Interrelations between effective population size and other pedigree tools for the management of conserved populations

ARMANDO CABALLERO[1]* AND MIGUEL A. TORO[2]

[1] *Departamento de Bioquímica, Genética e Immunología, Facultad de Ciencias, Universidad de Vigo, 36200 Vigo, Spain*
[2] *Departamento de Mejora, Genética y Biotecnología, INIA, Carretera de la Coruña Km. 7, 28040 Madrid, Spain*

Summary

Genetic parameters widely used to monitor genetic variation in conservation programmes, such as effective number of founders, founder genome equivalents and effective population size, are interrelated in terms of coancestries and variances of contributions from ancestors to descendants. A new parameter, the effective number of non-founders, is introduced to describe the relation between effective number of founders and founder genome equivalents. Practical recommendations for the maintenance of genetic variation in small captive populations are discussed. To maintain genetic diversity, minimum coancestry among individuals should be sought. This minimizes the variances of contributions from ancestors to descendants in all previous generations. The method of choice of parents and the system of mating should be independent of each other because a clear-cut recommendation cannot be given on the latter.

1. Introduction

Maintaining genetic diversity is one of the primary goals in the management of populations in captivity. In recent years, a growing number of studies have been devoted to developing techniques for the analysis of genealogies, and to proposing parameters to monitor the amount of genetic diversity. For example, MacCluer *et al.* (1986) proposed the method of gene dropping analysis, a Monte Carlo simulation procedure to calculate expected gene frequencies and probabilities of allele losses in pedigrees. Different alleles are assigned to every founder and the genotypes of all descendants along the actual pedigree are assumed to be generated according to simple Mendelian rules. The entire procedure is repeated many times and the information from the genotypes of the actual generation summarized over replicates.

Lacy (1989) and Rochambeau *et al.* (1989) have defined the effective number of founders in order to measure the overall founder representation in a managed population accounting for the loss of genetic variability from unequal founder contributions. Lacy

(1989, 1995) also introduced the concept of founder genome equivalents as the theoretically expected number of founders that would be required to provide the genetic diversity in the actual population if the founders were equally represented and had lost no alleles. This parameter is directly related to genetic diversity, defined as the expected frequency of heterozygotes by descent, as well as to group coancestry (Cockerham, 1967), the average pairwise coancestry of a given group of individuals of the pedigree including reciprocals and self-coancestries (Lacy, 1989; Rodrigañez *et al.*, 1998).

Practical recommendations for the maintenance of genetic variability in captive populations have been made based on the above analyses and parameters. Alderson (1991) proposed computing the gene origin probabilities or each potential candidate for breeding with reference to the founders, and then selecting animals with the highest effective number of founders as a way of equalizing founder contributions. Ballou & Lacy (1995) noted, however, that maximizing gene diversity is not simply a matter of equalizing founder contributions, because subsequent generations are a source of drift. Ballou & Foose (1995) proposed calculating a target founder contribution for each founder, consisting of the expected proportion of the

* Corresponding author. Tel: +34 986 812568/78. Fax: +34 986 812556. e-mail: armando@uvigo.es

founder's alleles that have survived to the current generation, and preferentially breeding individuals from founders whose contribution currently falls below their targets. From the relation between average coancestry and genetic diversity it is deduced that choosing animals for breeding so that individuals in the next generation have the lowest average coancestry maximizes genetic diversity (Lacy, 1995; see also Lindgren *et al.*, 1996). Ballou & Lacy (1995) have shown by simulation that minimizing coancestries is more effective in maintaining genetic variation than other strategies such as equalizing founder contributions or breeding animals most likely to contain unique alleles; and Montgomery *et al.* (1997) showed that *Drosophila* retains significantly more gene diversity based on six microsatellites and seven allozyme loci.

Another issue in the management of captive populations is which system of mating to follow. Random mating, circular mating or avoidance of mating between relatives have been proposed, but there is no clear agreement on the technique that should be used except that very close inbreeding (half-sibs or closer) should be avoided because of the high probability of inviable or infertile offspring.

Although much progress has been made on the above issues, the relationship between some of the concepts is not completely clear in the literature. In this paper we contribute to this clarification by spelling out some of the interrelations between effective population size and other genetic parameters used in genealogical analysis. We will connect the definitions of effective number of founders, number of genome equivalents and effective population size, in terms of average coancestries and variances of contributions from ancestors to descendants, extending previous work. We will also relate the measures of diversity obtained from gene dropping analysis with these parameters. As a result of such interrelations, practical recommendations for the maintenance of genetic variation in the management of small captive populations will be discussed in terms of choice of breeding individuals and in terms of mating systems.

2. Basic genetic tools for the analysis of genealogies: inbreeding, coancestry and genetic contributions

We will first summarize the most basic concepts and tools in the analysis of genealogies and breeding systems. Let F_x be the inbreeding coefficient of individual x, the probability of identity by descent of the two genes carried by this individual at a given locus, and f_{xy} the coancestry (kinship; Malécot, 1948) between individuals x and y, the probability of identity by descent of two genes taken at random from each individual at the locus (with replacement, if taken from the same individual). Consider individuals p

(with parents w and x and q (with parents y and z). Then,

$$f_{wx} = F_p, \quad f_{yz} = F_q,$$

$$f_{pq} = \frac{1}{4}(f_{wy} + f_{wz} + f_{xy} + f_{xz}) \tag{1}$$

and

$$f_{pp} = \frac{1}{2}(1 + F_p) = \frac{1}{2}(1 + f_{wx})$$

$$= \frac{1}{4}(f_{ww} + f_{wx} + f_{xw} + f_{xx}) + \frac{1}{2} - \frac{f_{ww}}{4} - \frac{f_{xx}}{4}$$

$$= \frac{1}{4}(f_{ww} + f_{wx} + f_{xw} + f_{xx}) + \frac{1}{4}\left(1 - \frac{F_w + F_x}{2}\right). \tag{2}$$

The first term in the final expression of (2) is equivalent to (1), regarding the Mendelian flow of genes from parents to offspring. The second term in (2) relates to the Mendelian sampling of genes with a correction for the way in which the mean level of inbreeding reduces it.

Consider now a given pedigree. The full additive relationship matrix **A** is a symmetrical matrix relating all individuals in the pedigree such that the element for individuals i and j is $a_{ij} = 2f_{ij}$. The average pairwise coancestry of a given group N of individuals of the pedigree including reciprocals and self-coancestries is then

$$\bar{f} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}/2N^2. \tag{3}$$

It is well known (Henderson, 1976; Thompson, 1977; Wray & Thompson, 1990) that Matrix **A** can be obtained as

$$\mathbf{A} = \mathbf{Z}\,\mathbf{D}\,\mathbf{Z}', \tag{4}$$

where **Z** is a triangular matrix that describes the flow of genes from one individual to the descendants, incorporating elements from (1) and the first term in (2). The jth column of this matrix gives the proportion of genes (genetic contributions, c_i; James, 1962) contributed by all previous ancestors (i) to individual j (including itself). For example, if i is a parent of j, $z_{ij} = 0\cdot5$; if i is a grandparent of j, $z_{ij} = 0\cdot25$; etc. **D** is a diagonal matrix with elements as the second term in (2), such that $d_i = 1$ if both parents of animal i are unknown (founders), or $d_i = 0\cdot5[1 - (F_{Si} + F_{Di})/2]$ otherwise, where F_{Si} and F_{Di} are the inbreeding coefficient of the sire and dam of i, respectively.

