Effective Size of Populations Under Selection

Enrique Santiago[1] and Armando Caballero

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

ABSTRACT

Equations to approximate the effective size (N_e) of populations under continued selection are obtained that include the possibility of partial full-sib mating and other systems such as assortative mating. The general equation for the case of equal number of sexes and constant number of breeding individuals (N) is $N_e = 4N/[2(1 - \alpha_I) + (S_k^2 + 4Q^2C^2)(1 + \alpha_I + 2\alpha_O)]$, where S_k^2 is the variance of family size due to sampling without selection, C^2 is the variance of selective advantages among families (the squared coefficient of variation of the expected number of offspring per family), α_I is the deviation from Hardy-Weinberg proportions, α_O is the correlation between genes of male and female parents, and Q^2 is the term accounting for the cumulative effect of selection on an inherited trait. This is obtained as $Q = 2/[2 - G(1 + r)]$, where G is the remaining proportion of genetic variance in selected individuals and r is the correlation of the expected selective values of male and female parents. The method is also extended to the general case of different numbers of male and female parents. The predictive value of the formulae is tested under a model of truncation selection with the infinitesimal model of gene effects, where C^2 and G are a function of the selection intensity, the heritability and the intraclass correlation of sibs. Under random mating $r = \alpha_I = -1/(N - 1)$ and $\alpha_O = 0$. Under partial full-sib mating with an average proportion β of full-sib matings per generation, $r \approx \beta$ and $\alpha_O \approx \alpha_I \approx \beta/(4 - 3\beta)$. The prediction equation is compared to other approximations based on the long-term contributions of ancestors to descendants. Finally, based on the approach followed, a system of mating (compensatory mating) is proposed to reduce rates of inbreeding without loss of response in selection programs in which selected individuals from the largest families are mated to those from the smallest families.

IN the absence of selection, all the individuals of a population have the same expected number of offspring, and differences in family size are due only to random sampling of individuals among families under a given reproductive system. In practice, however, populations are under selection, and, in general, families do not have equal probabilities of contribution to the next generation because of inherited or noninherited causes. This leads to an increase in homozygosity and random changes in gene frequency. Thus, the effective population size [N_e, the size of an ideal population that would give rise to the variance of change in gene frequency or the rate of inbreeding observed in the actual population under consideration (WRIGHT 1931)] is smaller than the number N of reproductive individuals. WRIGHT (1939) derived a formula for the effective population size in terms of variance of family size (S_k^2), $N_e \approx 4N/(2 + S_k^2)$, that can be applied when selection is acting on a noninherited trait. Since then expressions have been developed to predict effective population size for a number of different cases (see CABALLERO 1994 for a review).

When selection acts on an inherited trait, the effective size cannot be predicted solely from the variance of the family size at a given generation, because a fraction of the selective advantage of individuals remains in descendants over generations. Thus, the progeny of an individual with a high selective value tends to produce a large progeny number. If a neutral allele is randomly associated with that individual, its frequency will tend to increase, but eventually the rate of increase will slow down as the average selective advantage of descendants is diluted by segregation. Therefore, a fraction of random changes in gene frequency in a given generation is positively correlated with changes in previous generations, and WRIGHT's formula overpredicts the effective population size. The problem was first discussed in the context of artificial selection by ROBERTSON (1961), who gave a formula to predict the effective size of populations comprising full-sib families when the asymptotic state of correlated changes is reached under steady selection.

Corresponding author: Armando Caballero, Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JT, Scotland.
E-mail: eang60@castle.edinburgh.ac.uk

[1] *Present address:* Departamento de Biología Funcional, Universidad de Oviedo, Oviedo 33071, Spain.

He introduced the idea of the accumulation of selective advantages of individuals over generations. Thus, the selective advantage of an individual is expected to be reduced by one-half each generation in its descendants (the average selective advantage of the offspring is half that of the parents), and the total selective advantage over generations increases in a series $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$ up to a limiting value of twice the selective advantage present in the initial generation. By using ROBERTSON's derivation, NEI and MURATA (1966) reached an equivalent expression for the case of inherited fertility differences. BURROWS (1984) derived a formula for the effective size in the first generation of truncation selection, in which ROBERTSON's expression is only the first order approximation. He did not address, however, subsequent generations of selection and, consequently, the fundamental problem of correlated changes in gene frequencies over generations.

The solution given by ROBERTSON clearly underpredicts the effective size of selected populations, especially for high heritability and intense selection (HILL 1985; WRAY and THOMPSON 1990), because the increasing competitiveness of contemporaries as selection proceeds is not accounted for (WRAY and THOMPSON 1990). To overcome this deficiency, WRAY and THOMPSON (1990), working on the infinitesimal model (BULMER 1980), developed a recursive method to approximate the effective size as a function of the mean and variance of the contributions of ancestors in the first generation (when selection starts) to descendants in the limit. The method, however, does not allow easy predictions of the effective size. Simpler methods, which consider changes in gene frequencies over two generations, have been developed by WRAY *et al.* (1990), but, as expected, these methods overpredict the effective size. More recently, WOOLLIAMS *et al.* (1993) and WRAY *et al.* (1994) have derived equations to predict means and variances of the contributions of ancestors to descendants for mass and index selection, respectively. Thus, following the approach of WRAY and THOMPSON (1990), good approximations to the effective size of populations under selection and random mating of selected individuals can be obtained.

In this paper, equations to predict effective size under selection are obtained by an alternative method. An approach similar to that of ROBERTSON is followed, correcting and extending his predictions, and the results are discussed in connection with the WOOLLIAMS *et al.* (1993) approach. The method is also an extension of the derivation of CABALLERO and HILL (1992a), which allows the equations to be integrated into a general framework (KIMURA and CROW 1963; CROW and DENNISTON 1988; CABALLERO and HILL, 1992a) where selection and nonrandom mating of selected individuals (such as positive or negative assortative mating, or deliberate mating between full sibs) can be considered simultaneously. Finally, based on the principles of the approach followed, a new system of mating (which we will call compensatory mating) is proposed to reduce rates of inbreeding with practically no loss of response to selection.

DERIVATION OF EXPRESSIONS

The genetic model: We assume first a population with the same number of breeding males and females in which parents are pair mated to create $N/2$ families each generation. Extension to the general case of different numbers of male and female parents will be given later. Let f_i be the relative fitness of family i, with mean $\Sigma_{i=1}^{N/2} f_i / (N/2) = 1$ and variance $C^2 = [\Sigma_{i=1}^{N/2} f_i^2 / (N/2)] - 1$. As population size is assumed to be constant over generations, $2f_i$ is the expected number of offspring contributed by family i. We consider an autosomal neutral allele unlinked to the selected genes with mean frequency p_0 in generation 0, which is the initial unselected generation with a full-sib family structure. Effective population size is computed from the variance of change in gene frequency of this neutral allele among infinite (conceptual) replicates of the population. We will consider first the variance of change in gene frequency after one generation of selection. Later, the asymptotic effective population size under continued selection, when the variance of change in gene frequency (or the rate of inbreeding) is approximately constant over generations, will be addressed. Discrete generations are assumed.

First generation: Changes in gene frequency are the result of three independent processes acting each generation: random association between the neutral allele and families with a selective advantage or disadvantage, random sampling of individuals among families and Mendelian sampling of the neutral allele in the families in which one or both parents are heterozygotes. The two last processes comprise nonselective causes of drift, and expressions to predict effective size in this case have been derived by a number of authors (see CABALLERO 1994 for a review). In what follows the derivation of CABALLERO and HILL (1992a) is repeated for completeness and extended to take into account the first process of inherited variation.

Let $(x_{im} + x_{if})/2$ be the frequency of the neutral allele in family i, where m and f denote male and female parent, respectively. Thus, x_{im} or x_{if} is 0, $\frac{1}{2}$, or 1 if the individual carries zero, one or two copies of the gene,

respectively. The gene frequency in the population in generation 0 is

$$p_0 = \frac{1}{N} \sum_{i=1}^{N/2} (x_{im} + x_{if}).$$

Now, assume that family i contributes k_i offspring. The mean gene frequency in the first generation is then

$$p_1 = \frac{1}{N} \sum_{i=1}^{N/2} \left[k_i \left(\frac{x_{im} + x_{if}}{2} \right) + \frac{1}{2} \sum_{j=1}^{k_i} (\delta_{ijm} + \delta_{ijf}) \right],$$

where δ_{ijs} is the difference in gene frequency between the jth sampled gene from the individual i of sex s and its parental value x_{is}, i.e., δ_{ijs} is zero if the parent is a homozygote or $\pm\frac{1}{2}$ if a heterozygote. Therefore, the change in gene frequency is

$$p_1 - p_0 = \frac{1}{N} \sum_{i=1}^{N/2} \left(\frac{x_{im} + x_{if}}{2} \right)(k_i - 2)$$
$$+ \frac{1}{2N} \sum_{i=1}^{N/2} \sum_{j=1}^{k_i} (\delta_{ijm} + \delta_{ijf}). \quad (1)$$

Under selection the number of offspring contributed by the ith parent (k_i) can be partitioned into two terms: the expected contribution due to its selective value (the $2f_i$ defined above), plus a possible deviation due to random sampling (d_i) with average value zero among families, i.e., $k_i = 2f_i + d_i$. Thus,

$$p_1 - p_0 = \frac{1}{N} \sum_{i=1}^{N/2} \left(\frac{x_{im} + x_{if}}{2} \right)(2f_i - 2)$$
$$+ \frac{1}{N} \sum_{i=1}^{N/2} \left(\frac{x_{im} + x_{if}}{2} \right) d_i + \frac{1}{2N} \sum_{i=1}^{N/2} \sum_{j=1}^{k_i} (\delta_{ijm} + \delta_{ijf})$$
$$= S_1 + D_1 + H_1, \quad (2)$$

where S_1, D_1 and H_1 represent the change in gene frequency of the neutral gene from generation 0 to 1 because of random association between the neutral allele and a family with a selective advantage or disadvantage, random sampling of individuals among families and segregation of heterozygotes, respectively. These three terms are assumed to be independent of one another. First, we will consider the total contribution (k_i), and later we will partition this term into its two components ($2f_i$ and d_i). From (1) the variance of change in gene frequency is

$$V[p_1 - p_0] = \frac{1}{N^2} \left\{ V \left[\sum_{i=1}^{N/2} \left(\frac{x_{im} + x_{if}}{2} \right)(k_i - 2) \right] \right.$$
$$\left. + \frac{1}{4} V \left[\sum_{i=1}^{N/2} \sum_{j=1}^{k_i} (\delta_{ijm} + \delta_{ijf}) \right] \right\}.$$

