Effective Size of Nonrandom Mating Populations

Armando Caballero and William G. Hill

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland

Manuscript received May 31, 1991
Accepted for publication November 27, 1991

ABSTRACT

Nonrandom mating whereby parents are related is expected to cause a reduction in effective population size because their gene frequencies are correlated and this will increase the genetic drift. The published equation for the variance effective size, N_e, which includes the possibility of nonrandom mating, does not take into account such a correlation, however. Further, previous equations to predict effective sizes in populations with partial sib mating are shown to be different, but also incorrect. In this paper, a corrected form of these equations is derived and checked by stochastic simulation. For the case of stable census number, N, and equal progeny distributions for each sex, the equation is

$$N_e = \frac{4N}{2(1-\alpha) + S_k^2(1+3\alpha)},$$

where S_k^2 is the variance of family size and α is the departure from Hardy-Weinberg proportions. For a Poisson distribution of family size ($S_k^2 = 2$), it reduces to $N_e = N/(1+\alpha)$, as when inbreeding is due to selfing. When nonrandom mating occurs because there is a specified system of partial inbreeding every generation, α can be substituted by Wright's F_{IS} statistic, to give the effective size as a function of the proportion of inbred mates.

In most analyses of unstructured outbreeding natural populations mating is assumed to be at random, and in populations subdivided into small demes, random mating is assumed in each deme. Nonrandom mating, however, can be important in some species of plants and animals. Many plants reproduce by self-fertilization to varying degrees [see, for example, reviews by JAIN (1976) and SCHEMSKE and LANDE (1985)]. In certain parasitic Hymenoptera, males hardly disperse at all and mate with their sisters almost as soon as they emerge from the host (ASKEW 1968). SELANDER (1970) found a significant deviation from Hardy-Weinberg proportions for an esterase locus in the house mice of a barn. This implies that the population was reproductively subdivided into family groups, born out by behavioral studies indicating that house mice form small groups that repel outsiders, especially males. In humans and other primate social structures, in contrast, matings between relatives are usually avoided. This is also the policy normally carried out in management of domestic populations. Furthermore, deliberate matings of full-sibs can be useful in increasing fixation probabilities of recessive mutations without the disadvantages of delaying fixation by population subdivision (CABALLERO, KEIGHTLEY and HILL 1991).

The possible effect of nonrandom mating on effective population size (N_e) has been considered in the derivation of the variance effective size, which measures the amount of gene-frequency drift [CROW and MORTON (1955); KIMURA and CROW (1963); corrected for separate sexes by CROW and DENNISTON (1988)]. For stable census number, N, and monoecious diploids or separate sexes with the same progeny distribution for each sex, the effective size has been expressed as

$$N_{eKC} = \frac{4N}{2(1-\alpha) + S_k^2(1+\alpha)} \tag{1}$$

(KIMURA and CROW 1963), where S_k^2 is the variance of the number of gametes (which with no selfing is number of progeny) per parent, V_k, with the Gaussian correction, $S_k^2 = V_k N/(N-1)$, and α is a measure of the departure from Hardy-Weinberg ratios due to inbreeding or other causes, such that the frequency of heterozygotes is $2q(1-q)(1-\alpha)$, where q is the gene frequency in the parental generation. As will be shown, Equation 1 holds when inbreeding occurs due to selfing providing S_k^2 is interpreted as the variance of number of gametes contributed by each individual. By contrast, it does not hold when inbreeding is due to mating of relatives, but not selfing.

Alternative expressions for the effective sizes of populations with partial sib mating and Poisson distribution of family size have been developed by POLLAK (1987, 1988) following a system of recurrent equations derived by WRIGHT (1951). These expressions are summarized by

$$N_{eWP} = \frac{N_{eR}}{1 + (2a-1)F_{IS}} \tag{2}$$

[Pollak (1988), Based on Wright (1951)], where N_{eR} is the effective size with random mating, F_{IS} is the inbreeding coefficient at equilibrium when N is infinite, and a is a constant which depends upon the conditional probabilities of different types of sib mating. For example, with mixed full-sib mating and random mating, $a = 2$ and $N_{eWP} = N_{eR}/(1 + 3F_{IS})$. For comparison with (1), F_{IS} can be approximated by α if N is large, giving

$$N_{eWP} = \frac{N_{eR}}{1 + 3\alpha}. \tag{3}$$

Likewise, with a mix of half-sib mating and random mating where mothers of the half-sibs are again half-sibs, $a = 4$ and $N_{eWP} = N_{eR}/(1 + 7F_{IS})$ or

$$N_{eWP} = \frac{N_{eR}}{1 + 7\alpha}. \tag{4}$$

If Equation 1 held in these cases, it would show that when the expected distribution of progeny numbers is approximately Poisson and, therefore, $S_k^2 = 2$, the predicted variance effective size equals the actual census number irrespective of the magnitude of the departure from Hardy-Weinberg proportions. This departure is, however, expected to cause a reduction in effective size because, when mates are related, there is a covariance of their gene frequencies which could increment the genetic drift in the offspring. Equation 2, or more clearly its derivatives (3) and (4), show that reduction, but no attempt has been made to connect these equations with Kimura and Crow's expression. In fact, as will be shown, these expressions also give incorrect values of the effective size.