From expressions (3) and (4) it follows immediately that, in a pedigree with a total of M individuals, where N_0 are founders, the average pairwise coancestry of a given group of N individuals (for example the current cohort of individuals in the pedigree) is

$$\bar{f} = \frac{1}{2N^2}\sum_{i=1}^{N_0} c_i^2 + \frac{1}{4N^2}\sum_{i=N_o+1}^{M} c_i^2\left(1 - \frac{F_{Si} + F_{Di}}{2}\right), \tag{5}$$

where the first term is the sum of genetic contributions from founders (N_0) to the current cohort of individuals, and the second term is the sum of genetic contributions from non-founders ($M - N_0$, including the current cohort) to the current cohort of individuals. Note that $\sum_{i=1}^{M} c_i$ could be called the *number of discrete-generation equivalents*, because in a pedigree with non-overlapping generations it will obviously equal to the number of generations (Woolliams & Mäntysaari, 1995).

Expressions (4) and (5) are general relations applicable to any type of pedigree, but in order to arrive at further generalizations let us assume that the population is a single undivided one, with discrete generations, constant population size and a regular breeding system. Generation 0 is the generation of unrelated founders, and each generation the population consists of N individuals. Because the population size is assumed to be constant over generations, the mean contribution of ancestors in generation k to descendants in generation t is one, with variance

$$V_{k,t} = \sum_{i=1}^{N} \frac{c_{i(k,t)}^2}{N} - 1. \tag{6}$$

After a few generations all descendants will have the same contribution from a particular ancestor but the contributions will differ between ancestors, with a variance $V_{k,\infty}$, the variance of long-term contributions (Wray & Thompson, 1990). Although the long-term state is asymptotically approached, in practice it is approximately reached in a few generations for small population sizes, except when the amount of non-random mating is large. This can be seen from the sequence of terms in (16) below.

Thus, with the above assumptions, the average pairwise coancestry of the population at generation t is, from (5) and (6),

$$\bar{f}_t = \frac{1}{2N}(1 + V_{0,t}) + \frac{1}{4N} \sum_{k=1}^{t} (1 + V_{k,t})(1 - \bar{F}_{k-1}), \tag{7}$$

where $\bar{F}_k$ is the average coefficient of inbreeding in generation k. The above equation shows that the average coancestry of the actual population depends, on the one hand, on the variance of contributions from founders ($V_{0,t}$) and, on the other, on the variance of contributions from non-founders ($V_{k,t}$ with $k > 0$). However, as the non-founders may be related, their expected contributions can be redundant (in the absence of mutation), and have to be weighted by the inbreeding coefficient in the previous generation.

3. Effective population size

The effective size of a population (Wright, 1931) is defined as the size of an idealized population which would give rise to the rate of inbreeding (ΔF), or the rate of change in variance of gene frequencies ($\Delta V[q]$) observed in the population under consideration:

$$N_{eI} = \frac{1}{2\Delta F} \quad \text{or} \quad N_{eV} = \frac{1}{2\Delta V(q)} = \frac{1}{2\Delta f}, \tag{8}$$

which correspond to the so-called inbreeding and variance effective sizes, respectively (see Kimura & Crow, 1963a; Caballero, 1994). The right-hand side equality in (8) arises because the variance of gene frequencies is related to the average coancestry by

$$V(q_t) = \bar{f}_t q(1 - q) \tag{9}$$

(Cockerham, 1969). Thus, the inbreeding effective size measures the rate of increase in inbreeding and the variance effective size measures the rate of increase in coancestry.

In a single population randomly mated, inbreeding (F) and coancestry (f) coefficients will increase with generations such that $\bar{f}_t = \bar{F}_{t+1}$ (or $\bar{F}_{t+2}$ without self-fertilization). In non-random mating populations there may be a larger delay between $\bar{f}$ and $\bar{F}$. The degree of non-random mating is measured by the correlation of genes within individuals relative to the correlation of genes taken at random from the population (α). This coefficient gives an indication of the degree of deviation from Hardy–Weinberg proportions, and it is related to the previous inbreeding coefficients by

$$(1 - F) = (1 - f)(1 - \alpha) \tag{10}$$

(Wright, 1969). In a regular breeding system, α soon reaches an asymptotic value which mainly depends on the proportion of inbred matings (see e.g. Caballero & Hill, 1992), and although $\bar{f}$ and $\bar{F}$ can be very different in a given generation, their rates of increase will eventually converge to the same value so that $N_{eI} = N_{eV}$. Only in situations such as when the population is subdivided permanently in independent sublines with completely different pedigrees, or when the population is decreasing or increasing in size, will N_{eV} and N_{eI} differ permanently. Otherwise, they will be the same after a number of generations. Thus, we will not generally make a distinction between these two parameters.

The concept of effective size has usually an asymptotic meaning in a regular system, and it is more frequently used for predictive purposes rather than for analysing realized genealogies. However, we can still use the concept to understand the relationship with the pedigree tools explained above and those that will be explained below. In the context of genealogical analysis we can consider the increase in average coancestry between the founder generation and a given generation t,

$$\Delta f_{0,t} = (\bar{f}_t - \bar{f}_0)/(1 - \bar{f}_0).$$

Noting that $\bar{f}_0 = 1/2N$ if founders are unrelated, and using (7).

$$\Delta f_{0,t} \approx \frac{2V_{0,t} + \sum_{k=1}^{t}(1+V_{k,t})(1-\bar{F}_{k-1})}{4N}.$$

Equating this to the expected rate of increase in average coancestry over t generations in an ideal population of size N_e, i.e. $t/2N_e$, we obtain

$$N_e \approx \frac{2Nt}{2V_{0,t} + \sum_{k=1}^{t}(1+V_{k,t})(1-\bar{F}_{k-1})}. \tag{11}$$

Note that this is a linear approximation and, therefore, only gives accurate predictions of N_e if $\bar{F}_{k-1}$ is small.