Gene frequencies and numbers of offspring are expected to be uncorrelated in generation 0, and Men-

delian sampling terms are also uncorrelated and have equal variance ($V[\delta_{ij}]$). We take the same retrospective approach as CROW and KIMURA (1970, p. 353) in defining the effective size from the observed distribution of offspring number, such that the value k_i for family i is fixed, describing the actual number of offspring from family i, and not a random variable. We, however, define the variance of family size ($V[k]$) as

$$V(k) = \frac{\sum_{i=1}^{N/2} (k_i - 2)^2}{N/2} = \frac{\sum_{i=1}^{N/2} k_i^2}{N/2} - 4. \quad (3)$$

Finally, we ignore terms in $1/N^2$ from correlations among the x_i because their sum is fixed. Hence,

$$V[p_1 - p_0] = \frac{1}{N^2} \left\{ \frac{N}{2} V(k) V \left(\frac{x_{im} + x_{if}}{2} \right) + \frac{N}{2} V(\delta_{ij}) \right\}.$$

Now, we take expectations for the variances and covariances of gene frequencies and Mendelian sampling terms over an infinite number of conceptual replicates of the population. Thus, $V(x_{im}) = V(x_{if}) = p_0(1 - p_0)(1 + \alpha_I)/2$, where α_I is the deviation from Hardy-Weinberg proportions or the correlation between genes within individuals (approximately equal to WRIGHT's F_{IS} statistic [WRIGHT 1969, pp. 294–295]). The expectation of $\text{cov}(x_{im}, x_{if}) = p_0(1 - p_0)\alpha_O$, where α_O is the correlation between genes in pairs of parents or the correlation between genes within individuals in their offspring. Finally, the expectation of $V(\delta_{ij})$ equals the expected frequency of heterozygotes $2p_0(1 - p_0)(1 - \alpha_I)$ times the variance generated from them ($\frac{1}{4}$). Thus, substituting above we obtain

$$V[p_1 - p_0]$$
$$= \frac{p_0(1 - p_0)}{2N} \frac{V(k)(1 + \alpha_I + 2\alpha_O)}{4}$$
$$+ \frac{p_0(1 - p_0)}{2N} \frac{1 - \alpha_I}{2}. \quad (4)$$

The second term in (4) is due to segregation of heterozygotes. The first term is due to the variation in family size, which, as stated above, will have two components, due to differences in fitnesses among families and due to sampling variation. Now we partition $V(k)$ into its two components. Because of the different relative fitnesses of the families, the expected number of offspring contributed by family i is $E[k_i] = 2f_i$. Its variance ($\sigma_{k_i}^2$) will depend on the distribution of the number of individuals available for selection in each family. We consider three distributions.

Poisson: The total number of individuals available for selection is not fixed, and for a given family with an expected contribution $2f_i$, the actual number of descendants is assumed to be Poisson distributed with mean $2f_i$. Thus, the variance of the possible total num-

ber of individuals contributed by that family is, of course,

$$\sigma_{k_i}^2 = 2f_i. \tag{5}$$

Constant: Under artificial selection the same number of progeny is often evaluated per family. If n individuals of each sex are evaluated per family ($2n$ for both sexes), the probability of selecting a given individual of a family with an expected contribution $2f_i$ is $2f_i/2n$. The variance of the possible number of individuals contributed by the family will be approximately

$$\sigma_{k_i}^2 = 2f_i(1 - f_i/n). \tag{6}$$

Binomial: If the total number of individuals that are available for selection is fixed previous to selection but there is variation between families, the variance of contributions will have two components. One is due to differences in the number of individuals evaluated per family. This is assumed to be binomially distributed with mean n for one sex, and variance $n(1 - 2/N) \approx n$, when N is not very small. Thus, dividing by the squared mean number of individuals evaluated in each family (n^2, because the mean number of individuals of one sex contributed by one family is one), this component is $1/n$ and, considering both sexes together, $2/n$. The second component is equivalent to the variance of contributions when the same number of individuals is evaluated per family (Expression 6) and, therefore,

$$\sigma_{k_i}^2 = (2/n) + 2f_i(1 - f_i/n). \tag{7}$$

Then, taking expectations over families in the expression for $V(k)$ (Equation 3),

$$V(k) = \frac{\sum_{i=1}^{N/2} E[k_i^2]}{N/2} - 4 = \frac{\sum_{i=1}^{N/2} (\sigma_{k_i}^2 + E[k_i]^2)}{N/2} - 4.$$

Substituting $E[k_i]$ by $2f_i$ and $\sigma_{k_i}^2$ by (5), (6) or (7) and noting that, as stated before, $\sum_{i=1}^{N/2} f_i/(N/2) = 1$ and $\sum_{i=1}^{N/2} f_i^2/(N/2) = C^2 + 1$, the variance of family size after one generation of selection when the number of individuals available for selection in each family follows a Poisson, constant or binomial distribution is

$$V(k) = 2 + 4C^2, \tag{8}$$

$$V(k) = 2\left(1 - \frac{1 + C^2}{n}\right) + 4C^2 \tag{9}$$

or

$$V(k) = 2\left(1 - \frac{C^2}{n}\right) + 4C^2, \tag{10}$$

respectively. In these expressions the first term denotes the sampling component of variation (which we will denote V_n), and the second ($4C^2$) is the selective com-

ponent. Thus, substituting $V(k) = V_n + 4C^2$ from (8–10) into (4),

$$V[p_1 - p_0] = \frac{p_0(1 - p_0)}{2N} C^2(1 + \alpha_I + 2\alpha_O)$$

$$+ \frac{p_0(1 - p_0)}{2N} \frac{V_n}{4}(1 + \alpha_I + 2\alpha_O)$$

$$+ \frac{p_0(1 - p_0)}{2N} \frac{1 - \alpha_I}{2}$$

$$= E(S_1 + D_1 + H_1)^2$$

$$= E(S_1^2) + E(D_1^2) + E(H_1^2), \tag{11}$$

where E denotes expectation and the three terms are the variance of change in gene frequency due to random association between the neutral allele and a family with a selective advantage or disadvantage, random sampling of individuals among families and segregation of heterozygotes, respectively (cf. Equation 2). Equating the variance of change in gene frequency in an ideal population $[p_0(1 - p_0)/2N_e]$ to Equation 11, we obtain the effective size in the first generation of selection

$$N_{e,1} = \frac{4N}{2(1 - \alpha_I) + (V_n + 4C^2)(1 + \alpha_I + 2\alpha_O)}. \tag{12}$$

For a Poisson distribution of sampling variation ($V_n = 2$), and ignoring α_I and α_O terms (*i.e.*, assuming random mating and large N), (12) reduces to

$$N_{e,1} = \frac{N}{1 + C^2}$$

(*cf.* ROBERTSON 1961).

Cumulative effect of selection: Changes in gene frequency due to noninherited causes (random sampling of individuals among families, D_1, and segregation of heterozygotes, H_1) are not expected to persist over generations. However, the change in gene frequency due to random association between the neutral allele and families with a given selective advantage or disadvantage (S_1) will partially remain over generations, as will be shown next. This change in gene frequency due to the different fitnesses of families can also be expressed as a covariance (PRICE 1970) of relative fitnesses (f_i) and gene frequencies ($p_i = [x_{im} + x_{if}]/2$) of families,

$$S_1 = \frac{\sum_{i=1}^{N/2} f_i p_i}{N/2} - E(p_i) = \text{cov}_1(f_i, p_i).$$

Assuming that differences between f_i values of the families correspond to an additive component of genetic variation, the expected value of this covariance

in offspring before selection will be the same as the original covariance in parents. However, it is expected that selection reduces the genetic variance in selected offspring; selected individuals are less variable than the whole group in the same generation (BULMER 1980). The remaining proportion (G_1) of genetic variance will depend on the selection procedure and genetic system (its prediction for truncation selection on a normally distributed trait will be given later). Therefore, the expected remaining association between selective values and gene frequencies in selected individuals is

$$\text{cov}_1 (f_{i'}, p_{i'}) = G_1 \, \text{cov}_1 (f_i, p_i), \quad (13)$$

where the prime denotes selected individuals. After the selected individuals are mated, a fraction of this covariance will remain in the newly formed families of generation 2. This can be represented as

$$\text{cov}_2 (f_i, p_i) = \text{cov}_1 \left(\frac{f_{im'} + f_{if'}}{2}, \frac{p_{im'} + p_{if'}}{2} \right)$$

$$= {}^1\!/_4 \, [\,\text{cov}_1 (f_{im'}, p_{im'}) + \text{cov}_1 (f_{if'}, p_{if'})$$
$$+ \text{cov}_1 (f_{im'}, p_{if'}) + \text{cov}_1 (f_{if'}, p_{im'})\,],$$

where $p_{im'}$ ($p_{if'}$) are the gene frequency of male (female) parents (from generation 1) of the new families (in generation 2) and $f_{im'}$ ($f_{if'}$) are their expected selective values. All the individuals in generation 2 that come from the same family in generation 1 have the same expected values of f and p. The first two terms in brackets are the covariances between f and p values in each parent (which are expected to be the same for males and females, $\text{cov}_1[f_{im'}, p_{im'}] = \text{cov}_1[f_{if'}, p_{if'}] = \text{cov}_1[f_{i'}, p_{i'}]$), and the two last terms are the covariances between f values in one parent and p values in the other. These latter can be obtained as the product of the covariance between f and p values in the same parent ($\text{cov}_1[f_{i'}, p_{i'}]$) and the correlation between the expected f values of male and female parents, which we will denote r. Thus

$$\text{cov}_2 (f_i, p_i) = {}^1\!/_4 [2 \, \text{cov}_1 (f_{i'}, p_{i'}) + 2 \, \text{cov}_1 (f_{i'}, p_{i'}) \, r]$$

$$= {}^1\!/_2 (1 + r) \, \text{cov}_1 (f_{i'}, p_{i'}),$$

and substituting (13),

$$\text{cov}_2 (f_i, p_i) = \frac{G_1}{2} (1 + r) \, \text{cov}_1 (f_i, p_i). \quad (14)$$

Equation 14 is the expected change of gene frequency in the second generation due to an association in the first generation. The expected changes in the following generations due to an association in the first generation will also be in the same direction as S_1, but the magnitude of these changes will decrease

by a proportion $G_{t-1}(1 + r)/2$ every generation t. Thus, summarizing,

$$\text{cov}_1 (f_i, p_i) = S_1$$

is the expected change of gene frequency in the first generation due to an association between the neutral allele and families with a given selective advantage or disadvantage,

$$\text{cov}_2 (f_i, p_i) = \frac{G_1}{2} (1 + r) \, \text{cov}_1 (f_i, p_i) = \frac{G_1}{2} (1 + r) \, S_1$$

is the expected change in the second generation given an association in the first generation,

$$\text{cov}_3 (f_i, p_i) = \frac{G_2}{2} (1 + r) \, \text{cov}_2 (f_i, p_i)$$

$$= \frac{G_2}{2} (1 + r) \, \frac{G_1}{2} (1 + r) \, S_1$$

is the expected change in the third generation given an association in the first generation, and so forth. Under steady selection the proportional reduction in genetic variance per generation will be approximately constant over generations, so we assume that $G_1 = G_2 = G_3 = \cdots = G$. Therefore, the expected changes in gene frequency due to association between the neutral allele and families with a selective advantage or disadvantage in the first generation are

$$S_1, \quad S_1 \left[\frac{G}{2} (1 + r) \right], \quad S_1 \left[\frac{G}{2} (1 + r) \right]^2, \cdots$$

in generation 1, 2, 3, ...