In this paper, Equation 1 is corrected for the case when inbreeding is due to mating of relatives without selfing, and the new expression checked by means of stochastic simulation for several situations of unselected populations with nonrandom mating. By a different approach based on inbreeding grounds (Wright 1951, 1969), an equivalent equation is derived to predict effective sizes where a specified proportion of full-sib matings is performed every generation, correcting Pollak's equation. Throughout the paper we shall deal just with the case of stable census number for simplicity. Other assumptions are discrete generations, autosomal inheritance and no correlation between the fertility of a parent and that of its offspring.

EFFECTIVE SIZE WITH A DEPARTURE FROM THE HARDY-WEINBERG RATIOS

Inbreeding due to mating of relatives without selfing. Because mates are likely to be related in a nonrandom mating population, there is a correlation between gene frequency of male and female parents. To take such a correlation into account in the derivation, we use the average frequency of a couple and not, as is usual, that of every individual. The argument is given in terms of males and females, but also applies to monoecious individuals. Then, if in family i, q_{im} and q_{if} are the gene frequencies of male and female parents, respectively, k_i is the number of offspring coming from that family and there are $N/2$ families, the mean gene frequency of the total offspring in the population is following the methodology of Latter (1959) and Hill (1979)]

$$Q = \frac{1}{N}\sum_{i=1}^{N/2}\left[k_i\left(\frac{q_{im} + q_{if}}{2}\right) + \frac{1}{2}\sum_{j=1}^{k_i}(\delta_{ijm} + \delta_{ijf})\right],$$

where the second term in brackets accounts for the random segregation in heterozygotes, and δ_{ijm} (δ_{ijf}) is the difference in frequency between the jth sampled gene and its male (female) parental value q_{im} (q_{if}), where δ_{ijm} (δ_{ijf}) is 0 if the parent is a homozygote or ±1/2 if a heterozygote.

The genetic drift is $\delta_q = Q - q$, where $q = (1/N)\sum_{i=1}^{N/2}(q_{im} + q_{if})$. Thus, the variance in gene-frequency drift, V_{δ_q}, is

$$V_{\delta_q} = \frac{1}{N^2}\left\{\left[\frac{N}{8}V(q_{im} + q_{if})V_k\right] + \frac{N}{4}[(V\delta_{ijm}) + V(\delta_{ijf})]\right\}. \tag{5}$$

The variance of the gene frequency of the parents is

$$V(q_{im}) = V(q_{if}) = q(1 - q)(1 + \alpha_I)/2, \tag{6}$$

α_I being their inbreeding coefficient, and the covariance between their gene frequencies is

$$\mathrm{cov}(q_{im}, q_{if}) = q(1 - q)\alpha_O,$$

where α_O is the inbreeding of the offspring or the coancestry of the parents. Therefore,

$$V(q_{im} + q_{if}) = q(1 - q)(1 + \alpha_I) + 2q(1 - q)\alpha_O. \tag{7}$$

The variance due to segregation, $V(\delta_{ijm}) = V(\delta_{ijf})$, equals the product of the frequency of heterozygotes, $2q(1 - q)(1 - \alpha_I)$ and the variance generated from them, 1/4, that is

$$V(\delta_{ijm}) = V(\delta_{ijf}) = q(1 - q)(1 - \alpha_I)/2. \tag{8}$$

Substituting (7) and (8) into (5)

$$V_{\delta_q} = \frac{1}{8N}q(1 - q)[2(1 - \alpha_I) + V_k(1 + \alpha_I + 2\alpha_O)]. \tag{9}$$

In the ideal population, the drift variance per generation is $V_{\delta_q} = q(1 - q)/2N_e$ [see, *e.g.*, Crow and Kimura (1970), p. 357], and equating this to (9), gives

$$N_{e\alpha} = \frac{4N}{2(1 - \alpha_I) + V_k(1 + \alpha_I + 2\alpha_O)}.$$

In the steady state of constant mating structure, $\alpha_I = \alpha_O = \alpha$ and,

$$N_{e\alpha} = \frac{4N}{2(1-\alpha) + V_k(1+3\alpha)}, \qquad (10)$$

which, for large N when the Gaussian correction for the variance of family size is irrelevant, is a corrected form of Equation 1 with the term $(1 + 3\alpha)$ instead of $(1 + \alpha)$. The additional term (from α_O) comes because KIMURA and CROW (1963) assumed random mating of inbred individuals.

For the case of a Poisson distribution of family size ($S_k^2 = 2$), Equation 10 reduces to the simple form

$$N_{e\alpha} = \frac{N}{1+\alpha}, \qquad (11)$$

which agrees with the case of a monoecious population with a mixture of selfing and random mating and Poisson family size [LI (1976), p. 562; POLLAK (1987, 1988)] but differs from Equations 3 or 4.