In (11) we can replace the term $(1-\bar{F}_{k-1})$ by $(1-\bar{f}_{k-1})(1-\alpha)$ using (10). If we assume that the population size is not very small (say $N > 10$), so that second-order terms in N can be neglected, the factor $(1-\bar{f}_{k-1})$ can be ignored when k is not too large. Thus, to a good approximation,

$$N_e \approx \frac{2Nt}{2V_{0,t} + \sum_{k=1}^{t}(1+V_{k,t})(1-\alpha)}. \tag{12}$$

This is an expression for the effective size considering the variance of family sizes over consecutive generations. For a single generation ($t = 1$) under random mating ($\alpha = 0$), (12) reduces to the classical expression for the effective size considering a single generation (Wright, 1969),

$$N_e = \frac{2N}{1+2V_{0,1}} = \frac{4N}{2+S^2_{(1)}}, \tag{13}$$

because $V_{0,1} = S^2_{(1)}/4$, where $S^2_{(1)}$ is the variance of the number of offspring per parent and $V_{1,1} = 0$. For $t = 2$, noting that $V_{0,1} = S^2_{(1)}/4$, $V_{0,2} = S^2_{(2)}/16$ (in general, $V_{0,t} = S^2_{(t)}/4^t$), and $V_{2,2} = 0$, (12) reduces to that derived by Wray *et al.* (1990),

$$N_e = \frac{4N}{2+(S^2_{(2)}/8)+(S^2_{(1)}/4)}, \tag{14}$$

where $S^2_{(1)}$ is again the variance of the number of offspring from parents, and $S^2_{(2)}$ is the variance of the number of grandoffspring from grandparents.

After a number of generations ($t \to \infty$, but only a few in practice) the variance of contributions reaches an asymptotic value ($V_{0,\infty}$) and (12) becomes the equation derived by Wray & Thompson (1990) for random mating ($\alpha = 0$), and generalized by J. Woolliams (unpublished result; see Woolliams & Thompson, 1994, p. 129) for non-random mating, i.e.

$$N_e \approx \frac{2N}{(1+V_{0,\infty})(1-\alpha)}. \tag{15}$$

In a population where the same breeding system is applied regularly every generation, with absence of selection, and with a proportion β of inbred matings (random mating otherwise), the variance of family sizes over generations is increased with time by two causes: the build up of covariances between the numbers of descendants (see Wray *et al.*, 1990), and the correlation between mates (see Caballero & Santiago, 1995). Thus,

$$V_{k,t} =$$

$$\frac{S^2_{(1)}}{4}\left[1 + \frac{1+\beta}{2} + \left(\frac{1+\beta}{2}\right)^2 + \left(\frac{1+\beta}{2}\right)^3 + \cdots + \left(\frac{1+\beta}{2}\right)^{t-1-k}\right]$$

$$= \frac{S^2_{(1)}}{2(1-\beta)}\left[1 - \left(\frac{1+\beta}{2}\right)^{t-k}\right]. \tag{16}$$

Under random mating ($\beta = 0$), for Poisson distribution of family size, $S^2_{(1)} = 2$, and substituting the above into (12), we obtain $N_e = N$, as expected. For $k = 0$ and $t \to \infty$,

$$V_{0,\infty} = \frac{S^2_{(1)}}{2(1-\beta)}. \tag{17}$$

In the case of a proportion β of full-sib matings, $\alpha = \beta/(4-3\beta)$ and

$$V_{0,\infty} = S^2_{(1)}(1+3\alpha)/2(1-\alpha),$$

and substituting into (15),

$$N_e \approx \frac{4N}{2(1-\alpha)+S^2_{(1)}(1+3\alpha)}, \tag{18}$$

which was obtained by Caballero & Hill (1992). For partial selfing with proportion β, $\alpha = \beta/(2-\beta)$ and substituting into (17) and (15) we obtain again (18) with a term $(1+\alpha)$ instead of $(1+3\alpha)$ (Kimura & Crow, 1963*a*).

4. Effective number of founders

One of the main parameters proposed for the analysis of pedigrees is the effective number of founders (N_{ef}). This was defined by Lacy (1989) and Rochambeau *et al.* (1989) as the number of equally contributing founders that would be expected to produce the same genetic diversity as in the population under study. In our notation,

$$N_{ef} = \frac{1}{\sum_{i=1}^{N}\left(\frac{c_{i(0,t)}}{N}\right)^2}, \tag{19}$$

where the summation is for the contributions of founders in generation 0 to descendants in t. In terms of coancestries,

$$N_{ef} = 1 \bigg/ \sum_{i=1}^{N}\left(2\bar{f}_{i,t}\right)^2,$$

where $\bar{f}_{i,t}$ is the average coancestry between founder i

and descendants in generation t. From (19) and using (6) we note that

$$N_{ef} = \frac{N}{1 + V_{0,t}}. \tag{20}$$

Because, as was explained above, $V_{0,t}$ reaches an asymptotic value, N_{ef} will also become a constant value after a short number of generations and will not change thereafter irrespective of the pedigree. Therefore, a management programme based on the maximization of N_{ef} (Alderson, 1991) will be only partially effective in the initial generations, but completely ineffective thereafter. We should note an important point here. In a regular system and after a number of generations, N_{ef} equals half the asymptotic effective population size (cf. (15) with $\alpha = 0$), as was shown by Wray & Thompson (1990). Thus, *a posteriori*, N_{ef} reflects the rate of increase in inbreeding. However, in order to minimize the rate of inbreeding *a priori* the procedure should be based on minimizing variances of contribution from all generations, not only that of founders.

The concept of N_{ef} can be related to that of genetic diversity (*GD*) or expected heterozygosity, a common measure of genetic variation (Nei, 1973). In a gene dropping analysis two distinct alleles (founder genes) are assigned to each founder, i.e. there are $2N$ different alleles at generation t. We can then define a measure of genetic diversity at generation t as

$$GD_t^* = 1 - \sum_{n=1}^{2N} E^2(q_{n,t}),$$

where $E(q_{n,t})$ is the average value over replicates of the frequency of the nth founder allele at generation t. Given that each founder has two founder alleles, $q_{n,0} = 1/2N$,

$$\sum_{n=1}^{2N} E^2(q_{n,t}) = \sum_{i=1}^{N} c_{i(0,t)}^2 / 2N^2,$$

and using (19),

$$GD_t^* = 1 - \frac{1}{2N_{ef}}. \tag{21}$$

As we can see, N_{ef} gives a measure of genetic diversity based on average frequencies over replicates, and ignoring variation within replicates. Note also that $2N_{ef}$ is the effective number of alleles defined by Crow & Kimura (1970, p. 324). Because, as was explained above, N_{ef} becomes constant after a number of generations, *GD** becomes a very poor descriptor of genetic diversity, and a wider definition is necessary (see below).