Analogously, we can apply the same arguments for the new random associations between the neutral allele and families with a given selective advantage or disadvantage in generation 2 and in further generations. Then, changes in gene frequency due to associations in the second generation are

$$S_2, \quad S_2 \left[\frac{G}{2} (1 + r) \right], \quad S_2 \left[\frac{G}{2} (1 + r) \right]^2, \cdots$$

in generation 2, 3, 4, . . . , and similarly for further generations. Therefore, the expected variance of gene frequency drift is

$$V[p_1 - p_0] = E(S_1 + D_1 + H_1)^2$$

in the first generation (which is Equation 11),

$$V[p_2 - p_0] = E\Bigg(S_1 + D_1 + H_1 + S_1 \left[\frac{G}{2} (1 + r) \right]$$
$$+ S_2 + D_2 + H_2 \Bigg)^2$$

in the second generation,

$$V[p_3 - p_0] = E\left(S_1 + D_1 + H_1 + S_1\left[\frac{G}{2}(1 + r)\right] \right.$$

$$+ S_2 + D_2 + H_2 + S_1\left[\frac{G}{2}(1 + r)\right]^2$$

$$\left. + S_2\left[\frac{G}{2}(1 + r)\right] + S_3 + D_3 + H_3 \right)^2$$

in the third generation, and so on. As the gene frequency in generation 0 is the same for all conceptual replicates of the population, $V[p_0] = 0$ and $V[p_t - p_0] = V[p_t]$. Let us denote $Q_1 = 1$, $Q_2 = 1 + [G(1 + r)/2]$, $Q_3 = 1 + [G(1 + r)/2] + [G(1 + r)/2]^2$, $Q_4 = 1 + [G(1 + r)/2] + [G(1 + r)/2]^2 + [G(1 + r)/2]^3$, *etc.* Thus,

$$V[p_1] = E(Q_1 S_1 + D_1 + H_1)^2,$$

$$V[p_2] = E(Q_2 S_1 + D_1 + H_1 + Q_1 S_2 + D_2 + H_2)^2,$$

$$V[p_3] = E(Q_3 S_1 + D_1 + H_1 + Q_2 S_2 + D_2 + H_2$$
$$+ Q_1 S_3 + D_3 + H_3)^2,$$

and so on. Assuming that population size and the selective (C^2) and nonselective (V_n) components of variance are constant over generations, $E(S_2^2) = E(S_1^2)(1 - 1/2N_{e,1})$, $E(S_3^2) = E(S_1^2)(1 - 1/2N_{e,1})(1 - 1/2N_{e,2})$, $E(S_4^2) = E(S_1^2)(1 - 1/2N_{e,1})(1 - 1/2N_{e,2})(1 - 1/2N_{e,3})$, *etc.*, and similarly for $E(D^2)$ and $E(H^2)$, where $N_{e,t}$ is the effective population size in generation t. Assuming that differences between $N_{e,t}$ in consecutive generations are small such that $(1 - 1/2N_{e,1}) \approx (1 - 1/2N_{e,2}) \approx \cdots \approx (1 - 1/2N_{e,t}) = (1 - 1/2N_e)$ and denoting $E(S_1^2)$, $E(D_1^2)$ and $E(H_1^2)$ by $E(S^2)$, $E(D^2)$ and $E(H^2)$,

$$V[p_1] = Q_1^2 E(S^2) + E(D^2) + E(H^2),$$

$$V[p_2] = [Q_2^2 E(S^2) + E(D^2) + E(H^2)] + [Q_1^2 E(S^2)$$
$$+ E(D^2) + E(H^2)][1 - 1/2N_e],$$

$$V[p_3] = [Q_3^2 E(S^2) + E(D^2) + E(H^2)]$$
$$+ [Q_2^2 E(S^2) + E(D^2) + E(H^2)][1 - 1/2N_e]$$
$$+ [Q_1^2 E(S^2) + E(D^2) + E(H^2)][1 - 1/2N_e]^2,$$

and so on. These equations give the variance of gene frequency among infinite conceptual replicates of the population in each generation. Under no selection ($C^2 = 0$), binomial distribution of family size with large N and random mating ($V_n = 2$, $\alpha_I \approx \alpha_O = 0$) and substituting $E(S^2)$, $E(D^2)$ and $E(H^2)$ by the three terms in Equation 11, we obtain

$$V[p_t] = p_0(1 - p_0)\left[1 - \left(1 - \frac{1}{2N}\right)^t\right],$$

as it is expected in the ideal population.

Effective population size considering changes in variance of the gene frequency from generation 0 to 1 can be obtained from

$$\frac{V[p_1] - V[p_0]}{p_0(1 - p_0) - V[p_0]} = \frac{Q_1^2 E(S^2) + E(D^2) + E(H^2)}{p_0(1 - p_0)}$$

$$= \frac{1}{2N_{e,1}}.$$

Substituting the three terms in (11) and rearranging, we obtain Equation 12, as expected. Analogously, we can compute an effective size considering changes in variance of the gene frequency from generation 1 to 2

$$\frac{V[p_2] - V[p_1]}{p_0(1 - p_0) - V[p_1]} = \frac{1}{2N_{e,2}}$$

$$= \{[Q_2^2 E(S^2) + E(D^2) + E(H^2)]$$
$$+ [Q_1^2 E(S^2) + E(D^2) + E(H^2)][1 - 1/2N_e]$$
$$- [Q_1^2 E(S^2) + E(D^2) + E(H^2)]\}/$$
$$\{p_0(1 - p_0) - [Q_1^2 E(S^2) + E(D^2) + E(H^2)]\},$$

which after substitution and rearrangement gives

$$N_{e,2} = \frac{4N}{2(1 - \alpha_I) + (V_n + 4Q_2^2 C^2)(1 + \alpha_I + 2\alpha_O)},$$

$$(15)$$

and so forth. In general for large t, an expression for the asymptotic effective size can be obtained as

$$N_e = \frac{4N}{2(1 - \alpha_I) + (V_n + 4Q^2 C^2)(1 + \alpha_I + 2\alpha_O)},$$

$$(16)$$

where

$$Q = 1 + \left[\frac{G}{2}(1 + r)\right] + \left[\frac{G}{2}(1 + r)\right]^2 + \cdots$$

$$= \sum_{i=0}^{\infty} \left[\frac{G}{2}(1 + r)\right]^i = \frac{2}{2 - G(1 + r)}. \quad (17)$$

Hence, if it is assumed that there is no reduction in variance due to selection (*i.e.*, $G = 1$) and ignoring r, this gives ROBERTSON's (1961) series $Q = 1 + 1/2 + 1/4 + 1/8 + \cdots = 2$.

In Equation 16 the term ($V_n + 4Q^2 C^2$) consists of the variance of the family size after one generation of selection ($V_n + 4C^2$) (see Equation 12) plus the cumulative effect of selection on an inherited trait ($4[Q^2 - 1]C^2$).

Common environmental variation of sibs and noninherited sources of variation in family numbers can also be included in the noncumulative term of this variance.

Equation 16 can be simplified by noting that the terms C^2/n in Equations 9 and 10 are, in general, very small. Thus, neglecting them, V_n equals the variance of family size with random selection, *i.e.*, 2 with binomial distribution of individuals available for selection and $2(1 - 1/n)$ with constant distribution, which we will denote S_k^2,

$$N_e = \frac{4N}{2(1 - \alpha_I) + (S_k^2 + 4Q^2C^2)(1 + \alpha_I + 2\alpha_O)}.$$

(18)

For sampling variation binomially distributed ($S_k^2 = 2$) and assuming random mating and large N ($\alpha_I \approx \alpha_O = 0$), (18) reduces to

$$N_e = \frac{N}{1 + Q^2C^2}$$

(*cf.* ROBERTSON 1961).

Equation 18 has been derived in a general way. Now, we specify the value of the expected correlation of fitnesses between male and female parents (r in Equation 17) and the expected correlation of genes within individuals (α_I) and between mates (α_O) for two particular cases, random mating and partial full-sib mating of selected parents.

Random mating of selected parents: With random mating and infinite population size the expected value of r in Equation 17 is zero. However, in a finite population a negative correlation between the expected selective values of mates is generated under random mating, because their sum is fixed. This can be obtained as follows. Because $E[f_i] = 1$, $\sum_{i=1}^{N} (f_i - 1) = 0$. Therefore, $\sum_{i=1}^{N} (f_i - 1)^2 + \sum_{i \neq j}^{N} (f_i - 1)(f_j - 1) = 0$ and $NC^2 + N(N - 1) \operatorname{cov}(f_i, f_j) = 0$. From this we obtain $\operatorname{cov}(f_i, f_j) = -C^2/(N - 1)$ and hence,

$$E(r) = -\frac{1}{N - 1},$$

(19)

as would be expected.

α_I in Equation 18 is the deviation from Hardy-Weinberg proportions in male or female parents. Under random mating this is not zero but $-1/(N - 1)$ (KIMURA and CROW 1963; see also ROBERTSON 1965). α_O is the correlation between genes of male and female parents, which is zero under random mating.

Partial full-sib mating of selected parents: The structure of the equations above allows the possibility of including a certain proportion of mating between relatives. In particular, we can consider the case where every generation an average proportion β of the mat-

ings are between full sibs and the remaining $1 - \beta$ are at random, the full-sib mating habit not being inherited. When mating is made between full sibs, the correlation between the expected selective values of both parents (r in Equation 17) is one (because full sibs have the same expected f value), whereas for random mating it is approximately zero. Thus, the expected correlation between the selective values of both parents in the new families is approximately $r = \beta \times 1 + (1 - \beta) \times 0 = \beta$ and $Q = 2/(2 - G[1 + \beta])$. With partial full-sib mating α_I asymptotes very quickly to a value that approximates $\beta/(4 - 3\beta)$ (GHAI 1969; FALCONER 1989, pp. 97–98) and, ignoring second-order terms, $\alpha_O \approx \alpha_I$. Thus, Equation 18 then holds except that the term $(1 + \alpha_I + 2\alpha_O) \approx (1 + 3\alpha_I)$ (CABALLERO and HILL 1992a).