Inbreeding due to selfing: A similar derivation to that above may be given for the case when inbreeding results from selfing in the population.

An individual passes two genes to each selfed offspring and one gene to each nonselfed offspring. Therefore, the mean gene frequency in the offspring is

$$Q = \frac{1}{2N}\sum_{i=1}^{N}\left[\sum_{j=1}^{s_i}(2q_i + \delta_{ij1} + \delta_{ij2}) + \sum_{j=1}^{n_i}(q_i + \delta_{ij})\right],$$

where s_i and n_i are the number of selfed and nonselfed offspring of individual i, respectively, and q_i and δ_{ij} are defined as above with the subscript 1 or 2 to denote the two genes passed to each selfed progeny.

The variance in gene-frequency drift, where now $q = (1/N)\sum_{i=1}^{N} q_i$, is

$$V_{\delta_q} = \frac{1}{4N^2}\left\{V\left[\sum_{i=1}^{N} q_i(2s_i + n_i - 2)\right]\right. \\ \left. + V\left[\sum_{i=1}^{N}\left\{\sum_{j=1}^{s_i}(\delta_{ij1} + \delta_{ij2}) + \sum_{j=1}^{n_i}\delta_{ij}\right\}\right]\right\}.$$

Noting that $\sum_{i=1}^{N}(2s_i + n_i) = 2N$ and substituting $V(q_i)$ and $V(\delta)$ from (6) and (8), respectively,

$$V_{\delta_q} = \frac{1}{8N}q(1-q)[2(1-\alpha) + V(2s+n)(1+\alpha)].$$

Denoting the variance of the number of gametes ($V[2s + n]$) as S_k^2, this gives Equation 1, in agreement with KIMURA and CROW (1963) and CROW and DENNISTON (1988). Thus, Equation 1 holds for the case of nonrandom selfing, but not nonrandom pair mating. KIMURA and CROW (1963) and CROW and DENNISTON (1988), however, assumed that with Poisson distribution, the variance of the number of gametes equals its mean, *i.e.*, $S_k^2 = 2$ and from (1), $N_e = N$, irrespective of whether the parents are in Hardy-Weinberg proportions or not. This is not the case if selfing occurs, for if the number of selfed and nonselfed progeny are independently Poisson distributed, $S_k^2 = 2 + 2\beta$, where β is the proportion of selfing. As $\alpha = \beta/(2 - \beta)$ for partial selfing (see, *e.g.*, HEDRICK and COCKERHAM 1986), $\beta = 2\alpha/(1 + \alpha)$ and substituting into (1) yields Equation 11 again, as expected.

TABLE 1

Observed (N_e) and expected (N_{eKC}, N_{eWP} and $N_{e\alpha}$; see text) effective population sizes for populations with size N, random mating (R) or full-sib mating whenever possible (FS)

Population	*NFS*	α	N_e ± S.E.	N_{eKC}	N_{eWP}	$N_{e\alpha}$
$N = 64$						
R	1.0	−0.011	64.0 ± 0.07	64.0	66.2	64.7
FS	15.5	0.159	54.5 ± 0.05	64.0	43.3	55.2
$N = 200$						
R	1.0	−0.005	200.4 ± 0.22	200.1	202.8	200.9
FS	47.8	0.182	170.0 ± 0.46	200.0	129.4	169.2

NFS, number of full-sib matings achieved; α, observed departure from the Hardy-Weinberg proportions.

CHECK OF PREDICTIONS BY SIMULATION

To check these equations, stochastic simulation was carried out for a population subject to nonrandom mating. The simulated population consisted of N diploid individuals, half of each sex, mated in pairs to produce families with an average number of two males and two females following a multinomial distribution of family size for each sex. From the total offspring, parents for the next generation were chosen at random and were either mated at random (scheme R) or by crossing full-sibs whenever possible, otherwise at random (*e.g.*, if two males and one female from a family were sampled to breed, one male was mated with his sister, the other with a random female from a different family) (scheme FS).

Every simulation was run for 30 generations and 300 or 3000 replicates, depending on the population size. Values of α were obtained by including in the simulation a gene with initial frequency 0.5 in Hardy-Weinberg proportions and calculating each generation the relative deviation of observed homozygosity from expectation with the Hardy-Weinberg assumption. These values were averaged for all generations and replicates after generation 10, when an asymptote had been reached. Observed effective sizes were calculated from the change in variance of the gene frequency between generations 11 and 30.