5. Founder genome equivalents

A limitation of the previous concept is that it does not take into account loss of genetic variability by genetic drift in subsequent generations. This ignorance is particularly important in small conservation programmes with potential bottlenecks in the pedigree. To overcome this problem, Lacy (1989, 1995) introduced the concept of founder genome equivalents that, referring to generation t, is

$$N_{ge} = \frac{1}{2\bar{f}_t}. \tag{22}$$

(i) Founder genome equivalents and contributions

Lacy (1995) recognized that

$$N_{ge} \approx N_e/t, \tag{23}$$

showing that N_{ge} decreases as the generation number increases. From (22) and using (7) we immediately obtain that

$$N_{ge} = \frac{2N}{2 + 2V_{0,t} + \sum_{k=1}^{t} (1 + V_{k,t})(1 - \bar{F}_{k-1})}. \tag{24}$$

Comparing (24) with (11) we obtain the approximate relation (23).

Originally, Lacy (1989) defined N_{ge} in a way different from (22):

$$N_{ge} = \frac{1}{\sum_{i=1}^{N} (c_{i(0,t)}/N)^2 / r_{i(0,t)}}, \tag{25}$$

where $r_{i(0,t)}$ is the expected proportion of surviving alleles at generation t from the ith founder at generation 0 (what is called the allele retention). This parameter can be calculated by probability theory (Thomas & Thompson, 1984; Thompson, 1986), but it is more frequently done by gene dropping analysis (MacCluer *et al.*, 1986). More recently, Lacy (1995) proposed abandoning this definition of N_{ge} and staying with the above definition as half the inverse of average coancestry (22). A similar conclusion has been strongly supported by Lindgren *et al.* (1996), who refer to N_{ge} as the status number. To understand the relation between (22) and (25), note that the expression in the summation of (25) can be written as $(2\bar{f}_{i,t})^2 / r_{i(0,t)}$, where $\bar{f}_{i,t}$ is the average coancestry between ancestor i of generation 0 and descendants in generation t. Therefore, the denominator in (25) is the expected proportion of genes from founder i which are shared by individuals in generation t, the weights $r_{i(0,t)}$ being a correction accounting for the loss of genes from i in the pedigree. Expression (25) is a valid approximation only to twice the average coancestry among descendants in large pedigrees (Lacy, 1995), but they generally differ. In any case and, from a practical point of view, (22) should be used instead of (25), as it does not require complex probability calculations or Monte Carlo pedigree simulations but can readily be obtained from the additive relationship matrix.

Recalling expression (20) we can partition (24) into

$$\frac{1}{N_{ge}} = \frac{1}{N_{ef}} + \frac{1}{N_{enf}}.$$

Thus, N_{ge} has two components: N_{ef}, given by (20), that depends on the contributions of founders to the actual population; and N_{enf}, that we will call the *effective number of non-founders*,

$$N_{enf} = \frac{2N}{\sum\limits_{k=1}^{t} (1+V_{k,t})(1-\bar{F}_{k-1})},$$

which accounts for the contribution of non-founders, and is accumulating as time proceeds.

(ii) *Founder genome equivalents and allele frequencies*

In the framework of gene dropping analysis a measure of genetic diversity more complete than that from (21) is given by

$$GD_t = 1 - E\left(\sum_{n=1}^{2N} q_{n,t}^2\right) = 1 - \sum_{n=1}^{2N} E(q_{n,t}^2) = 1 - \bar{f}_t$$

$$= 1 - \frac{1}{2N_{ge}}, \tag{26}$$

$q_{n,t}$ being the frequency of allele n at generation t in a given simulation replicate, and the expectation is over replicates. As it is expressed by (26), N_{ge} can be calculated directly from pedigree information, without requiring gene drop analysis, as used recently by Boichard *et al.* (1997) and Sölkner *et al.* (1998).

The diversity measure from (26) is based on the value over replicates of the real heterozygosities by descent observed in each run and, therefore, it takes into account genetic drift; while in GD^* from (21), the heterozygosity by descent was calculated using expected values, over replicates, of gene frequencies. The difference between these quantities,

$$GD_t^* - GD_t = \sum_{n=1}^{2N} E(q_{n,t}^2) - \sum_{n=1}^{2N} E^2(q_{n,t}) = 2N\overline{V(q_t)}, \tag{27}$$

is a measure of the average variance of allelic frequencies, indicating that the effective number of non-founders measures precisely the amount of genetic drift that has occurred during the history of the population since its foundation.

Note that because $GD_t = 1 - \bar{f}_t$ and $E(q_{n,t}) = q = 1/2N$, then the expected $GD_t^* = 1 - 1/2N$, and from (27), $\overline{V(q_t)} = (\bar{f}_t/2N) - (1/4N^2)$. Substituting the expected frequency, $q = 1/2N$, into the above we obtain

$$\overline{V(q_t)} = \frac{2N\bar{f}_t - 1}{2N - 1} q(1-q). \tag{28}$$

This is not exactly the same as (9) for the following reason. Expression (28) refers to a founder population

where alleles are fixed (for a given pedigree, founder alleles are fixed), while (9) refers to a base population where alleles are randomly assigned. Therefore, (9) assumes an additional sampling process.

Another interesting partition that can be made is in genetic diversity (or coancestry) within and between individuals. From (26) we note that (removing generation subscripts for clarity),

$$\bar{f} = E\left(\sum_{n=1}^{2N} q_n^2\right) = E\left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{n=1}^{2N} q_{n,i}^2\right)$$

$$+ E\left(\frac{1}{N^2}\sum_{i \neq j=1}^{N}\sum_{n=1}^{2N} q_{n,i} q_{n,j}\right)$$

$$= E\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{n=1}^{2N} q_{n,i}^2\right) - E\left(\frac{1}{2N^2}\sum_{i \neq j}^{N}\sum_{n=1}^{2N} (q_{n,i} - q_{n,j})^2\right).$$

The first term of the final expression represents the coancestry within individuals, $\bar{f}_N$. The second one represents the coancestry between individuals, or the distance between gene frequencies among individuals, D, a measure of the degree of genetic differentiation (Nei, 1973). Thus, $\bar{f} = \bar{f}_N - D$. Denoting

$$G = (\bar{f}_N - \bar{f})/(1-\bar{f}) = D/(1-\bar{f}),$$

it is straightforward to show that

$$(1-\bar{f}_N) = (1-\bar{f})(1-G).$$

Noting that $f_N = \frac{1}{2}(1+F)$ (see expression (2)), and using (10) we obtain

$$G = \frac{1+\alpha}{2}. \tag{29}$$

Thus, we can partition the total diversity, $GD = 1 - \bar{f}$, into the diversity within individuals, $GD_W = 1 - \bar{f}_N$, and the diversity between individuals,

$$GD_B = GD - GD_W = \bar{f}_N - \bar{f} = D,$$

so that from the above relationships, $G = GD_B/GD$ is the proportion of diversity between individuals, and $1 - G = GD_W/GD$ is the proportion within individuals. Note that from (29) we observe that with random mating ($\alpha = 0$) it is expected that $G = 1/2$, so half the genetic diversity is within and half between individuals. With complete inbred matings ($\alpha = 1$) all genetic diversity is between individuals, as expected.

