0.800
Parameter estimate				
Independent mating events	0.451 (0.058)	0.276 (0.041)	0.503 (0.039)	0.789 (0.056)
Correlated mating events	0.495 (0.089)	0.258 (0.100)	0.467 (0.334)	0.507 (0.157)

Population values given are those used to generate simulated family distributions. Results represent mean and SDs (in parentheses) based on 1000 simulation-estimation trials.

Each family member receives a gamete from a common maternal parent. The second gamete is derived from the same maternal parent through self-fertilization, with probability s , or from a different parent through outcrossing, with probability $t = 1 - s$. We further assume that the genotypes of the potential pollen donors are distributed uniformly over maternal parents (i.e., there is no spatial or temporal heterogeneity in the distribution of pollen donor genotypes). The departure we make from the traditional mixed-mating model concerns the assumption of independent mating events. We assume that the conditional probability $I_p = t$. Consequently, all families are produced by one of nine different mating types. The conditional probabilities of observing the three progeny genotypes given the mating type are functions of t (Table 3). We denote them as Θ_{ijk} ($i, j, k = 1, 2, 3$), where Θ_{ijk} is the probability that the mating pair, consisting of maternal parent of genotype j and paternal parent of genotype k , gives rise to a progeny of genotype i . Below we refer to the combination of j th maternal and k th paternal genotypes as the jk th mating type. The probability of observing a particular progeny vector for the y th family given the jk th mating type is

$$P(N_y|jk) = R! \prod_{i=1}^3 \frac{\Theta_{ijk}^{n_{iy}}}{n_{iy}!}, \quad y = 1, 2, \dots, m.$$

We let M_j and P_k represent probabilities of drawing the j th maternal and k th paternal genotypes. The joint probability of choosing a particular mating type and observing the progeny vector N_y is

$$P(N_y, jk) = R! M_j P_k \prod_{i=1}^3 \frac{\Theta_{ijk}^{n_{iy}}}{n_{iy}!}.$$

The conditional probability of the jk th mating type, given the progeny vector N_y , is

$$P(jk|N_y) = \frac{M_j P_k \prod_{i=1}^3 \frac{\Theta_{ijk}^{n_{iy}}}{n_{iy}!}}{\sum_{j=1}^3 \sum_{k=1}^3 M_j P_k \prod_{i=1}^3 \frac{\Theta_{ijk}^{n_{iy}}}{n_{iy}!}} = \pi_{jky}.$$

The M_j s and P_k s, however, are not known, and so a provisional estimate must be used. New estimates of M_j and P_k are obtained as

$$\hat{M}_j = \left(\sum_{r=1}^m \sum_{k=1}^3 \pi_{jkr} \right) / m \quad \text{and} \quad \hat{P}_k = \left(\sum_{r=1}^m \sum_{j=1}^3 \pi_{jkr} \right) / m.$$

This procedure is based on the gene counting method of Cappelini *et al.* (25).

The second part of the procedure involves the estimation of t . The progeny distributions are distributed into a 9×3 array in which the ijk th element, e_{ijk} , represents the total

Table 3. Conditional probabilities of observing progeny genotype i given mating type jk ($i, j, k = 1, 2, 3$) based on the one pollen parent model

Mating type (♀ × ♂)	Progeny genotype		
	A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	$s + t = 1$	0	0
$A_1A_1 \times A_1A_2$	$s + t/2$	$t/2$	0
$A_1A_1 \times A_2A_2$	s	t	0
$A_1A_2 \times A_1A_1$	$(s/4) + t/2$	$(s + t)/2 = 1/2$	$s/4$
$A_1A_2 \times A_1A_2$	$(s + t)/4 = 1/4$	$(s + t)/2 = 1/2$	$(s + t)/4 = 1/4$
$A_1A_2 \times A_2A_2$	$s/4$	$(s + t)/2 = 1/2$	$s/4 + t/2$
$A_2A_2 \times A_1A_1$	0	t	s
$A_2A_2 \times A_1A_2$	0	$t/2$	$s + t/2$
$A_2A_2 \times A_2A_2$	0	0	$s + t = 1$

number of progeny of genotype i whose mating type is jk . The maternal and paternal parents of the progeny arrays are not known, and so the classification proceeds as

$$e_{ijk} = \sum_{r=1}^m n_{ir} \pi_{jkr}$$

The maximum likelihood estimate of t is obtained by finding the value \hat{t} that maximizes the log likelihood equation,

$$\ln L = C + \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 e_{ijk} (\ln \Theta_{ijk}),$$

where C is a constant. In our work, the Newton-Raphson procedure was employed to find the solution of this cubic equation.