Table 1 shows the observed values of effective size and the predictions by means of Equations 1, 3 and 11 as well as the observed values of α and the number of full-sib matings achieved for two population sizes

TABLE 2

Observed parameters for populations with size N and a fixed proportion of full-sib matings every generation

Population	*NFS*	S_k^2	$S_{km,f}$	α
$N = 64$	0 (0.0)	2.0	−0.022	−0.020
	8 (6.7)	2.1	0.096	0.051
	16 (12.8)	2.3	0.222	0.114
	24 (18.2)	2.6	0.358	0.211
	32 (23.0)	2.9	0.504	0.347
$N = 200$	0 (0.0)	2.0	−0.009	−0.007
	25 (20.8)	2.1	0.106	0.051
	50 (39.5)	2.3	0.225	0.131
	75 (56.5)	2.6	0.357	0.237
	100 (71.1)	2.8	0.499	0.368

NFS, intended number of full-sib matings (the number actually performed is shown in parenthesis); S_k^2, variance of family size; $S_{km,f}$, covariance of the number of male and female progeny; α, departure from the Hardy-Weinberg proportions.

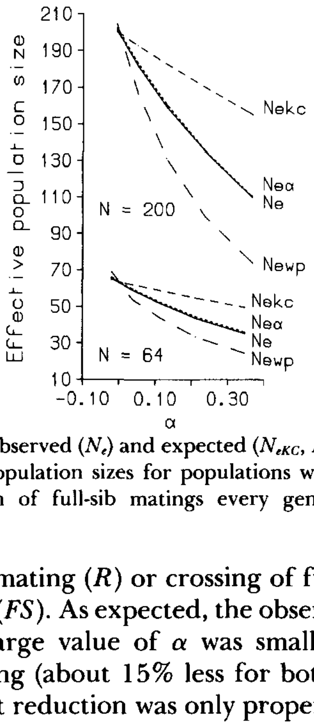

FIGURE 1.—Observed (N_e) and expected (N_{eKC}, N_{eWP} and $N_{e\alpha}$; see text) effective population sizes for populations with size N and a fixed proportion of full-sib matings every generation, plotted against α.

and random mating (R) or crossing of full-sibs whenever possible (FS). As expected, the observed effective size with a large value of α was smaller than with random mating (about 15% less for both population sizes) and that reduction was only properly taken into account by $N_{e\alpha}$.

An alternative scheme, which permitted a variable number of full-sib matings and consisted of choosing the reproductive males at random from the total offspring and crossing a fixed proportion of them with their sisters (if available) and the remainder with non-sisters is shown in Table 2 and Figure 1. The number of full-sib matings actually performed was smaller than intended because sibs sometimes were not available as the distribution of family size was multinomial. S_k^2 increased with the magnitude of α because of the increase in the covariance of the number of male and female progeny (although S_k^2 for the females was less than expected from the multinomial, because of impositions of the mating structure). Values of N_{eWP} were calculated from equation (3) substituting N_{eR} for the effective size with random mating and the appropriate S_k^2.

As clearly seen in Figure 1, $N_{e\alpha}$ (obtained from Equation 10) was a very accurate prediction of the effective size but N_{eKC} was again an overestimation and N_{eWP} an underestimation, especially with greater α. However, both the proportional over- or underestimation and the reduction in the effective size with increasing α were independent of the population size, the proportional reduction being approximately linear with α such that a 10% increase in α yielded about a 9% decrease in effective size.

Finally, a hierarchical scheme in which the number of reproductive males (25) was smaller than that of females (100) to give families of full and half-sibs, was run. This would give $N_{eR} = 80$ with random mating. Males and females were chosen at random for reproduction and mating was between half-sibs whenever possible (avoiding full-sib matings) and otherwise at random. With an average number of 49.7 half-sib matings, the observed effective size was 73.3 ± 0.5. The predicted value of N_{eWP} (Equation 4) was 49.1, whereas the predicted value of $N_{e\alpha}$ was 73.3, calculated from Equation 10 considering means, variances and covariances of family size for each sex separately, following CROW and DENNISTON (1988).

EFFECTIVE SIZE WITH A SPECIFIED PROPORTION OF FULL-SIB MATINGS

The previous discussion has been in terms of the effective size as a measure of the change in variance of the gene frequency, but under the assumed situation of stable census number, inbreeding and variance effective sizes have the same value. An approach in terms of inbreeding is also illuminating and can be used where the departure from the Hardy-Weinberg proportions is due to a specified number of full-sib matings.

Prediction of WRIGHT's F statistics: By allowing the possibility of nonrandom mating in the population, we can distinguish the effects of nonrandom mating from the effects of finite size on the inbreeding. This can be done by considering the whole group of replicates as a subdivided population in which every replicate is a subpopulation and then using WRIGHT's F statistics [WRIGHT (1969), pp. 294–295; HARTL and CLARK (1989), pp. 293–298]. In a subdivided population, F_{IS} is the correlation of uniting gametes relative to gametes drawn at random from within a subpopulation (a replicate) or, as stated above, the inbreeding when the population size is infinite, and is a function of the nonrandom mating in the subpopulation; F_{ST} is the correlation of gametes drawn at random within subpopulations relative to gametes drawn at random from the entire population (all replicates) and is a function of the effective population size; and F_{IT} is the