In the context of gene dropping it is assumed that all founders carry different alleles, and the above expressions of GD refer to heterozygosities by descent. In general, however, we can consider a marker locus with two or more alleles, where the initial frequency of allele i in the base population is p_i. The average coancestry estimated with this marker, $\bar{f}_m$, relates to the coancestry from the pedigree, $\bar{f}$, by

$$1 - \bar{f}_m = (1 - \sum p_i^2)(1-\bar{f}), \text{ and}$$

$$\overline{V(p)} = (1/n_a)\bar{f}(1-\sum p_i^2),$$

where n_a is the number of alleles. For two alleles, the latter expression reduces to (9), as expected.

6. Genetic management of conservation programmes

The objective of genetic management is the preservation of the genetic variation of the population from which the founders were drawn, as well as to deliver worthwhile improvement on it. Although several simple rules deduced from effective size theory are widely accepted, such as equalizing of sex ratio and family sizes, and avoidance of fluctuations in population size, there seems to be no general agreement on the best strategy to follow regarding other matters.

In the genetic management of a close captive population there are two main decisions that have to be taken: the first is how to select the animals that will contribute gametes to the next generation; the second is how the matings will be organized. In what follows we discuss the recommendations for selection decisions and systems of mating in the light of the theoretical relations made previously.

(i) *Choice of breeding individuals and their offspring contribution*

From the considerations presented in the preceding section it can be concluded that minimizing the average coancestry of individuals seems to be the most straightforward way of maintaining genetic variability (Lacy, 1995; see (26)). The method of choice of breeding individuals should be based on minimizing the average coancestry among the reproductive individuals weighted by their contributions to the next generation. For the sake of simplicity let us consider a monoecious population of N individuals in generation t. The number of progeny that each of these individuals should contribute to generation $t+1$ must be such that

$$\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j f_{ij} \quad \text{is minimum,} \tag{30}$$

where w_i is the contribution from the ith individual and

$$\sum_{i=1}^{N} w_i = N,$$

in order to maintain a constant population size.

The reasons for recommending this strategy can be summarized as follows.

1. It is intuitively appealing because minimizing the average coancestry in generation $t+1$ maximizes the population genetic diversity in terms of expected heterozygosity (26).

2. If the individuals of generation t are unrelated, or relations are uniform among individuals, the technique will minimize $\Sigma_{i=1}^{N} w_i^2$, that is, to equalize family size and, therefore, minimize the effective size in a single generation (13).

3. If individuals at generation t are related this criterion will maximize the effective population size expressed by (12). Maximization of N_e implies not only the equalization of contributions from founders (generation 0) to generation $t+1$, but also the contributions from all previous generations to generation $t+1$, i.e. from generation 1 to $t+1$, 2 to $t+1$, and so on. We may note that this is equivalent to minimizing the number of individuals in generation $t+1$ with common ancestors, because minimizing $V_{t,t+1}$ will minimize the number of full-sibs in generation $t+1$; minimizing $V_{t-1,t+1}$ will minimize the number of cousins in generation $t+1$; and so on.

The idea of minimizing coancestry in order adequately to manage genetic variability was initially proposed by Lacy (1995) and Ballou & Lacy (1995) in the context of conservation programmes, and by Wray & Goddard (1994) and Brisbane & Gibson (1995) in the context of animal breeding. In the same way, the idea of unequal contributions of selected individuals to the next generation in order to minimize genetic drift constitutes the basis of the weighted selection strategy (Toro & Nieto, 1984) that has been shown to be advantageous both by simulation (Toro *et al.*, 1988; Wray & Goddard, 1994; Meuwissen, 1997; Grundy *et al.*, 1998) and by experimental work with *Drosophila* (Nieto *et al.*, 1986; Sánchez *et al.*, 1999). It has also been recommended in the guidelines of the FAO (1998) (see also Oldenbroek, 1999).

Minimizing coancestry compares favourably with other strategies. Alderson (1991) proposed the equalization of founder contributions, i.e. maximization of N_{ef}. However, although much attention has been given to this strategy (e.g. Folch & Jordana, 1998), it should not be recommended because it minimizes the variance of contributions from founders but not from non-founders. Minimizing coancestry also gains over simpler techniques such as equalizing family sizes (Gowe *et al.*, 1959; Wang, 1997*b*), because it takes into account that individuals of the parental generation could be related and, therefore, that their contributions could be redundant (Ballou & Lacy, 1995; Montgomery *et al.*, 1997). However, after a number of generations when the relation among individuals becomes uniform, minimum coancestry will simply be the equalization of family sizes. Minimum coancestry also has advantages over other simple rules such as giving breeding priority to animals with the highest probability of carrying unique alleles,

or to animals with the lowest representation in the descent population (see Ballou & Lacy, 1995; Lacy, 1995).

Ballou & Foose (1995) proposed calculating a target founder contribution,

$$TF_i = \frac{r_{i(0,t)}}{\sum\limits_{i=1}^{N} r_{i(0,t)}}, \tag{31}$$

where $r_{i(0,t)}$ is the allele retention, or expected proportion of founder i's alleles that have survived to generation t, and preferentially breeding individuals from founders whose contribution currently fall below their targets in order to shift the observed founder contributions towards the target founder contributions. Thus N_{ge} expressed by (25) would be maximized when $c_{i(0,t)}/N = TF_i$, and substituting (31) into (25),

$$N_{ge} = \sum_{i=1}^{N} r_{i(0,t)},$$

i.e., N_{ge} equals what Lacy (1989) terms the number of founder genomes surviving. This strategy relates directly to that of minimizing coancestry (Ballou & Lacy, 1995). However, to evaluate target founder contributions it is necessary to calculate the allelic retention of founders, and this has to be done by probability theory or Monte Carlo simulation methods (gene dropping analysis). By contrast, average coancestries are straightforwardly obtained from the pedigree.

(ii) *Choice of mating system in a conservation programme*

The choice of the mating system in a conservation programme is less simple because it depends on the time scale of interest and other circumstances, such as the capacity of the species to cope with the effects of inbreeding. As stated by Lindgren & Mullin (1998), founder genome equivalents (status number in their nomenclature) depends only on relatedness in the concerned population, not how gametes unite. Thus, minimization of the average coancestry in a given generation $(t+1)$, as from condition (30), can be made irrespective of the system of mating in the previous one, t. However, the system of mating used will affect the average coancestry in generation $t+2$ and, hence, the founder genome equivalents.

In the short term, it is obvious that forcing mating between relatives will increase the average inbreeding in the population, and the opposite will occur if matings between relatives are avoided. However, in the long term, the effects can be the same or the opposite depending on the circumstances. This has been shown repeatedly in the literature (Kimura & Crow, 1963*b*; Robertson, 1964; Cockerham, 1969; Caballero, 1994; Wang, 1997*a*; Wang & Caballero, 1999), but there still seems to be misunderstandings in the conservation genetics literature about this issue. For example, Ballou & Lacy (1995) state that the Maximum Avoidance of Inbreeding (MAI) system proposed by Wright (1921) 'represents a strategy that maximizes the inbreeding effective population size'. However, if the variance of family size ($S^2_{(1)}$ in (18) is small (in MAI it is zero), avoidance of inbred matings will decrease, rather than increase, the effective population size, as explained below.