Different numbers of males and females: The previous approach can be extended to different numbers of sexes by considering males and females separately, because changes in gene frequencies of the neutral gene are uncorrelated in male and female parents. Assume a population with N_m males and N_f females, such that $N_m < N_f$ and each generation one male is mated to an integral number N_f/N_m females.

The effective number of individuals of sex s after one generation of selection can be predicted using a slightly corrected version of the equation of CROW and DENNISTON (1988), including a term $2\alpha_O$ from CABALLERO and HILL (1992a),

$$N_{es} = 4N_s \bigg/ \bigg\{ \left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}} \right] (1 - \alpha_{I,s})$$

$$+ \left[\frac{V(k_{sm})}{\mu_{sm}^2} + \frac{2 \operatorname{Cov}(k_{sm}, k_{sf})}{\mu_{sm}\mu_{sf}} + \frac{V(k_{sf})}{\mu_{sf}^2} \right]$$

$$\times (1 + \alpha_{I,s} + 2\alpha_O) \bigg\}, \quad (20)$$

and considering the appropriate variances and covariances of offspring number for the case of selection. N_s is the number of parents of sex s, $\mu_{sm}(= N_m/N_s)$ and $V(k_{sm})$ are the mean and variance of male offspring from parents of sex s, respectively, and analogously for female offspring. $\operatorname{Cov}(k_{sm}, k_{sf})$ is the covariance between the number of male and female offspring from parents of sex s, and $\alpha_{I,s}$ is the deviation from Hardy-Weinberg proportions in parents of sex s. The effective size of the population is obtained by

$$\frac{1}{N_e} = \frac{1}{4N_{em}} + \frac{1}{4N_{ef}}$$

(21)

(CROW and DENNISTON 1988).

Thus, we now derive the expected variances and covariances of offspring number with one generation of selec-

tion (expressions will be given only for numbers of male offspring but those for female offspring are analogous, substituting subscripts m by f). By definition,

$$V(k_{sm}) = \frac{\sum_{i=1}^{N_s} k_{ism}^2}{N_s} - \left(\frac{N_m}{N_s}\right)^2, \qquad (22)$$

where k_{ism} is the number of male offspring contributed by parent i of sex s, and

$$\text{Cov}(k_{sm}, k_{sf}) = \frac{\sum_{i=1}^{N_s} k_{ism} k_{isf}}{N_s} - \frac{N_m}{N_s}\frac{N_f}{N_s}. \qquad (23)$$

We define f_{ism} as the relative selective advantage of parent i of sex s in their contribution to male offspring. The mean and variance of f_{ism} over families is $\sum_{i=1}^{N_s} f_{ism}/N_s = 1$ and $C_{sm}^2 = [\sum_{i=1}^{N_s} f_{ism}^2/N_s] - 1$, respectively. Thus, the expected number of male offspring contributed by parent i of sex s is $(N_m/N_s) f_{ism}$. The variance of this number $(\sigma_{k_{ism}}^2)$ depends, as before, on the distribution of the number of offspring from parents of sex s available for selection. For a Poisson distribution,

$$\sigma_{k_{ism}}^2 = \frac{N_m}{N_s} f_{ism}. \qquad (24)$$

For a constant distribution,

$$\sigma_{k_{ism}}^2 = \frac{N_m}{N_s} f_{ism}\left[1 - \frac{N_m}{N_s}\frac{f_{ism}}{n_s}\right], \qquad (25)$$

where n_s is the number of offspring of one sex available for selection from parents of sex s. For a binomial distribution,

$$\sigma_{k_{ism}}^2 = \left(\frac{N_m}{N_s}\right)^2\frac{1}{n_s} + \frac{N_m}{N_s} f_{ism}\left[1 - \frac{N_m}{N_s}\frac{f_{ism}}{n_s}\right], \qquad (26)$$

where now n_s is the average number of offspring of one sex available for selection from parents of sex s. Finally, $\sigma_{k_{ism,sf}} = 0$ if there are not any sources of variation other than sampling.

Taking expectations in (22) and (23) over families,

$$V(k_{sm}) = \frac{\sum_{i=1}^{N_s} E[k_{ism}^2]}{N_s} - \left(\frac{N_m}{N_s}\right)^2$$

$$= \frac{\sum_{i=1}^{N_s} (\sigma_{k_{ism}}^2 + E[k_{ism}]^2)}{N_s} - \left(\frac{N_m}{N_s}\right)^2 \qquad (27)$$

and

$$\text{Cov}(k_{sm}, k_{sf}) = \frac{\sum_{i=1}^{N_s} E[k_{ism} k_{isf}]}{N_s} - \frac{N_m}{N_s}\frac{N_f}{N_s}$$

$$= \frac{\sum_{i=1}^{N_s} (\sigma_{k_{ism,sf}} + E[k_{ism}]E[k_{isf}])}{N_s} - \frac{N_m}{N_s}\frac{N_f}{N_s}. \qquad (28)$$

Substituting $E[k_{ism}] = (N_m/N_s) f_{ism}$ and $\sigma_{k_{ism}}^2$ from (24), (25) or (26) into (27), we obtain

$$V(k_{sm}) = \frac{N_m}{N_s} + C_{sm}^2\left(\frac{N_m}{N_s}\right)^2 \qquad (29)$$

for Poisson distribution,

$$V(k_{sm})$$

$$= \frac{N_m}{N_s}\left[1 - \frac{1 + C_{sm}^2}{n_s}\left(\frac{N_m}{N_s}\right)\right] + C_{sm}^2\left(\frac{N_m}{N_s}\right)^2 \qquad (30)$$

for constant distribution, or

$$V(k_{sm}) = \frac{N_m}{N_s}\left[1 - \frac{C_{sm}^2}{n_s}\left(\frac{N_m}{N_s}\right)\right] + C_{sm}^2\left(\frac{N_m}{N_s}\right)^2 \qquad (31)$$

for binomial distribution. Expressions for female offspring are the same as (29–31) substituting subscripts m by f. Finally, substituting $E[k_{ism}] = (N_m/N_s) f_{ism}$, $E[k_{isf}] = (N_f/N_s) f_{isf}$ and $\sigma_{k_{ism,sf}} = 0$ into (28) and assuming that the selected trait is controlled by the same set of genes in males and females,

$$\text{Cov}(k_{sm}, k_{sf}) = \frac{N_m}{N_s} C_{sm}\frac{N_f}{N_s} C_{sf}, \qquad (32)$$

for any distribution of available individuals for selection.

Again, the first term in (29–31) denotes the sampling component, which we will denote V_{sm} (and V_{sf} for the analogous expressions for female offspring), and the second term denotes the selective one. Equation 32 only has a selective component, assuming that there are not any sources of covariation other than sampling. Therefore, substituting (29–31) as $V(k_{sm}) = V_{sm} + C_{sm}^2(N_m/N_s)^2$, and analogously for females $V(k_{sf}) = V_{sf} + C_{sf}^2(N_f/N_s)^2$, as well as (32) into (20), we obtain an expression for the effective number of individuals of sex s in the first generation of selection,

$$N_{es} = 4N_s\Big/\Bigg\{\left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}}\right](1 - \alpha_{I,s})$$

$$+ \left[\frac{V_{sm}}{\mu_{sm}^2} + \frac{V_{sf}}{\mu_{sf}^2} + 4C_s^2\right](1 + \alpha_{I,s} + 2\alpha_O)\Bigg\}, \qquad (33)$$

where $C_s = (C_{sm} + C_{sf})/2$.

The cumulative change in gene frequency over generations is equivalent for both sexes because genes from male and female parents are selected with the same intensity in male and female offspring. Thus, although the remaining proportion of the genetic variance due to selection in sons (G_m) and daughters (G_f) can be different because selection can act with different intensities in both sexes, the remaining proportion of genetic variance in selected descendants of male parents is the

same as in selected descendants of female parents ($G = [G_m + G_f]/2$). As before, only changes in gene frequency due to selection on an inherited trait are expected to persist over generations. These changes can be expressed as $[\text{cov}_1(f_i, p_{im}) + \text{cov}_1(f_i, p_{if})]/2$, where $p_{im}(p_{if})$ is the frequency of the neutral gene in the male (female) parent of family i, and f_i is the expected selective value of family i. The offspring has the same covariance before selection, but after selection the covariance is reduced, $\text{cov}_1(f_{i'}, p_{is'}) = G\,\text{cov}_1(f_i, p_{is})$, where s is m or f and primes denote again selected individuals. After mating the selected individuals, the expected change in gene frequency from the first to the second generation within the set of genes from individuals of sex s in generation 0 is

$$\text{cov}_2(f_i, p_{is}) = \text{cov}_1\left(\frac{f_{im'} + f_{if'}}{2}, \frac{p_{ism'} + p_{isf'}}{2}\right)$$

$$= \frac{1}{4}[\text{cov}_1(f_{im'}, p_{ism'}) + \text{cov}_1(f_{if'}, p_{isf'})$$

$$+ \text{cov}_1(f_{im'}, p_{isf'}) + \text{cov}_1(f_{if'}, p_{ism'})],$$

where $f_{im'}$ and $f_{if'}$ are the expected selective values of male and female parents (from generation 1) of the new families (in generation 2), respectively, and $p_{ism'}$ and $p_{isf'}$ refer to frequency of genes that come from the individual of sex s of generation 0, through male and female parents in generation 1, respectively. Thus,

$$\text{cov}_2(f_i, p_{is}) = \frac{1}{4}[G_m\,\text{cov}_1(f_{im}, p_{ism})$$

$$+ G_f\,\text{cov}_1(f_{if}, p_{isf}) + G_m\,\text{cov}_1(f_{im}, p_{isf})$$

$$+ G_f\,\text{cov}_1(f_{if}, p_{ism})] = \frac{1}{4}[(G_m + G_f)$$

$$\times \text{cov}_1(f_i, p_{is}) + (G_m + G_f)\,\text{cov}_1(f_i, p_{is})\,r]$$

$$= \frac{G_m + G_f}{4}(1 + r)\,\text{cov}_1(f_i, p_{is})$$

$$= \frac{G}{2}(1 + r)\,\text{cov}_1(f_i, p_{is}), \quad (34)$$

where r is the correlation between the expected selective values of the families from which the male and female parents come.