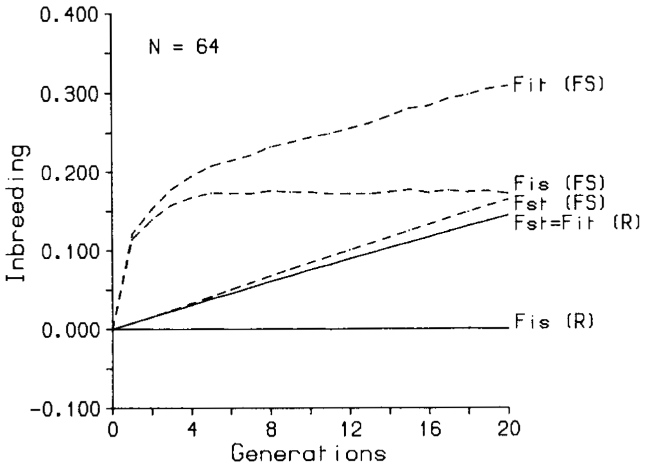

FIGURE 2.—Observed values of F_{IT}, F_{ST} and F_{IS} for a population with size 64 and random mating (*R*) or full-sib mating whenever possible (*FS*).

correlation of uniting gametes relative to gametes drawn at random from the entire population. The relationship between these three types of inbreeding coefficients is

$$(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST}). \qquad (12)$$

In the simulation, values of F_{IT} and F_{ST} were calculated by means of the symmetric matrix of the numerator of WRIGHT's coefficient of additive genetic relationship (or, equivalently, coancestry) and the values of F_{IS} from them by means of Equation 12.

F_{IT}, F_{ST} and F_{IS} values for the case with $N = 64$ and random (*R*) or full-sib mating whenever possible (*FS*) are plotted against generations in Figure 2. For random mating, F_{IS} is zero and, therefore, F_{IT} equals F_{ST}, while in the nonrandom mating scheme F_{IS} increases to an asymptote.

Using a similar derivation to that of WRIGHT (1951; 1969, p. 197), assume a proportion x of matings are between individuals which are full-sibs and let CM_t and CO_t be the coancestry between mated and non-mated individuals in generation t, respectively. The coancestry between two mated sibs in generation t is the probability of two copies of the same gene coming from one of the genes in one of their parents, *i.e.*, 1/4, plus the probability of two copies of the same gene coming from the two genes in one of their parents, *i.e.*, $(1/4)CM_{t-2}$, plus the probability of two copies of the same gene coming from different parents, *i.e.*, $(1/2)CM_{t-1}$. The coancestry between two non-sibs mated in generation t is the coancestry between unmated individuals in generation $t - 1$. Therefore,

$$CM_t = x\left[\frac{1}{4} + \frac{1}{4}\,CM_{t-2} + \frac{1}{2}\,CM_{t-1}\right] + (1 - x)CO_{t-1} \qquad (13)$$

and, analogously,

$$CO_t = y\left[\frac{1}{4} + \frac{1}{4}\,CM_{t-2} + \frac{1}{2}\,CM_{t-1}\right] + (1 - y)CO_{t-1}, \qquad (14)$$

where y is the probability that two individuals are sibs given that they are not mated. Thus, at generation $t + 1$, $F_{IT} = CM_t$ and $F_{ST} = CM_t/(N - 1) + (N - 2)CO_t/(N - 1)$, or, $F_{ST} \approx CO_t$ for large N.

For large populations an approximation for the asymptotic value of F_{IS} can be evaluated. Then, F_{ST} and, therefore, $CO_t \approx 0$ and all the inbreeding comes from F_{IS}; thus, from Equation 13, $CM_t = (x/4)(1 + CM_{t-2} + 2CM_{t-1})$. At equilibrium $CM_t = CM_{t-1} = CM_{t-2} = \widehat{CM}$, to give

$$\widehat{CM} = F_{IS} = \frac{x}{4 - 3x}, \qquad (15)$$

which is a well-known result (GHAI 1969; LI 1976, p. 245).

Prediction of *y*: y accounts for the probability that two individuals are full-sibs given that they are not mated. This can be evaluated in the following manner. In a population with $N/2$ families and k_i reproductive offspring per family, the probability that two individuals are full-sibs is

$$\frac{\sum_{i=1}^{N/2} k_i(k_i - 1)}{N(N - 1)},$$

which, for stable census number gives

$$\frac{(N/2)(V_k + 2)}{N(N - 1)}.$$

Since there are $N(N - 1)/2$ pairs, the total number of possible full-sib pairs is $N(V_k + 2)/4$. Of these, $xN/2$ are pairs of full sibs actually mated and thus $N(V_k + 2)/4 - xN/2$ are not mated. Analogously, from the $N(N - 1)/2$ possible pairs, $N/2$ are actually mated and $N(N - 1)/2 - N/2$ are not mated, therefore the probability that two individuals are full-sibs given that they are not mated is

$$y = \frac{N(V_k + 2)/4 - xN/2}{N(N - 1)/2 - N/2} = \frac{V_k + 2 - 2x}{2N - 4},$$

or approximately, for large N,

$$y \approx \frac{S_k^2 + 2 - 2x}{2N}. \qquad (16)$$

For simplicity, we have considered no sexes (*i.e.*, monoecious individuals) in this derivation but, obviously, the result equally applies for two sexes because, among non-mated pairs, there is no need to distinguish between those of the same or different sex.