As can be deduced from the denominator of (18), an increase in α (forcing inbred matings) has opposite effects on the effective size. The first term in the denominator accounts for the genetic drift caused by segregation of heterozygotes. Because forcing inbred matings reduce the frequency of heterozygotes, this results in a decrease in the amount of drift occurring for this reason and, therefore, in an increase in the effective size. The second term of the denominator refers to the genetic drift due to the variable contribution from parents. Because forcing inbred matings increase the frequency of homozygotes, this results in an increase in the amount of drift occurring if some individuals have larger contributions to the offspring than others (large $S^2_{(1)}$), as homozygotes will pass only one type of allele to all their offspring. The results of these two antagonistic forces will depend on the actual value of $S^2_{(1)}$. If $S^2_{(1)}$ is small, such as in conservation programmes, where this variance will be intended to be as low as possible (zero with MAI), the first term can be more important, and forcing (avoiding) inbred matings will increase (reduce) the effective size.

Thus, although MAI will reduce or delay inbreeding in the short term, it will give higher inbreeding in the long term. The opposite effect will occur if matings between relatives are forced and there is no variance in family sizes. Inbreeding will be larger in the short term but smaller in the long term. However, because short-term inbreeding may have negative consequences in terms of inbreeding depression, avoidance of matings between relatives may be more appropriate. The maximum expression of this system of mating is through minimum coancestry matings, in which matings occur among the least related animals. This can generally be implemented using linear programming techniques (Toro *et al.*, 1988).

Other systems of mating have been proposed for controlling inbreeding in the context of selected populations: for example, factorial mating designs (Woolliams, 1989), in which only half-sib and no (or few) full-sib families are obtained in the next generation, and compensatory mating (Santiago & Caballero, 1995), in which individuals from large families are mated to individuals of smaller families.

Simulations have shown that these two methods are useful for controlling inbreeding in selected populations (Villanueva *et al.*, 1994, and Grundy *et al.*, 1994, respectively). Compensatory mating can also be implemented by ordering males and females according to their average coancestry with all other individuals, and mating males with the highest average to females with the lowest (Caballero *et al.*, 1996). Its effectiveness in selected populations occurs because transmission lines of families with low and high selective success are mixed up (Santiago & Caballero, 1995). Ballou & Lacy (1995), however, have advised not using such a system of mating in unselected populations. According to them, this results in mixing rare and common alleles and, thereafter, the number of copies of the rare alleles cannot be increased without also increasing that of the over-represented alleles. They advise, on the contrary, mating individuals with similar average coancestries. Although compensatory mating is not expected *a priori* to be effective in non-selected populations, we do not see clearly why it should be disadvantageous. An ideal outcome of conservation management is that of each descendant having the same proportion of its genome from each founder and the maximum allele retention from these. Thus, mixing of rare and common alleles will have to occur in order to equalize contributions and to avoid loss of rare alleles. In fact, compensatory mating causes a slight decrease in the rate of inbreeding in non-selected populations (A. Caballero, unpublished results). On the contrary, when the reverse is carried out (males and females with similar average coancestries are mated to each other), a substantial increase in the rate of inbreeding is observed relative to random mating.

7. Discussion

We have tried to clarify some of the interrelations between effective population size and genetic tools frequently used in the management of conserved populations, such as the effective number of founders or the number of genome equivalents. The effective number of founders, N_{ef}, is a function of the expected contributions from founders to descendants. After a few generations all descendants will have the same contribution from a particular ancestor, so the variance of ancestors' contributions becomes constant, and the effective number of founders becomes useless.

The actual (not the expected) contribution from founders will depend on the Mendelian segregation occurring every generation in the pedigree. The number of genome equivalents, N_{eg}, represents a compound of contributions from founders (N_{ef}), and from all other individuals in the genealogy (the effective number of non-founders, N_{enf}), and relates directly to the effective population size. In terms of gene frequencies, the difference between the measure of genetic diversity expressed by N_{ef} and that represented by N_{ge} is the inclusion in the latter of the genetic drift occurring during pedigree development.

Boichard *et al.* (1997) calculated N_{ge} in a slightly different way from that in (26), as can be deduced from the example given in their table 2. This is

$$N_{ge} = (1/2) \, E\left[1 \Big/ \sum_{n=1}^{2N} q_n^2 \right],$$

i.e. half the average of the inverses, instead of half the inverse of the averages. This might be a mistake in the calculations. If not, the new definition does not seem to provide any additional information over (26). Boichard *et al.* (1997) also defined the concept of *effective number of ancestors* as the minimum number of ancestors (founders or not) necessary to explain the genetic diversity under study. This verbal definition coincides with that given initially by Lacy (1989) for the concept of founder genome equivalents (25). However, Boichard *et al.* (1997) proposed an approximate method to calculate it, based on computing the marginal contribution of an ancestor, i.e. the contribution not yet explained by other ancestors. Obviously, there is no need for such a complex procedure because, as previously shown, contributions from all founder and non-founder individuals corrected to avoid redundancies can be calculated in an exact way (5). Moreover, Boichard *et al.* (1997) showed that although the effective number of ancestors accounted for bottlenecks in the pedigree, it did not account for additional random losses of genes during the segregations, which are, however, accounted for by the number of genome equivalents.

By definition of genetic diversity, it is evident that the most advisable method of choice of parents is to minimize the average pairwise coancestry every generation and, by definition, to maximize the number of genome equivalents. From the relations derived in this paper, this is, in turn, the same as minimizing the variance of contributions from all previous generations to the current one, i.e. maximizing the effective population size. When comparing maximization of effective population size with other strategies, Ballou & Lacy (1995) concluded that 'Maximizing N_e might not be the most effective strategy for maintaining genetic diversity in populations with known pedigree. Quite possibly, a strategy that utilizes all the information contained within a pedigree could preserve genetic variation better than one that is based on maximization of N_e but ignores the ancestry of each individual'. This is, however, a by-product of using an incomplete prediction (or definition) of N_e. For example, Ballou & Lacy (1995) & Ballou & Foose (1995) use predictions of N_e (e.g. equation 26·4 of Ballou & Foose, 1995) equivalent to (13) in this paper, accounting for one generation of family sizes. When

predictions of N_e accounting for multiple generations are given (e.g. 12), maximization of N_e becomes equivalent to maximization of genetic diversity.

When only pedigree information is available, the most effective method is to minimize average group coancestries, as explained above. When molecular genetic markers are also available, this information can also be incorporated for estimating the true coancestry relationships. Coancestries should be calculated conditional on marker information (Toro *et al.*, 1999). A procedure to do this is through Monte Carlo Markov chains. The idea is to calculate the probability of identity by descent in a random point of the genome given the pedigree and the marker information. A similar suggestion has been made by Wang & Hill (2000).