Once t is obtained, we repeat the procedure for estimating maternal and paternal frequencies and the conditional probabilities π_{ijk} . This results in a new array $[e_{ijk}]$ and, hence, a new maximum likelihood estimate of t . The two-stage process is repeated until convergence is achieved at both stages. Because the number of iterations required to reach convergence at both stages will depend on the provisional estimates of the M_j s, P_k s, and t , we use the following initial values,

$$M_j = P_k = \left(\sum_{r=1}^m n_{jr} \right) / \left(\sum_{r=1}^m \sum_{j=1}^3 n_{jr} \right)$$

and

$$t = M_2 / [4M_1M_3 + M_2(1 - M_2)].$$

A test for goodness-of-fit of the estimated parameters to the data can be performed if P_k and M_j are assumed to be equal for $j = k = 1, 2, 3$. The test statistic is

$$\chi^2 = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \frac{(E_{ijk} - e_{ijk})^2}{E_{ijk}},$$

where $E_{ijk} = TM_jP_k\Theta_{ijk}$ and T is the total number of progeny assayed—i.e.,

$$T = m \cdot R.$$

The test statistic is distributed as χ squared with 5 degrees of freedom.

To find the lower-bound variance estimates of the parameter estimates \hat{t} , \hat{M}_j , and \hat{P}_k , it is necessary to consider the complete likelihood expression for the outcome of a given experiment (the collection of family vectors). This likelihood can be formulated by noting that the total probability of observing the y th family vector $N_y = (n_{1y}, n_{2y}, n_{3y})$ is

$$P(N_y) = R! \sum_{j=1}^3 \sum_{k=1}^3 M_j P_k \prod_{i=1}^3 \frac{(\Theta_{ijk})^{n_{iy}}}{n_{iy}!}.$$

Therefore, that the probability of observing the collection of family vectors N_y ($y = 1, 2, \dots, m$) is

$$P(N_1, N_2, \dots, N_m) = C \cdot \left[\prod_{y=1}^m \sum_{j=1}^3 \sum_{k=1}^3 M_j P_k \prod_{i=1}^3 (\Theta_{ijk})^{n_{iy}} \right],$$

where C is a constant representing the products of the multinomial coefficients from the probability expressions of the m families. The log likelihood is thus

$$L(\Theta|N_1, \dots, N_m) = C + \sum_{y=1}^m \ln \left[\sum_{j=1}^3 \sum_{k=1}^3 M_j P_k \prod_{i=1}^3 (\Theta_{ijk})^{n_{iy}} \right].$$

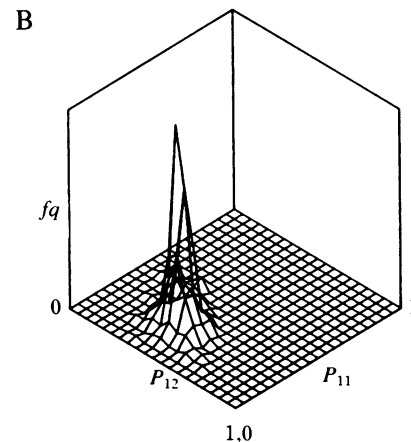
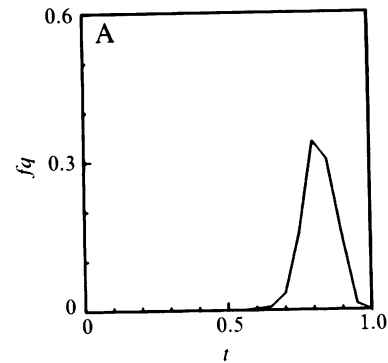


FIG. 2. Distribution of estimated parameters for 1000 simulation-estimation trials in which family data were generated by the correlation method and estimation was based on the one pollen parent model (see Table 4 for population values of parameters). (A) Distribution of outcrossing rate estimates (t). (B) Distribution of paternal genotype frequencies [$P_{11} = f(A_1A_1\delta)$, $P_{12} = f(A_1A_2\delta)$].