WRIGHT (1969, p. 197) gave equations similar to (13) and (14) for changes in F_{IT} and F_{ST} (*F* and *E*,

TABLE 3

Observed and predicted [assuming that the probability that two individuals are full-sibs given that they are not mated, y, is $2/N$ as in WRIGHT-POLLAK's equations or $(2 - x)/N$, where x is the proportion of full-sib matings; see text] values of F_{IT} and F_{ST} for a population with size $N = 64$, multinomial distribution of family size ($S_k^2 = 2$) and full-sib mating whenever possible

	Observed		$y = 2/N$		$y = (2 - x)/N$	
Gen	F_{IT}	F_{ST}	F_{IT}	F_{ST}	F_{IT}	F_{ST}
1	0.120	0.008	0.121	0.008	0.121	0.008
4	0.194	0.033	0.198	0.039	0.194	0.034
7	0.220	0.059	0.231	0.072	0.222	0.060
10	0.243	0.085	0.259	0.104	0.244	0.085
13	0.268	0.109	0.284	0.136	0.265	0.111
16	0.284	0.134	0.309	0.166	0.286	0.135
19	0.303	0.157	0.334	0.195	0.305	0.159

Gen, generation.

respectively, in his notation) and POLLAK (1988) derived his expression for the effective size from these equations. However, they assumed that the probability that two random copies of a gene in random mates that are not sibs are derived from parents that were sibs (our y) was $2/N$, a result which is correct for random mating but not for nonrandom mating as shown by equation (16). In Table 3 values for F_{IT} and F_{ST} calculated as stated above from equations (13) and (14) as well as those calculated from the WRIGHT-POLLAK's equations are compared to their simulated values for the case with $N = 64$ and full-sib mating whenever possible. It is clear that the WRIGHT-POLLAK's equations are only approximate in the early generations and the bias accumulates rapidly.

Approximation of the effective size for large N: From Equation 14 and letting $\pi_t = 1 - CM_t$ and $P_t = 1 - CO_t$,

$$P_t = y\left[\frac{1}{4}\pi_{t-2} + \frac{1}{2}\pi_{t-1}\right] + (1 - y)P_{t-1}, \tag{17}$$

and from Equation 12, recalling that $CM_t = F_{IT}$ and $CO_t \approx F_{ST}$, then $\pi_t = P_t\Phi$, where $\Phi = (1 - F_{IS})$.

Letting λ be the asymptotic rate of decline in heterozygosity, the value of P in generation t will be $P_t = \lambda P_{t-1} = \lambda^2 P_{t-2} = \ldots$. Substituting into Equation 17

$$\lambda^2 P_{t-2} = y\left[\frac{1}{4}\Phi P_{t-2} + \frac{1}{2}\Phi\lambda P_{t-2}\right] + (1 - y)\lambda P_{t-2},$$

which factoring out and rearranging gives $\lambda^2 - y + y\Phi/2]\lambda - y\Phi/4 = 0$. Solving for λ and neglecting terms with y^2 (as shown in Equation 16, $y \propto 1/N$ and, therefore, square terms are negligible for large N), $\lambda = 1 - y(1 - 3\Phi/4)$. From (15), $\Phi = 1 - F_{IS} = (4 - 4x)/(4 - 3x)$ plus terms of order $1/N$, and therefore, $\lambda = 1 - y/(4 - 3x)$. By definition, $(1 - \lambda) \approx 1/2N_e$ (see CROW and KIMURA 1970, p. 104), so ignoring second order terms,

$$N_e = \frac{4 - 3x}{2y}. \tag{18}$$

By substituting (16) into (18) an estimate of the effective size is obtained in terms of the population census, the proportion of full-sib matings achieved, and the variance of family size,

$$N_e = \frac{N(4 - 3x)}{S_k^2 + 2 - 2x}. \tag{19}$$

In the case of a Poisson distribution of family size, with $S_k^2 = 2$, (16) reduces to $y = (2 - x)/N$, and (19) to

$$N_e = \frac{N(4 - 3x)}{4 - 2x}. \tag{20}$$

Recalling that $\alpha \approx F_{IS}$ when N is large (actually, $\alpha = F_{IS} - \alpha_R$, where $\alpha_R = -1/2N - 1/2T$ is the value of α for the random mating case with multinomial distribution of family size and T scored individuals; ROBERTSON 1965), x can be given from (15) as a function of α, $x = 4\alpha/(1 + 3\alpha)$. Substituting into (19) and (20), it gives equation (10) and (11), respectively, as expected.