Thus, following the same arguments as in the case for equal numbers of male and female parents, a general expression for the asymptotic effective size under selection is

$$N_{es} = 4N_s \Big/ \left\{ \left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}}\right](1 - \alpha_{I,s}) \right.$$

$$\left. + \left[\frac{V_{sm}}{\mu_{sm}^2} + \frac{V_{sf}}{\mu_{sf}^2} + 4Q^2 C_s^2\right](1 + \alpha_{I,s} + 2\alpha_O) \right\}, \quad (35)$$

combined for the two sexes by means of (21).

As before, if we neglect the terms in C_s^2/n_s in expres-

sions (30 and 31), V_{sm} and V_{sf} give the variances of offspring number under random selection, i.e., N_m/N_s for a binomial distribution of individuals available for selection, and $(N_m/N_s)[1 - (N_m/N_s n_s)]$ for a constant distribution (and analogously for female offspring). We will denote these variances as S_{sm}^2 and S_{sf}^2, respectively. Thus,

$$N_{es} = 4N_s \Big/ \left\{ \left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}}\right](1 - \alpha_{I,s}) \right.$$

$$+ \left[\left(\frac{S_{sm}^2}{\mu_{sm}^2} + \frac{2S_{sm,sf}}{\mu_{sm}\mu_{sf}} + \frac{S_{sf}^2}{\mu_{sf}^2}\right) + 4Q^2 C_s^2\right]$$

$$\left. \times (1 + \alpha_{I,s} + 2\alpha_O) \right\}. \quad (36)$$

The term $S_{sm,sf}$ does not appear in the derivation because it is zero in the absence of any causes of variation of offspring numbers other than sampling, but it is included for completeness, because these sources of non-inherited variation can also occur. Equation (36) is thus equivalent to CROW and DENNISTON's (1988) expression for the variance effective size with an extra term accounting for selection. Note that if $N_m = N_f = N/2$, (21 and 36) reduce to (18), as expected.

APPLICATION TO PARTICULAR SELECTIVE SYSTEMS

The derivations shown above were carried out in a general framework in which no reference to a particular selective and genetic model was made. To predict the effective population size under selection, two parameters must be known, G and C^2. The magnitude of these parameters is dependent on the selective system involved. We will consider here the application to an infinitesimal model of gene effects (BULMER 1980) in which descendants are selected according to their phenotypic values (a viability model) and there is random mating of selected parents, but the analysis could also be extended to other models (e.g., for a fertility model see NEI and MURATA [1966]).

Under truncation selection on a normally distributed trait, it has been demonstrated that selection reduces the additive genetic variance of the selected individuals by a proportion $G = 1 - kh^2$ (BULMER 1980), where h^2 is the heritability of the trait, $k = i(i - x)$, and x is the truncation point in the standardized normal distribution of phenotypes. Moreover, under phenotypic selection $C^2 \approx i^2\rho$ (MILKMAN 1978), where ρ is the intraclass correlation of full sibs.

As both h^2 and ρ decline over generations and we are interested in the asymptotic effective size (the effective size under continued selection when the rate of increase in inbreeding reaches its asymptotic value), it

is more suitable to use the asymptotic h^2 and ρ, i.e., their values in the same period for which the effective size is to be predicted. Strictly, h^2 and ρ decline continuously over generations until they are exhausted in a finite population, but we can consider the nearly steady state where this decline is negligible.

If population size is not very small (say $N > 10$), a good approximation for the additive genetic variance after a few generations of truncation selection with the infinitesimal model of gene effects (V_A^*, asterisks will denote asymptotic values) is given by expression (9.34) of BULMER (1980),

$$(1 + k)V_A^{*2} + (V_E - V_A)V_A^* - V_A V_E = 0, \quad (37)$$

where V_A and V_E are the initial additive and environmental variances, respectively. Solving for the positive solution of V_A^*, the asymptotic heritability (h^2*) is (see also GOMEZ-RAYA and BURNSIDE 1990)

$$h^2* = \frac{V_A^*}{V_A^* + V_E} \quad (38)$$

and the asymptotic intraclass correlation of full sibs is

$$\rho^* = \frac{V_b^*}{V_A^* + V_E}, \quad (39)$$

where V_b^* is the asymptotic additive genetic variance between families before selection. Assuming that the additive genetic variance within families is constant over generations and equal to $V_A/2$,

$$V_b^* = V_A^* - (V_A/2) = \frac{1}{2}V_A^*(1 - kh^2*). \quad (40)$$

For different numbers of male and female parents, similar approximations can be obtained for $G_m = (1 - k_m h^2*)$, where $k_m = i_m(i_m - x_m)$, i_m is the selection intensity, and x_m is the truncation point in the standardized normal distribution for males, and analogously for females. From the average value of G_m and $G_f(G)$, Q can be calculated with the same formula as with equal number of sexes (Equation 17). Accordingly, $C_{sm}^2 \approx i_m^2 \rho_s^*$ and $C_{sf}^2 \approx i_f^2 \rho_s^*$, where ρ_s^* is the asymptotic intraclass correlation of families of sex s. However, because with $N_m \ll N_f$ selection intensity in males can be very high, it is appropriate to estimate C^2 with more precision. BURROWS (1984) showed that $C^2 = i^2[\rho + \rho^2 x^2/2 + \rho^3(x^2 - 1)^2/6 + \cdots]$. Thus, we can use a better approximation to C^2,

$$C_{sm}^2 = i_m^2[\rho_s^* + \frac{1}{2}\rho_s^{*2} x_m^2 + \frac{1}{6}\rho_s^{*3}(x_m^2 - 1)^2], \quad (41)$$

and similarly for C_{sf}^2. More accurate approximations can also be obtained from the work of MENDELL and ELSTON (1974).

Again, the asymptotic variance can be obtained from the quadratic equation of BULMER (Equation 37) substituting $k = (k_m + k_f)/2$. Asymptotic heritabilities are

approximated by (38) and asymptotic intraclass correlations of families from parents of sex s by

$$\rho_s^* = \frac{V_{bs}^*}{V_A^* + V_E}. \quad (42)$$

With random mating

$$V_{bs}^* = \frac{1}{4} V_A^*(1 - k_m h^2*) + \frac{1}{4} V_A^*(1 - k_f h^2*) \frac{N_s}{N_f}, \quad (43)$$

because the expected genetic value of any family is the average of the genetic values of male and female parents, which were selected with different intensities.

The equations above can be used in a general way for other kinds of selection. The selection coefficient of a genotype is approximately equal to the product of the standard selection differential (i) and the standard genotypic effect (see e.g., FALCONER 1989, p. 202). This holds for any kind of selection in which fitness is a nondecreasing function of a distribution of phenotypes (MILKMAN 1978). The proportion of genetic variance that is still remaining in selected individuals (G) relative to the unselected ones in the same generation is the result of selection acting on genotypes (i.e., their fitness values). Therefore, all the fitness functions that produce the same selection differential i will give the same reduction in genetic variance. We can, thus, use the same expression for G above for any kind of selection with intensity i. The corresponding x value for any of these systems of selection would be the cut-off point in the normal distribution when truncation selection is carried out with the same intensity. Furthermore, the property that the fitness value is a function of intensity of selection and genotypic effect also allows the utilization of $i^2\rho$ to approximate C^2 under any kind of phenotypic selection in which fitness is a nondecreasing function of phenotypes. It must be noted, however, that this formula is only an approximation and it does not hold for very high selection intensities and heritabilities.

Effective population size for systems of mating such as assortative mating can also be predicted by the equations derived in this paper. The appropriate value of r in Equation 17 is the correlation between the expected f values of mates. Given that all the descendants from the same family have the same expected f value, r can be approximated under assortative mating as $r = \rho r_P$, where ρ is the intraclass correlation of family members, and r_P is the phenotypic correlation between male and female parents. Thus, maximum positive or negative assortative mating can be approximated by using $r_P \approx 1$ or -1, respectively. Strictly, the approximation only holds when matings are arranged between unselected individuals. Therefore, it is expected that it will not be very accurate with strong selection.

SYSTEM OF MATING TO REDUCE INBREEDING

In the model proposed above, it is assumed that there is no correlation between the changes in gene frequency produced by the three random processes in the same generation, *i.e.*, association of the neutral allele with selective values of the families (S), sampling variation in the selection of individuals to be parents of the next generation (D) and random sampling of the neutral gene within families in which one or both parents are heterozygotes (H). However, this does not hold for consecutive generations. We have seen that a fraction of the change in gene frequency in one generation is correlated with changes in the previous ones when the selective values of the families are inherited. This can be considered as the consequence of a deterministic change. If we expand the population to an infinite size, these changes will happen in that generation and in the next ones. As a deterministic or predictable change, one could imagine that that change can be manipulated to generate negative correlations between changes in gene frequency over generations and, therefore, reduce the overall effect of the drift process.

As was mentioned above, the number of offspring contributed by a given parent $i(k_i)$ can be partitioned into two terms: the contribution due to its selective value ($2f_i$) and a deviation due to sampling of individuals between families (d_i), *i.e.*, $k_i = 2f_i + d_i$. The terms correspond to S and D changes in gene frequency, respectively. If in a given generation, selected individuals from the families with the largest contributions are mated to selected individuals from the families with the lowest contributions, a negative correlation is generated in the new families between the contribution they had due to their selective value ($2f_i$) and the deviation due to sampling of individuals between families (d_i). This is a consequence of the fact that, under this system of mating, all the new families in one generation had the same $2f_i + d_i$ value in the previous generation. In other words all the couples produced by this system of mating had the same contributions the generation before (if we average the contribution of the families from which both parents came). Thus, if the parents of a given family had a high selective advantage in the generation before, their contributions due to sampling variation were small, and vice versa. The consequence is that, in addition to the positive correlation between changes due to association between selective values of families and gene frequencies (S changes) over generations, a negative correlation is generated between the S changes in one generation and the D changes due to sampling of individuals between families one generation before. These opposing correlations compensate each other every generation, and, as a result, the cumulative effect of selection approximately vanishes. Approximate pre-

dictions of this situation can, therefore, be made by using $Q^2 = 1$, the first term in the series. In practice this system of mating (compensatory mating) is performed by ordering families according to the total number of selected individuals and mating males with the highest ranking to females with the lowest, in sequence. With different numbers of male and female parents, individuals can be ordered according to the sum of the total number of their selected half-sibs and the total number of their selected full sibs with males given a weight N_f / N_m that of females to reflect their contributions.