Double differentiation of this expression with respect to the individual parameters yields the information matrix for a given experiment, from which the variance estimates may be obtained through numerical methods of matrix inversion (26).

When the one pollen parent model is applied to family data produced using the correlation method, greatly improved estimates of the mating system parameters are obtained (Fig. 2 and Table 4). In a separate paper, we will address the question of how well the variances derived via this procedure agree with the variances obtained by repeated application of the estimation algorithm to family data generated by Monte Carlo simulation.

Discussion

The independence and correlation assumptions represent extremes of a continuum. Nevertheless, the primary advantage of the two models is that they allow a complete and simple specification of how progeny genotypes are formed. Usually one of the models will provide a more accurate description of mating than the other and will therefore yield estimates that are closer to the true population values. In deciding which model to apply, several biological and statistical factors should be considered. These factors include the mode of pollination—whether wind- or insect-mediated, the number of ovules per ovary, and the number of flowers that are open simultaneously. For example, when progeny for mating system estimation are collected from insect-pollinated plants bearing few multiseeded fruit, or from several fruits that de-

Table 4. Estimates of maternal frequencies, paternal frequencies, and outcrossing rates based on the one parent pollen model: Family data generated using the correlation method

Population value and parameter estimate	Parameter				
	$M_1 = f(A_1A_1\varphi)$	$M_2 = f(A_1A_2\varphi)$	$P_2 = f(A_1A_1\delta)$	$P_2 = f(A_1A_2\delta)$	t
Population value	0.278	0.444	0.278	0.444	0.800
Parameter estimate	0.280 (0.051)	0.435 (0.053)	0.291 (0.072)	0.436 (0.077)	0.794 (0.055)

Population values given are those used to generate simulated family distributions. Results represent mean and SDs (in parentheses) based on 1000 simulation-estimation trials.

velop from flowers that are open simultaneously, the opportunity for correlated mating events is greatly increased. In contrast, seed families collected from wind-pollinated plants with single-seeded fruits that develop from flowers open at different times should often conform to the independence assumption. It is of interest to note that the more successful applications of the mixed-mating model have been with wind-pollinated plants.

χ^2 tests associated with the goodness-of-fit of the estimated parameters to the data can be calculated as part of the estimation procedure associated with each model. As illustrated above, when the independence assumption is violated, estimation results based on the mixed-mating model are characterized by excess heterozygous progeny in families of heterozygous maternal parents. On the other hand, when the correlation assumption is violated, estimation based on the one pollen parent model leads to a bias toward inferring heterozygous parentage when the true parentage is homozygous. In fact, the estimate of $f(A_1A_2\delta)$ ($=P_2$) approaches unity in most cases where the outcrossing rate is even moderately high ($t > 0.5$) (Fig. 3 and unpublished data). These statistical indicators can be used to provide additional guidance in making an appropriate choice between the two models.

The notion that the seeds of any one fruit (or plant) may have multiple paternal parents has figured prominently in recent discussions of mate selection in plants (27–29). It has been suggested that female sporophytic control, and competition between pollen tubes, may influence male reproductive success. Moreover, it has been reported that progeny that result from fertilizations involving more successful male

gametes have increased vigor (30, 31). However, because estimates of the degree of multiple paternity are scarce, it is difficult to assess the opportunities for selection of mates through intergametophytic competition. Besides providing a basis for the estimation of mating system parameters, application of the mixed-mating and one pollen parent models may help to provide preliminary estimates of the degree of multiple paternity in certain plant species.

We thank Dr. Jonathan Arnold for discussing this problem with us and Dr. Kermit Ritland for comments on the manuscript. This research was funded by National Science Foundation Grant BSR-8304796.