DISCUSSION

The effective population number is usually smaller than the actual number of adults of reproductive age, the reasons for this being, traditionally, unequal numbers of males and females, temporal variation in population number and greater than Poisson variability in the number of progeny per parent (CROW and KIMURA 1970, pp. 109–110). Selection also causes a reduction in effective population size (ROBERTSON 1961; WRAY and THOMPSON 1990).

With unselected populations where matings between relatives are more frequent than at random, there is also a reduction in effective size because of the additional increase in genetic drift due to the correlation of the gene frequency in the parents. This reduction is independent of the population size and approximately linear with the departure from the Hardy-Weinberg proportions.

The causes of this reduction in the effective size can be seen in a different manner. When mates are related, the variance of the number of offspring is not affected but the variance of the number of grandoffspring is enhanced because there is a covariance between the number of offspring generated by the sons and daughters of a given family. Likewise, the variance of the number of subsequent descendants is also accordingly enhanced for the same reason. The classical equation of KIMURA and CROW (1963) for the effective size using means and variances of the number of offspring per parent (Equation 1) is inappropriate in this case as it takes into account only the distribution of descendants in one generation. WRAY and THOMP-

SON (1990) derived a method to predict effective sizes considering the long-term contribution from ancestors to descendants. For the conditions dealt with in this paper, the predicted effective size would be $N_e \approx 2N/(1 + S_r^2)$, where S_r^2 is the variance of long-term contributions of ancestors to descendants. This method was, however, derived assuming random mating which is why, perhaps, their predictions are severe underestimations for large departures from Hardy-Weinberg proportions (data not shown).

CROW and DENNISTON (1988) also gave an equation for the inbreeding effective size which considers means and covariances in the progeny. For the same situation as above, this equation would reduce to $N_e = N/(1 + S_{km,f})$, where $S_{km,f}$ is the covariance of the number of male and female progeny per parent. Nevertheless, this equation does not consider the possibility of nonrandom mating which can be clearly seen because when there is no covariance of male and female offspring (*e.g.*, with an independent distribution of family size for each sex), the predicted effective size equals the census number irrespective of the existence of nonrandom mating. Furthermore, this equation is, in general, not valid for more than one generation either for a multihypergeometric distribution of family size or, for example, in the extreme case of constant family size ($S_{km,f} = 0$), where the predicted effective size equals the actual size but it should be approximately double.

Finally, POLLAK (1987, 1988) derived an equation to predict effective sizes with a mix of sib mating and random mating. His derivation, however, was based on a recurrent equation for the inbreeding of the population with an incorrect coefficient and, therefore, gave incorrect roots.

In this paper, we have derived and checked by simulation a corrected form of the classical equation of KIMURA and CROW for the variance effective size when inbreeding is due to mating between relatives, excluding selfing (Equation 10). Further, following a derivation in terms of inbreeding, we have also derived an equation to predict effective sizes in the particular case where a certain proportion of full-sib matings is performed every generation (Equation 19), correcting POLLAK's equation. Equation 10 is applicable to any situation but requires knowledge of the magnitude of the departure from the Hardy-Weinberg ratios, a parameter which is not usually known. Equation 19, on the contrary, is applicable to a more restricted case of nonrandom mating, but has the advantage of permitting a direct estimation of the effective size provided we know the average number of full-sib matings which will be performed. A mating structure where the maximum number of full-sib matings is performed every generation has been shown to be a valuable method for increasing fixation probabilities of recessive mutants without causing delays in times to fixation or decreasing fixation probabilities for genes with different gene action (CABALLERO, KEIGHTLEY and HILL 1991) and Equation 19 would be of practical direct use in that case.

It is easy to obtain an equation like Equation 19 for other cases of partial inbreeding. For example, with a mix of half-sib mating and random mating when mothers of half-sibs are themselves half-sibs, $F_{IS} = z/(8 - 7z)$ (GHAI and KEMPTHORNE 1971; HEDRICK and COCKERHAM 1986), where z is the proportion of half-sib matings. Again approximating α by F_{IS} and substituting into (10), considering each sex separately, we obtain

$$N_e = \frac{N_{eR}(8 - 7z)}{4(1 - z) + S_k^2(2 - z)}$$

or, for a Poisson family size ($S_k^2 = 2$),

$$N_e = \frac{N_{eR}(8 - 7z)}{8 - 6z} = \frac{N_{eR}}{1 + F_{IS}} \approx \frac{N_{eR}}{1 + \alpha},$$

as with partial selfing and partial full-sib mating.