In order to minimize the average coancestry among individuals (condition 30) several procedures can be followed. Ballou & Lacy (1995) suggested a recursive method implying selection not only of parents but also of their mating. However, we believe that a separation between the two procedures (choice of parents and system of mating) is advisable. As was explained above, the type of mating system to be used will depend on the time scale of interest and other issues, such as the potential inbreeding depression of the species. Wray & Goddard (1994), Wang *et al.* (1994) and Brisbane & Gibson (1995) proposed approximate methods for selection of individuals with a minimum coancestry in selected populations. Simulated annealing can be another way of obtaining approximate results, and an exact method can be implemented by quadratic or linear integer programming techniques (Fernández & Toro, 1999).

Minimum coancestry is the most effective method of maintaining genetic diversity when the complete pedigree is known. When individuals have unknown origin or there are uncertainties in their ancestry the efficiency of the method is diminished, and several procedures can be followed. The classical one is to assume that these individuals are founders. Other possibilities are to ignore them, or to consider only the known part of their genomes (Ballou & Lacy, 1995). Finally, probabilities of uncertain coancestry can also be calculated (Pérez-Enciso, 1990).

In this paper we have assumed that the population to be managed is a single one without subdivision into small subpopulations. Wang & Caballero (1999) have discussed the genetic consequences of population subdivision and migration in terms of genetic drift and effective population size. These and other non-genetic issues are also discussed by Lacy (1994).

The criterion given in (30) does not take into account the age structure of the population. With overlapping generations, Ballou & Lacy (1995) proposed calculating the coancestry value of an individual as the weighted mean of the coancestry coefficient between this individual and all members of the group weighted by the reproductive value of the last. The reproductive value is defined as the expected future lifetime production of progeny, assuming that animals perform demographically according to a life-table. Coancestry values will be lower than the mean if most of the relatives are early post-reproductive and greater if most of them are at good breeding age. Ballou & Lacy (1995) assumed a fixed age structure but it would also be possible to simultaneously optimize average coancestry and other demographic parameters. It is not clear which would be the optimal criterion in such a case but it can be conjectured that it will be of the form $\Sigma_i \Sigma_j w_i f_{ij} w_j / g_{ij}$, where g_{ij} is the average age of individuals i and j. In fact, Wang *et al.* (1994) simulated a breeding scheme comparing several selection indices, and those indices combining the age of the selected individuals and their coancestry to other individuals of the population were clearly superior.

We have also assumed that the only neutral genetic variability to be maintained is that originally from the base population, and there is no input of new mutations. Assuming an infinitesimal model of mutations, it is also possible to include this new diversity into the coancestry matrix (Wray, 1990). Nevertheless, it is likely that a general procedure for minimizing coancestry will also effectively maintain genetic diversity arisen by neutral mutation, by minimizing the chances of allelic loss through genetic drift.

By equalizing contributions from parents, the intensity of natural selection is reduced and one could argue that deleterious mutations would accumulate more frequently in the population than if differences in contributions among parents were allowed. A recent theoretical investigation by Schoen *et al.* (1998) comparing a method of maintaining plant seeds that allows for differential contributions from parents versus others in which contributions are equalized showed that the expected decline in fitness can be much larger with the latter. This throws into question the practice of equalizing contributions in conservation programmes. The higher genetic variability maintained and the lower inbreeding depression allowed when contributions are equalized can be offset by a larger decline in reproductive performance due to the weaker purging of deleterious genes, both from the base population, or arising as new mutations. However, the analytical studies and simulations by Schoen *et al.* (1998) do not allow for natural selection within families, and consider an unrealistic model of mutations with constant effects and dominances. When selection within families is considered under mutational models of variable effects, the decline in fitness observed under equalization of family sizes is substantially reduced, and comes closer to that

occurred with variable family sizes (J. Fernández & A. Caballero, unpublished data).

Equalization of family sizes has been used by Shabalina *et al.* (1997) as a procedure to minimize selection and estimate the rate of input of mutational damage into fly populations with large census sizes. A decline in fitness was observed in this experiment and ascribed to accumulation of mutations. However, other causes may have contributed to this decline, such as inbreeding depression, increased fitness in the control along the experiment and, in particular, adaptation to captivity (see Keightley *et al.*, 1998; and also the related experiment by Gilligan *et al.*, 1997). Further, in this experiment the experimental population size was very large ($N_e \approx 400$), diminishing the advantageous effects of equalizing family sizes relative to the negative ones of mutation accumulation. In conservation programmes it is expected that population sizes are generally much lower. Loebel *et al.* (1992) and Borlase *et al.* (1993) have run experiments with low population sizes. In these experiments populations with random contributions from parents showed no difference, or even worse fitness performances, than those obtained by equalizing individuals or founder contributions.

Conservation of genetic resources is one important field of study in conservation biology. The objective of this paper has been to interrelate different genetic concepts widely used in this field and to discuss the general practical implications. The strategy to be applied, however, will depend on the particular species and genealogy to be handled, and the results and relations outlined in this paper should be used only as a general guide.

References

Alderson, G. H. L. (1991). A system to maximise the maintenance of genetic variability in small populations. In *Conservation of Domestic Livestock* (ed. L. Alderson & L. Bodo), pp. 18–19. Wallingford, UK: CAB International.

Ballou, J. D. & Foose, T. J. (1995). Demographic and genetic management of captive populations. In *Wild Mammals in Captivity* (ed. D. G. Kleinman, S. Lumpkin, M. Allen, H. Harris & K. Thompson), pp. 263–283. Chicago: University of Chicago Press.

Ballou, J. D. & Lacy, R. C. (1995). Identifying genetically important individuals for management of genetic diversity in captive populations. In *Population Management for Survival and Recovery* (ed. J. D. Ballou, M. Gilpin & T. J. Foose), pp. 76–111. New York: Columbia University Press.

Boichard, D., Maignel, L. & Verrier, E. (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genetics, Selection, Evolution* **29**, 5–23.

Borlase, S. C., Loebel, D. A., Frankham, R., Nurthen, R. K., Briscoe, D. A. & Dagard, G. E. (1993). Modeling problems in conservation genetics using captive *Drosophila* populations: consequences of equalization of family sizes. *Conservation Biology* **7**, 122–131.

Brisbane, J. R. & Gibson, J. P. (1995). Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theoretical and Applied Genetics* **91**, 421–431.

Caballero, A. (1994). Developments in the prediction of effective population size. *Heredity* **73**, 657–679.

Caballero, A. & Hill, W. G. (1992). Effective size of nonrandom mating populations. *Genetics* **130**, 909–916.

Caballero, A. & Santiago, E. (1995). Response to selection from new mutation and effective size of partially inbred populations. I. Theoretical results. *Genetical Research* **66**, 213–225.