TEST OF PREDICTIONS BY SIMULATION

Stochastic simulation was carried out to check the prediction equations. Artificial selection was made on a trait controlled by an infinitesimal additive model of gene effects (BULMER 1980) with initial heritability h^2. In generation 0 genotypic values were assumed to be normally distributed with mean zero and variance h^2. Phenotypes were obtained by adding to the genotypic value a random environmental deviation normally distributed with mean zero and variance $1 - h^2$. Truncation selection on the phenotypes was carried out each generation, such that the best N_s individuals of sex s out of the $n_s N_s$ evaluated of sex s were chosen as parents of the next generation, where n_s is the average number of individuals evaluated of each sex from parents of sex s. Mating of selected individuals followed different procedures: random mating; mating between full sibs whenever possible, random otherwise; maximum positive or negative assortative mating and compensatory mating. Discrete generations were assumed. Genotypic values of the offspring were obtained as the average of the genotypic values of their parents plus a random Mendelian deviation normally distributed with mean zero and variance $(h^2/2)[1 - (F_m + F_f)/2]$, where F_s is the inbreeding coefficient of the parent of sex s, obtained from pedigrees. Because population size is assumed to be constant over generations and the population is not permanently subdivided, asymptotic inbreeding and variance effective sizes must be equal, because the genotypic frequencies can be predicted both from the variance of gene frequencies and the inbreeding coefficients. For each generation WRIGHT's (1922) numerator relationship matrix was constructed, and average inbreeding coefficients per generation were calculated from the average coancestry of parents. Effective population sizes were obtained from the average rate of inbreeding ($\Delta F = [F_{t+1} - F_t]/[1 - F_t]$, where F_t is the inbreeding coefficient in generation t), generally between generations 5 and 14 (when this is approximately constant), as $N_e = (1/2\Delta F) - \frac{1}{2}$ (FALCONER 1989). Moreover, a neutral gene with initial fre-

TABLE 1

Simulated (*sim*) and predicted (*pre*) effective population size for a population
with 20 male and 20 female parents

n	h^2	B			C		
		sim	*pre*[1]	*pre*[2]	*sim*	*pre*[1]	*pre*[2]
2	0.1	37.2	37.0	36.5	48.0	47.7	47.0
	0.2	35.5	35.3	34.6	45.4	45.0	43.9
	0.4	33.4	33.6	32.8	42.5	42.2	41.0
	0.6	32.7	33.1	32.4	41.5	41.4	40.3
	0.8	32.9	33.4	32.8	41.6	41.9	41.0
3	0.1	34.8	34.8	34.1	40.9	40.5	39.5
	0.2	32.3	32.2	31.3	37.3	37.1	35.8
	0.4	29.7	29.8	29.1	33.9	33.9	32.9
	0.6	28.8	29.4	28.8	32.9	33.4	32.6
	0.8	29.2	30.2	29.9	33.7	34.4	34.0
6	0.1	31.8	31.2	30.2	33.7	33.3	32.2
	0.2	27.9	27.6	26.5	29.3	29.2	28.0
	0.4	24.8	24.8	24.0	25.9	26.1	25.3
	0.6	23.9	24.4	24.1	25.0	25.6	25.3
	0.8	25.0	25.9	26.0	26.2	27.3	27.5

pre[1], full prediction obtained with Equation 16 and simulated asymptotic heritability and intraclass correlation of full-sibs; *pre*[2], prediction obtained with Equation 18 and simplified or predicted parameters; *n*, number of individuals scored per sex, family and generation; h^2, initial heritability; *B*, binomial distribution of scored individuals per family; *C*, constant number of scored individuals per family. Standard errors ~0.1.

quency 0.5 and Hardy-Weinberg proportions in generation 0 was also included in all simulations. Effective size was computed from the variance of the gene frequencies of the neutral gene among replicates and from the reduction in the frequency of heterozygotes. Both estimates were always neither significantly different from each other nor from the estimates from pedigrees, as expected. These latter are, however, those shown below, because they had the lowest standard errors. Two thousand replicates were run for each case simulated.

Table 1 shows simulated (*sim*) and predicted (*pre*) values of effective size for a population with 20 male and 20 female parents, a range of initial heritabilities, three selection intensities and two distributions of number of scored individuals, binomial and constant (multinomial and multihypergeometric distribution of selected individuals, respectively).

Two predictions are made. The first (*pre*[1]) is aimed at checking the validity of the equations, and, therefore, it uses the full prediction with some parameters obtained from the simulations. Thus, (*pre*[1]) uses Equation 16, where V_n is given by the first term in (9) or (10), Q is given by (17) with *r* from (19), and $G = 1 - kh^{2*}$, where h^{2*} is the asymptotic heritability obtained from simulations as the average between generations 5 and 14. Finally, $C^2 = i^2 \rho^*$, where ρ^* is the asymptotic intraclass correlation of full sibs obtained from the simulations (average between generations 5 and 14). The

value of α_I for random mating is $-1/(N-1)$ and $\alpha_O = 0$. The second prediction (*pre*[2]) is aimed at giving a practical simple approximation. It uses Equation 18 (*i.e.*, terms in C^2/n in (9) and (10) are neglected). *r* in Equation 17 is also neglected and the asymptotic heritability and intraclass correlation are predicted by means of (38) and (39 and 40), respectively, with V_A^* from (37). An example of prediction follows. For $N = 40$, $n = 3$ and $h^2 = V_A = 0.4$, from standard statistical tables $i = 1.076$, $x = 0.408$ and $k = 0.719$. The asymptotic additive genetic variance can be obtained by solving (37), $V_A^* \approx 0.320$, and from (38–40), $h^{2*} \approx 0.348$, $V_b^* \approx 0.120$ and $\rho^* \approx 0.130$. Thus, $G \approx 1 - kh^{2*} = 0.750$, $Q \approx 2/(2 - G) = 1.60$ and $C^2 \approx i^2 \rho^* = 0.151$. From the variance of a multihypergeometric distribution (with random selection), $S_k^2 \approx 2(1 - 1/n) = 1.33$ and, finally, $\alpha_I = -1/(N-1) = -0.026$. Thus, from Equation 18 $N_e = 32.9$.

Simulated values have standard errors of ~0.1. The average absolute error in the predictions relative to the simulated values is 1.3% for *pre*[1] and 2.5% for *pre*[2]. Thus, *pre*[1] illustrates the validity of the equations, and *pre*[2] gives approximations that can be used for most practical purposes.

A case was also run where there was an initial heritability $h^2 = 0.4$ and a common environmental variance $V_C = 0.2$, with 20 pairs of parents and six individuals of each sex evaluated per family. The observed N_e was 20.5,

TABLE 2

Simulated (*sim*) and predicated (*pre*) effective population size for a population
with N individuals (half of each sex)

| | $n = 3$ | | | | $n = 6$ | | | |
| | $h^2 = 0.1$ | | $h^2 = 0.4$ | | $h^2 = 0.1$ | | $h^2 = 0.4$ | |
N	*sim*	*pre*	*sim*	*pre*	*sim*	*pre*	*sim*	*pre*
10	10.4	10.7	9.1	9.2	9.3	9.2	7.9	7.5
20	20.6	20.8	17.4	17.5	17.6	17.4	14.2	13.7
40	40.9	40.5	33.9	33.9	33.7	33.3	25.9	26.1
80	81.0	80.5	67.0	67.5	65.8	65.7	49.4	51.4
160	161.8	159.6	133.1	133.6	128.9	129.9	95.1	101.8

Prediction of N_e by Equation 16 and simulated asymptotic heritability and intraclass correlation of full-sibs. h^2, initial heritability. Constant number (n) of individuals scored per sex, family and generation. Standard errors of simulated values ~ 0.1.

and its prediction with Equation 16 where a term $4i^2 V_C$ is added to the term $4Q^2 C^2$ was 20.3.

Table 2 illustrates the predictions of N_e for different population sizes. It is worth noting that the ratio N_e/N is not constant for all population sizes, but it is reduced with increasing N until a limiting value is reached. Thus, for $n = 6$ and $h^2 = 0.4$, for example, the ratio is 0.79, 0.71, 0.65, 0.62 and 0.59 for increasing N (from 10 to 160 in Table 2, respectively). The main reason for this is that the asymptotic intraclass correlation among full sibs is smaller with smaller N, or, more precisely, the intraclass correlation decays faster with small population size. For the cases above the simulated correlation (average between generations 5 and 14) was 0.094, 0.109, 0.119, 0.121 and 0.123, respectively. Predictions in Table 2 use these asymptotic correlations, and, therefore, this effect is mostly accounted for.

Table 3 shows simulations and predictions for maximum negative and positive assortative mating of selected individuals. These were obtained by substituting r by $\rho^* r_P$ in Equation 17, where r_P is the phenotypic correlation among mates obtained from the simulation. This correlation was ~ -0.85 for negative assortative mating and ~ 0.90 for positive assortative mating. Predictions are very close to simulations except for positive assortative mating with very high heritability (0.8). The reason for this is that the approximation $r = \rho^* r_P$ is strictly valid with no selection and breaks down with very high heritability and intense selection.

Table 4 shows simulations and predictions for populations with 20 males and 40 or 100 females. In each generation one male was mated to two or five females, respectively, and n offspring of each sex was evaluated per mating. The first prediction (pre^1), aimed at checking the validity of the equations, is the full prediction using Equations 21 and 35 with V_{sm} and V_{sf} from the first term in (30). $C_s^2 = [(C_{sm} + C_{sf})/2]^2$, with C_{sm} and C_{sf} from (41) and using the asymptotic intraclass

correlation of sibs for sex s obtained from the simulations; $\alpha_s = -1/(2N_s - 1)$. Q is obtained by Equation 17 with $G = (G_m + G_f)/2$, where $G_m = 1 - k_m h^{2*}$ and $G_f = 1 - k_f h^{2*}$, h^{2*} being the asymptotic heritability obtained from the simulations and r from (19) substituting N by $(N_m + N_f)/2$ (as an approximation). The second prediction (pre^2) is the simple practical approximation obtained by (21) and (36), where C_s^2 is obtained with the first term in (41), r is neglected in Equation 17, and the asymptotic heritability and intraclass correlation are approximated by (38), (42) and (43), respectively, with V_A^* from (37) using $k = (k_m + k_f)/2$. Prediction errors average 1.2% for the full prediction (pre^1) and 3.2% for the approximation (pre^2).

TABLE 3

Simulated (*sim*) and predicted (*pre*) effective population size for a population with 20 male and 20 female parents and maximum negative or positive assortative mating

| | $n = 3$ | | $n = 6$ | |
h^2	*sim*	*pre*	*sim*	*pre*
		Negative		
0.1	41.2	41.0	34.2	34.0
0.2	38.4	38.3	30.4	30.4
0.4	36.2	36.1	27.7	27.9
0.6	36.8	36.3	28.0	28.0
0.8	39.1	38.3	31.2	30.6
		Positive		
0.1	39.9	39.5	33.1	32.8
0.2	35.3	35.0	28.0	27.7
0.4	29.4	30.5	22.5	23.4
0.6	25.6	26.1	20.3	22.6
0.8	20.4	28.0	16.8	22.6

Footnote as in Table 2.