The results obtained in this paper have an important implication on the discussion by POLLAK (1988) of the reduction of the effective size caused by different mating systems of inbreeding. POLLAK argued that under partial selfing $N_e = N_{eR}/(1 + F_{IS}) = N_{eR}(1 - \beta/2)$, under partial full-sib mating $N_e = N_{eR}/(1 + 3F_{IS}) = N_{eR}(1 - 3\beta/4)$, and under partial half-sib mating if mothers of half-sibs are themselves half-sibs $N_e = N_{eR}/(1 + 7F_{IS}) = N_{eR}(1 - 7\beta/8)$, where β is the proportion of the corresponding inbreeding mating. Thus, he concluded that the more intense the inbreeding, the less the reduction in effective size with respect to the random case for $(1 - \beta/2) > (1 - 3\beta/4) > (1 - 7\beta/8)$. The result for partial selfing is correct but we have shown that in the two other cases with $S_k^2 = 2$, (a Poisson family size with the number of male and female offspring uncorrelated), $N_e = N_{eR}(4 - 3\beta)/(4 - 2\beta)$ and $N_e = N_{eR}(8 - 7\beta)/(8 - 6\beta)$, respectively, or $N_e = N_{eR}/(1 + F_{IS})$ in all cases. Therefore, as $(1 - \beta/2) < (4 - 3\beta)/(4 - 2\beta) < (8 - 7\beta)/(8 - 6\beta)$, the expected result that an increase in the intensity of the local inbreeding leads to a greater decrease in effective size is obtained: quite simply it is a function only of F_{IS}.

Most of the discussion on this paper has focused on cases where the number of inbred matings is more frequent than at random. It is, however, common to avoid matings between relatives in human and other primate social structures and in breeding programs applied to livestock production. This results in $F_{IS} < 0$ and, therefore, an increase in effective size. Nevertheless, this increase is very small, especially for large population sizes. A simulation was run where matings between relatives were minimized by mating individuals with the smallest coancestry. With $N = 64$, the value of F_{IS} was -0.019 and N_e was 67.5, while for $N = 200$, they were -0.010 and 201.5, respectively. This is easy to understand from the fact that, with random mating, the number of inbred matings is independent of N. For example, one full-sib mating is expected on average for any value of N. Thus, if N is large there is not much effect of whether or not that mating is avoided.

We are grateful to J. F. CROW, E. POLLAK and an anonymous referee for helpful comments and suggestions.

LITERATURE CITED

ASKEW, R. R., 1968 Considerations on speciation in *Chalcidoidea* (*Hymenoptera*). Evolution **22:** 642–645.

CABALLERO, A., P. D. KEIGHTLEY and W. G. HILL, 1991 Strategies for increasing fixation probabilities of recessive mutations. Genet. Res. **58:** 129–138.

CROW, J. F., and C. DENNISTON, 1988 Inbreeding and variance effective population numbers. Evolution **42:** 482–495.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper & Row, New York.

CROW, J. F., and N. E. MORTON, 1955 Measurement of gene frequency drift in small populations. Evolution **9:** 202–214.

GHAI, G. L., 1969 Structure of populations under mixed random and sib mating. Theor. Appl. Genet. **39:** 179–182.

GHAI, G. L., and O. KEMPTHORNE, 1971 The dynamics of populations under partial inbreeding. J. Indian Soc. Agric. Stat. **23:** 142–151.

HARTL, D. L., and A. G. CLARK, 1989 *Principles of Population Genetics*, Ed. 2. Sinauer Associates, Sunderland, Mass.

HEDRICK, P. W., and C. C. COCKERHAM, 1986 Partial inbreeding: equilibrium heterozygosity and the heterozygosity paradox. Evolution **40:** 856–861.

HILL, W. G., 1979 A note on effective population size with overlapping generations. Genetics **92:** 317–591.

JAIN, S. K., 1976 The evolution of inbreeding in plants. Annu. Rev. Ecol. Syst. **7:** 469–495.

KIMURA, M., and J. F. CROW, 1963 The measurement of effective population number. Evolution **17:** 279–288.

LATTER, B. D. H., 1959 Genetic sampling in a random mating population of constant size and sex ratio. Aust. J. Biol. Sci. **12:** 500–505.

LI, C. C., 1976 *First Course in Population Genetics.* Boxwood, Pacific Grove, Calif.

POLLAK, E., 1987 On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics **117:** 353–360.

POLLAK, E., 1988 On the theory of partially inbreeding finite populations. II. Partial sib mating. Genetics **120:** 303–311.

ROBERTSON, A., 1961 Inbreeding in artificial selection programmes. Genet. Res. **2:** 189–194.

ROBERTSON, A., 1965 The interpretation of genotypic ratios in domestic animal populations. Anim. Prod. **7:** 319–324.

SCHEMSKE, D. W., and R. LANDE, 1985 The evolution of self-fertilization and inbreeding depression in plants. 2. Empirical observations. Evolution **39:** 41–52.

SELANDER, R. K., 1970 Behavior and genetic variation in natural populations. Am. Zool. **10:** 53–66.

WRAY, N. R., and R. THOMPSON, 1990 Prediction of rates of inbreeding in selected populations. Genet. Res. **55:** 41–54.

WRIGHT, S., 1951 The genetical structure of populations. Ann. Eugenics **15:** 323–354.

WRIGHT, S., 1969 *Evolution and the Genetics of Populations, Vol. 2. The Theory of Gene Frequencies.* University of Chicago Press, Chicago.

Communicating editor: B. S. WEIR