1. Wright, S. (1921) *Genetics* 6, 124–143.
2. Allard, R. W. (1975) *Genetics* 79, 115–126.
3. Jones, D. F. (1916) *Science* 43, 509–510.
4. Fyfe, J. L. & Bailey, N. T. J. (1951) *J. Agric. Sci.* 41, 371–378.
5. Harding, J. & Tucker, C. L. (1964) *Heredity* 19, 369–381.
6. Vasek, F. C. (1964) *Evolution* 18, 213–218.
7. Brown, A. H. D. & Allard, R. W. (1970) *Genetics* 66, 113–145.
8. Harding, J. (1970) *Evolution* 24, 120–127.
9. Allard, R. W., Babbal, G. R. & Clegg, M. T. (1972) *Proc. Natl. Acad. Sci. USA* 69, 3043–3048.
10. Brown, A. H. D., Matheson, A. C. & Eldridge, K. G. (1975) *Aust. J. Bot.* 23, 931–949.
11. Brown, A. H. D., Zohary, D. & Nevo, E. (1978) *Heredity* 41, 49–62.
12. Clegg, M. T., Kahler, A. L. & Allard, R. W. (1978) *Genetics* 89, 765–792.
13. Ellstrand, N. C., Torres, A. M. & Levin, D. A. (1978) *Syst. Bot.* 3, 403–407.
14. Rick, C. M., Holle, M. & Thorp, R. W. (1978) *Plant Syst. Evol.* 129, 31–44.
15. Jain, S. K. (1979) *Crop Sci.* 19, 23–26.
16. Ritland, K. & Jain, S. (1981) *Heredity* 47, 35–52.
17. Shaw, D. V., Kahler, A. L. & Allard, R. W. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1298–1302.
18. Schoen, D. J. (1982) *Evolution* 26, 352–360.
19. Hayman, B. I. (1953) *Heredity* 7, 185–192.
20. Ennos, R. A. & Clegg, M. T. (1982) *Heredity* 48, 283–292.
21. Turner, M. E., Stephens, J. C. & Anderson, W. W. (1982) *Proc. Natl. Acad. Sci. USA* 79, 203–207.
22. Ritland, K. (1984) *Genetics* 106, 139–152.
23. Ennos, R. A. (1981) *Genetica* 57, 93–98.
24. Haldane, J. B. S. (1924) *Proc. Cambridge Philos. Soc. Biol. Sci.* 1, 158–163.
25. Ceppellini, R., Siniscalco, M. & Smith, C. A. B. (1955) *Ann. Hum. Genet.* 20, 97–115.
26. Elandt-Johnson, R. (1971) *Probability Models and Statistical Methods in Genetics* (Wiley, New York), p. 306.
27. Janzen, D. H. (1977) *Am. Nat.* 111, 365–371.
28. Willson, M. F. & Burley, N. (1983) *Mate Choice in Plants* (Princeton Univ. Press, Princeton, NJ).
29. Bertin, R. I. (1982) *Am. Nat.* 119, 694–709.
30. Mulcahy, D. L. (1974) *Nature (London)* 249, 491–493.
31. Mulcahy, D. L. (1979) *Science* 206, 20–23.

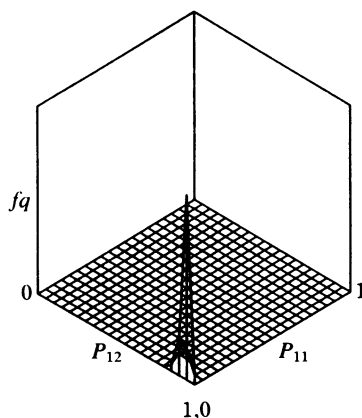


FIG. 3. Joint distribution of paternal frequencies for 1000 simulation-estimation trials in which family data were generated by the independence method and estimation was based on the one pollen parent model [$P_{11} = f(A_1A_1\delta)$, $P_{12} = f(A_1A_2\delta)$]. Population values of parameters are as given in Fig. 2 and Table 4.