TABLE 4

Simulated (*sim*) and predicted (*pre*) effective population size for a population with 20 male and *F* female parents

	h^2		$n = 3$			$n = 6$	
		sim	*pre*[1]	*pre*[2]	*sim*	*pre*[1]	*pre*[2]
$F = 40$	0.1	50.5	49.8	48.5	43.5	43.0	41.7
	0.2	45.8	45.1	43.8	37.5	37.5	36.4
	0.4	41.5	41.2	40.3	32.8	33.1	33.0
	0.6	40.5	40.6	40.2	32.3	32.6	33.3
	0.8	41.8	42.6	42.5	34.3	35.3	36.4
$F = 100$	0.1	59.0	57.8	56.8	52.6	51.9	51.0
	0.2	53.1	52.0	51.2	45.9	45.2	44.8
	0.4	47.9	47.3	47.4	40.2	40.0	40.9
	0.6	47.4	46.9	47.7	39.6	39.7	41.6
	0.8	49.5	49.9	51.0	42.7	43.5	45.9

pre[1], full prediction obtained with Equations 21 and 35 and simulated parameters; *pre*[2], prediction obtained with Equations 21 and 36 and simplified or predicted parameters; h^2, initial heritability. Constant number (*n*) of individuals scored per sex, family and generation. Standard errors of simulated values ~0.1.

An example of prediction with different numbers of male and female parents follows. Consider $N_m = 20$ males, $N_f = 100$ females, $n_f = 6$ individuals of each sex scored per full-sib family and $n_m = 30$ per half-sib family, random mating of selected parents and initial heritability 0.4. From standard statistical tables $i_m = 2.23$, $i_f = 1.50$, $x_m = 1.83$ and $x_f = 0.97$. Hence, $k_m = 0.89$, $k_f = 0.80$ and $k = (k_m + k_f)/2 = 0.84$. The asymptotic additive genetic variance can be obtained from (37), $V_A^* \approx 0.310$. From (38) $h^{2*} = 0.341$ and from (42) and (43) $V_{bm}^* \approx 0.065$, $V_{bf}^* \approx 0.110$, $\rho_m^* \approx 0.072$ and $\rho_f^* \approx 0.121$. Now, using the first term in (41), $C_{mm}^2 \approx i_m^2 \rho_m^* = 0.357$, $C_{mf}^2 \approx i_f^2 \rho_m^* = 0.161$, $C_{fm}^2 \approx i_m^2 \rho_f^* = 0.603$, $C_{ff}^2 \approx i_f^2 \rho_f^* = 0.273$, $C_m^2 = [(C_{mm} + C_{mf})/2]^2 = 0.249$ and $C_f^2 = [(C_{fm} + C_{ff})/2]^2 = 0.422$. Now, $G_m = 1 - k_m h^{2*} = 0.696$, $G_f = 1 - k_f h^{2*} = 0.727$, $G = (G_m + G_f)/2 = 0.711$ and $Q \approx 2/(2 - G) = 1.552$. Finally, $\mu_{mm} = \mu_{ff} = 1$, $\mu_{mf} = 1/\mu_{fm} = 5$, $S_{mm}^2 \approx 1 - 1/n_m = 0.967$, $S_{mf}^2 \approx (N_f/N_m)(1 - 1/n_f) = 4.167$, $S_{fm}^2 \approx (N_m/N_f)(1 - 1/n_m) = 0.192$, $S_{ff}^2 \approx 1 - 1/n_f = 0.833$, $S_{mm,mf} = S_{fm,ff} = 0$, $\alpha_{I,m} = -1/(2N_m - 1) = -0.026$, $\alpha_{I,f} = -1/(2N_f - 1) = -0.005$ and $\alpha_O = 0$. Substituting into Equation 36, $N_{em} = 17.1$, $N_{ef} = 25.6$, and from (21) $N_e = 4N_{em}N_{ef}/(N_{em} + N_{ef}) = 40.9$.

Table 5 illustrates the case where selected full sibs are mated whenever possible, at random otherwise. With this system of mating, rates of inbreeding take more generations to reach constant values than with random mating (especially for low heritability). Thus, effective population sizes and other asymptotic parameters were obtained from generation 15 to 24 for $h^2 = 0.1$ and 0.2 and from generation 10 to 19 for larger h^2. Predictions are obtained by means of Equation 18 with $\alpha_O \approx \alpha_I \approx \beta/(4 - 3\beta))$, where β is the average propor-

tion of full-sib matings per generation (~0.5). In addition r in Equation 17 is replaced by β. Effective sizes with ~50% full-sib mating are about half the corresponding effective sizes with random mating (cf. Table 1). Predictions are very accurate even for very high heritabilities.

Finally, Table 6 shows the results of the system of mating proposed to reduce inbreeding without loss in selection response (compensatory mating). In the simulations families were arranged in order according to the number of selected individuals they had. Males from the families with more selected members were mated with females from the families with less selected members, following a sequential procedure. Columns one and four of Table 6 show the effective size and the average rate of response to selection between genera-

TABLE 5

Simulated (*sim*) and predicted (*pre*) effective population size for a population with 20 male and 20 female parents and ~50% of the matings between full sibs

h^2	$n = 3$		$n = 6$	
	sim	*pre*	*sim*	*pre*
0.0	41.8	42.5	37.0	37.9
0.1	21.0	20.2	17.4	16.4
0.2	18.0	17.6	14.9	14.4
0.4	15.7	16.0	12.9	12.7
0.6	16.7	17.1	13.0	13.2
0.8	19.3	19.4	15.7	15.3

Prediction of N_e by Equation 18 and simulated asymptotic intraclass correlation of full-sibs. h^2, initial heritability. Constant number (*n*) of individuals scored per sex, family and generation. Standard errors of simulated values ~0.1

TABLE 6

Simulated (*sim*) and predicted (*pre*) effective population size and rate of response to selection for a population with 20 male and 20 female parents

n	h^2	N_e			ΔG	
		sim^R	sim^C	pre^C	sim^R	sim^C
3	0.1	40.9	44.2	44.9	0.09	0.09
	0.2	37.3	42.4	42.8	0.17	0.17
	0.4	33.9	40.2	39.7	0.33	0.32
	0.6	32.9	39.0	37.8	0.48	0.47
	0.8	33.7	38.3	36.8	0.62	0.63
6	0.1	33.7	38.6	39.3	0.12	0.12
	0.2	29.3	36.1	36.2	0.23	0.23
	0.4	25.9	33.2	32.4	0.43	0.43
	0.6	25.0	31.6	30.4	0.63	0.63
	0.8	26.2	31.0	29.9	0.83	0.84

N_e, effective population size; ΔG, rate of response to selection per generation in phenotypic standard deviation units (between generations 5 and 14); R, random mating; C, compensatory mating. Other parameters as in footnote of Table 2.

tions 5 and 14 with random mating of selected parents (R). Columns two and five show the corresponding values with compensatory mating (C). Rates of response for the two methods are nearly the same in all cases, whereas effective sizes are significantly larger with the system of mating. These are between 10 and 20% larger for the lower selection intensity, and between 15 and 30% larger for the higher selection intensity.

Increases in effective size are considerably larger with compensatory mating than with negative assortative mating (cf. Table 3), for which there is also an appreciable loss in selection response. A small loss in selection response was also observed (data not shown) with compensatory mating when inbreeding is not accounted for in the simulations to generate the genotypic values each generation. This is likely to be due to a certain component of negative assortative mating in the system of mating proposed, and the loss in response is compensated with the large reduction in inbreeding.

Predictions of effective size with compensatory mating are also shown in Table 6 (column three). These were obtained by means of Equation 16 with $Q^2 = 1$ and the intraclass correlation after one generation of selection, as it is expected that with this system of mating the cumulative effect of selection effectively disappears.

DISCUSSION

WRIGHT (1939) derived a formula to predict the effective size as a function of the variance of family size. The derivation assumes that the process of genetic drift is a consequence of uncorrelated changes in gene frequency over generations. This assumption does not hold when selection acts on an inherited trait, because associations between the neutral gene and selective values in a given generation are not removed in one generation. A fraction of genetic changes in one generation is due to remaining associations created in previous generations. As a result, WRIGHT's formula overpredicts the effective size of populations under selection. Obviously, the cumulative effect of changes fades as the association is removed by segregation and recombination. The problem was first addressed by ROBERTSON (1961). He introduced the term Q to take account of this cumulative process. By using a derivation based on the change of variance in gene frequency of a neutral gene unlinked to the genetic system, he obtained the equation $N/N_e = 1 + Q^2C^2$ for a multinomial distribution of family size, and to $N/N_e = (1 - 1/2n) + (1 - 1/n) Q^2C^2$ for a multihypergeometric distribution of family size, where the limiting value of Q is 2. Experimental checks of these predictions showed that they are reasonable approximations (JONES 1969; YOO 1980; GALLEGO and GARCIA-DORADO 1986; GALLEGO and CABALLERO 1990), but in most cases heritabilities were low or selection was not continued long enough for effective size to reach an asymptote. Simulation data, however, have made clear that ROBERTSON's predictions are underestimations of the asymptotic effective size, which become severe with large heritability and selection intensity (HILL 1985; WRAY and THOMPSON 1990).

WRAY and THOMPSON (1990) derived a method to predict rates of inbreeding based on the long-term contributions of ancestors in generation 1 to descendants in the limit. They arrived at the equation $N/N_e = (\mu_r^2 + \sigma_r^2)/2$, where μ_r and σ_r^2 are the mean and variance of the long-term contributions of ancestors to descendants, respectively. For a multinomial distribution of

family size and constant census size each generation, we can equate ROBERTSON's and WRAY and THOMPSON's equations, arriving at the approximate relation $\sigma_r^2 = 1 + 2Q^2C^2$. WRAY and THOMPSON (1990) also showed that ROBERTSON's prediction had a few deficiencies, the most important of which was that the increasing "competitiveness" of contemporaries as selection proceeds was not accounted for. In other words what was missing in ROBERTSON's derivation was the inclusion of the reduction in genetic variance due to selection (G) in the series to account for the cumulative effect of selection (Equation 17). Thus, the cumulative effect of changes fades faster than predicted by ROBERTSON, and the limiting value of Q is <2. WRAY and THOMPSON (1990) calculated the correct Q value as the ratio of the regression of number of descendants in generation t from ancestors in generation 1 to the corresponding regression from descendants in generation 2 to ancestors in generation 1. This gives the same result as the term S_∞ in WOOLLIAMS et al. (1993) and Equation 17 in this paper. Other minor problems in the derivation of ROBERTSON pointed out by WRAY and THOMPSON (1990) are the use of an incorrect intraclass correlation of full sibs after one generation of selection and a confusion between number of families and number of individuals in the derivation.

In contrast to WRAY and THOMPSON's arguments, our derivation is based on the same principles as that of ROBERTSON, *i.e.*, in terms of variance of change in gene frequency generation after generation. We have included a term accounting for selection in the general expression for the effective population size of CROW and DENNISTON (1988). Our prediction Equation 18 reduces to the same as ROBERTSON's for multinomial family size ($S_k^2 = 2$), large N and random mating ($\alpha_I \approx \alpha_O = 0$), but for the multihypergeometric case the equation is different, $N/N_e = (1 - 1/2n) + Q^2C^2$, the last term being reduced by ($1/n$) in ROBERTSON's expression. This is precisely what allows us to have just a single equation for any distribution of family size.

We have split the variance of the contributions of the families into two terms: selective and nonselective components. This partition has been made under an additive model of gene action. More generally, the selective component should include only the fraction of the variance of the contributions that is due to the selective values transmitted to descendants (the additive component of the genetic variance). Nonadditive genetic components and environmental effects of the selective values should be included into the nonselective component of the variance of the contributions because these effects are not cumulative.

WRAY and THOMPSON's (1990) method had the practical problem that some recurrence computing was needed to obtain predictions of long-term contributions. However, WOOLLIAMS et al. (1993) and WRAY et al. (1994) have recently derived expressions to predict variances of long-term contributions and thus, effective population sizes under mass and index selection. When their basic equation is reduced to the simplest case of equal numbers of males and females and multinomial family size, the result is $N/N_e = 1 + (1 + Q^2)C^2/2$ (WRAY et al. 1994), *i.e.*, it has the term $(1 + Q^2)/2$ instead of Q^2. This difference between our equation (or ROBERTSON's equation) and the latter reflects the fact that approximations are based on different approaches, as will be shown in what follows.

Although the effective size of interest in this paper is that in the asymptotic stage under continued selection (when the rate of change in inbreeding is approximately constant over generations), we can express the effective size that would correspond to the change of variance in different periods of the process. The effective size for the first generation of selection (Equation 12) under the simple conditions dealt with above is given by $N/N_e = 1 + C^2$. We can also obtain an expression for the effective size considering two generations of selection as the harmonic mean of $N_{e,1}$ (Equation 12) and $N_{e,2}$ (Equation 15), where $Q_2 = 1 + [G(1 + r)/2]$. Assuming random mating and large N ($\alpha_I \approx \alpha_O = 0$) and considering $G = 1$ and $r = 0$, for simplicity, the effective size considering two generations of drift is, then, given by $N/N_e = 1 + [1 + (1.5)^2]C^2/2$. Analogously, with the harmonic mean of $N_{e,1}$, $N_{e,2}$ and $N_{e,3}$ we obtain $N/N_e = 1 + [1 + (1.5)^2 + (1.75)^2]C^2/3$ for three generations, $N/N_e = 1 + [1 + (1.5)^2 + (1.75)^2 + (1.875)^2]C^2/4$ for four generations, and so on. In the limit the result would tend to $N/N_e = 1 + 4C^2$, as found by ROBERTSON. Therefore, the simplified equation of WRAY et al. (1994) is the same as the equation considering two generations of drift ($N/N_e = 1 + [1 + (1.5)^2]C^2/2$), except that the limiting term Q replaces that for the second generation ($Q_2 = 1 + \frac{1}{2}$). WRAY and THOMPSON's (1990) derivation is made in terms of the contributions from ancestors in the first generation to descendants in the limit. WOOLLIAMS et al. (1993) and WRAY et al. (1994) worked with this same approach. Thus, they use the equation in the second generation and make corrections to obtain a limiting solution (J. A. WOOLLIAMS, personal communication). The use of parameters (intraclass correlation and heritability) after one generation of selection is then needed to get accurate predictions with their equation. We are using the asymptotic equation where the corresponding asymptotic parameters are more suitable. It is worth noting, however, that the value of h^2 used to predict Q has little effect on the predictions of N_e, and h^2 after one generation of selection would be a good enough approximation (data not shown).

Prediction errors using our equations are ~3% within the range of selection intensities and heritabilities investigated by simulation. These are of about the same order of those obtained by the method of WOOL-LIAMS *et al.* (1993).

Our expressions for the effective population size in terms of KIMURA and CROW's (1963) and CROW and DENNISTON's (1988) equations allows us to include the possibility of partial full-sib mating of selected parents. Partial full-sib mating can be useful in increasing the fixation probability of recessive mutations (CABALLERO *et al.* 1991) as well as improving selection response in breeding programs (CABALLERO and HILL 1992b; TORO 1993). To make predictions with this system of mating, it is necessary both to account for correlations of genes between mates (CABALLERO and HILL 1992a) and to include an appropriate term in the value of Q. This term is the correlation of expected selective values of male and female parents, which approximately equals the average proportion of full-sib matings per generation. The inclusion of this correlation into Q also allows predictions for other systems of mating, as negative or positive assortative mating. With random mating Q practically reaches its asymptotic value in four generations or less. However, under some systems of nonrandom mating, especially those that increase the frequency of mating between relatives, a longer period of time is needed to reach the asymptotic Q value. Derivations have been made for species with separate sexes. Equation 18, however, can be readily applied to monoecious species under random mating.

Finally, a system of mating has been proposed in which individuals from high contribution families are mated to individuals from low contribution families. The consequence of this system of mating is the effective elimination of the cumulative effect of selection. Thus, effective population sizes are increased up to 30% with little or no loss of response. A system of mating like this can be very useful in breeding schemes, especially if combined with a selection strategy to reduce rates of inbreeding, like using an upwardly biased estimate of heritability in BLUP (Best Linear Unbiased Prediction) evaluations (GRUNDY *et al.* 1994).

A similar result to that obtained with compensatory mating, *i.e.*, the disappearance of the cumulative effect of selection, would occur if we could generate a negative correlation between the selective values of male and female parents ($r = -1$ in Equation 17). With negative assortative mating we generate a negative phenotypic correlation between mates (r_P) close to -1, but the corresponding genotypic correlation is approximately $r = \rho r_P$, as was stated above. Thus, if $\rho \approx 0.2$ (for, say, an initial heritability of 0.4), then $r \approx -0.2$, and the decrease in Q^2 is much smaller than with compensatory mating (cf. Tables 3 and 6).

Approximations for the variance of selective values and the reduction of genetic variance with selection used in this paper refer to the infinitesimal model for mathematical convenience and testing purposes, but they can be generalized to other selective and genetic systems. Equations are based on an initial standing variation before selection is carried out, where there is no linkage between the neutral and selected genes. N. H. BARTON (personal communication) has investigated situations where genetic variance is maintained by deleterious mutation-selection balance or fluctuation of the selective forces to produce a stable polymorphism.

LITERATURE CITED

BULMER, M. G., 1980 *The Mathematical Theory of Quantitative Genetics.* Clarendon Press, Oxford.

BURROWS, P. M., 1984 Inbreeding under selection from unrelated families. Biometrics **40**: 357–366.

CABALLERO, A., 1994 Developments in the prediction of effective population size. Heredity **73**: 657–679.

CABALLERO, A., and W. G. HILL, 1992a Effective size of nonrandom mating populations. Genetics **130**: 909–916.

CABALLERO, A., and W. G. HILL, 1992b Use of inbreeding in large populations. 43rd Annual Meeting of the E.A.A.P., Madrid, Spain.

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 and Row, New York.

FALCONER, D. S., 1989 *Introduction to Quantitative Genetics,* Ed. 3, Longman, London.

GALLEGO, A., and CABALLERO, A., 1990 The cumulative effect of artificial selection on the reduction of population effective size. Experimental and simulated data. J. Anim. Breed. Genet. **107**: 180–187.

GALLEGO, A., and GARCIA-DORADO, A., 1986 Evolution de l'effectif génétique de lignées de *Drosophila melanogaster* soumises à une sélection artificielle. Genet. Sel. Evol. **18**: 249–260.

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

GOMEZ-RAYA, L., and E. B. BURNSIDE, 1990 The effect of repeated cycles of selection on genetic variance, heritability, and response. Theor. Appl. Genet. **79**: 568–574.

GRUNDY, B., A. CABALLERO, E. SANTIAGO and W. G. HILL, 1994 A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. Anim. Prod. **59**: 465–468.

HILL, W. G., 1985 Fixation probabilities of mutant genes with artificial selection. Genet. Sel. Evol. **17**: 351–358.

JONES, L. P., 1969 Effect of artificial selection on rates of inbreeding in populations of *Drosophila melanogaster*. 1. Effect on early generations. Aust. J. Biol. Sci. **22**: 143–155.

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

MENDELL, N. R., and R. C. ELSTON, 1974 Multifactorial qualitative

traits: genetic analysis and prediction of recurrence risks. Biometrics **30:** 41–57.

MILKMAN, R., 1978 Selection differentials and selection coefficients. Genetics **88:** 391–403.

NEI, M., and M. MURATA, 1966 Effective population size when fertility is inherited. Genet. Res. **8:** 257–260.

PRICE, G. R., 1970 Selection and covariance. Nature **227:** 520–521.

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.

TORO, M. A., 1993 A new method aimed at using the dominance variance in closed breeding populations. Genet. Sel. Evol. **25:** 63–74.

WOOLLIAMS, J. A., N. R. WRAY and R. THOMPSON, 1993 Prediction of long-term contributions and inbreeding in populations undergoing mass selection. Genet. Res. **62:** 231–242.

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

WRAY, N. R., J. A. WOOLLIAMS and R. THOMPSON, 1990 Methods for predicting rates of inbreeding in selected populations. Theor. Appl. Genet. **80:** 503–512.

WRAY, N. R., J. A. WOOLLIAMS and R. THOMPSON, 1994 Prediction of rates of inbreeding in populations undergoing index selection. Theor. Appl. Genet. **87:** 878–892.

WRIGHT, S., 1922 Coefficients of inbreeding and relationship. Am. Nat. **56:** 330–338.

WRIGHT, S., 1931 Evolution in Mendelian populations. Genetics **16:** 97–159.

WRIGHT, S., 1939 Statistical Genetics in Relation to Evolution. Exposés de Biométrie et de Statistique Biologique, Herman & Cie, Paris, France.

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

YOO, B. H., 1980 Long-term selection for a quantitative character on large replicate populations of *Drosophila melanogaster*. 5. The inbreeding effect of selection. Aust. J. Biol. Sci. **33:** 713–723.