Caballero, A., Santiago, E. & Toro, M. A. (1996). Systems of mating to reduce inbreeding in selected populations. *Animal Science* **62**, 431–442.

Cockerham, C. C. (1967). Group inbreeding and coancestry. *Genetics* **56**, 89–104.

Cockerham, C. C. (1969). Variance of gene frequencies. *Evolution* **23**, 72–84.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper & Row.

Food and Agricultural Organization (1998). *Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans. Management of Small Populations at Risk*. Initiative for Domestic Animal Diversity (IDAD). Geneva: FAO.

Fernández, J. & Toro, M. A. (1999). The use of mathematical programming to control inbreeding in selection schemes. *Journal of Animal Breeding and Genetics*, in press.

Folch, P. & Jordana, J. (1998). Demographic characterization, inbreeding and maintenance of genetic diversity in the endangered Catalonian donkey breed. *Genetics, Selection, Evolution* **30**, 195–201.

Gilligan, D. M., Woodworth, L. M., Montgomery, M. E., Briscoe, D. A. & Frankham, R. (1997). Is mutation accumulation a threat to the survival of endangered populations? *Conservation Biology* **11**, 1235–1241.

Gowe, R. S., Robertson, A. & Latter, B. D. H. (1959). Environment and poultry breeding problems. 5. The design of poultry control strains. *Poultry Science* **38**, 462–471.

Grundy, B., Caballero, A., Santiago, E. & Hill, W. G. H. (1994). A note on using biased parameter values and nonrandom mating to reduce rates of inbreeding in selection programmes. *Animal Production* **59**, 465–468.

Grundy, B., Villanueva, B. & Woolliams, J. (1998). Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genetical Research* **72**, 159–168.

Henderson, C. R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* **32**, 69–83.

James, J. W. (1962). The spread of genes in random mating control populations. *Genetical Research* **3**, 1–19.

Keightley, P. D., Caballero, A. & García-Dorado, A. (1998). Surviving under mutation pressure. *Current Biology*, **8**, R235–R237.

Kimura, M. & Crow, J. F. (1963*a*). The measurement of effective population number. *Evolution*, **17**, 279–288.

Kimura, M. & Crow, J. F. (1963*b*). On the maximum avoidance of inbreeding. *Genetical Research* **4**, 399–415.

Lacy, R. C. (1989). Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalence. *Zoo Biology* **8**, 111–124.

Lacy, R. C. (1994). Managing genetic diversity in captive populations of animals. In *Restoration of Endangered Species* (ed. M. L. Bowles & C. J. Whelan), pp. 63–89. Cambridge: Cambridge University Press.

Lacy, R. (1995). Clarification of genetic terms and their use in the management of captive populations. *Zoo Biology* **14**, 565–578.

Lindgren, D. & Mullin, T. J. (1998). Relatedness and status number in seed orchard crops. *Canadian Journal of Forest Research* **28**, 276–283.

Lindgren, D., Gea, L. D. & Jefferson, P. A. (1996). Status number for measuring genetic diversity. *Forest Research* **4**, 69–76.

Loebel, D. A., Nurthen, R. K., Frankham, R., Briscoe, D. A. & Craven, D. (1992). Modeling problems in conservation genetics using captive *Drosophila* populations: consequences of equalizing founder representation. *Zoo Biology* **11**, 319–332.

MacCluer, J. W., VandeBerg, J. L., Read, B. & Ryder, O. A. (1986). Pedigree analysis by computer simulation. *Zoo Biology* **5**, 147–160.

Malécot, G. (1948). *Les Mathématiques de l'Hérédité*. Paris: Masson et Cie.

Meuwissen, T. H. E. (1997). Maximizing the response of selection with a predefined rate of inbreeding. *Journal of Animal Science* **75**, 934–940.

Montgomery, M., Ballou, J. D., Nurthen, R. K., England, P., Briscoe, D. & Frankham, R. (1997). Minimizing kinship in captive breeding programs. *Zoo Biology* **16**, 377–389.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA* **70**, 3321–3323.

Nieto, B., Salgado, C. & Toro, M. A. (1986). Optimization of artificial selection response. *Journal of Animal Breeding and Genetics* **103**, 199–204.

Oldenbroek, J. K. (1999). Genebanks and the conservation of farm animal genetic resources. Lalystad, The Netherlands: DLO Institute for Animal Sciences and Health.

Pérez-Enciso, M. (1990). Use of the uncertain relationship matrix to compute effective population size. *Journal of Animal Breeding and Genetics* **112**, 327–332.

Robertson, A. (1964). The effect of nonrandom mating within inbred lines on the rate of inbreeding. *Genetical Research* **5**, 164–167.

Rochambeau, H. de, La Fuente, L. F. de, Rouvier, R. & Ouhayoun, J. (1989). Sélection sur la vitesse de croissance post-sevrage chez le lapin. *Genetics, Selection, Evolution* **21**, 527–546.

Rodrigañez, Toro, M. A., Rodríguez, C. & Silió, L. (1998). Effect of founder allele survival and inbreeding depression on litter size in a closed line of Large White pigs. *Animal Science*, **67**, 573–582.

Sánchez, L., Toro, M. A. & García, C. (1999). Improving the efficiency of artificial selection: more selection pressure with less inbreeding. *Genetics* **151**, 1103–1114.

Santiago, E. & Caballero, A. (1995). Effective size of populations under selection. *Genetics* **139**, 1013–1030.

Schoen, D. J., David, J. L. & Bataillon, T. M. (1998). Deleterious mutation accumulation and the regeneration of genetic resources. *Proceedings of the National Academy of Sciences of the USA* **95**, 394–399.

Shabalina, S. A., Yampolsky, L. Y. & Kondrashov, A. S. (1997). Rapid decline of fitness in panmictic populations of *Drosophila melanogaster*. *Proceedings of the National Academy of Science of the USA* **94**, 13034–13039.

Sölkner, J., Filipcic, L. & Hampshire, N. (1998). Genetic variability of populations and similarity of subpopulations in Austrian cattle breeds determined by analysis of pedigrees. *Animal Science* **67**, 249–256.

Thomas, A. & Thompson, E. A. (1984). Gene survival in an isolated population: the number of distinct genes on Tristan da Cunha. *Annals of Human Biology* **11**, 101–112.

Thompson, E. A. (1986). Ancestry of alleles and extinction of genes in populations with defined pedigrees. *Zoo Biology* **9**, 259–274.

Thompson, R. (1977). The estimation of heritability with unbalanced data. II. Data available on more than two generations. *Biometrics* **33**, 497–504.

Toro, M. A. & Nieto, B. M. (1984). A simple method for increasing the response to artificial selection. *Genetical Research* **44**, 347–349.

Toro, M. A., Nieto, B. M. & Salgado, C. (1988). A note on minimisation of inbreeding in small scale breeding programmes. *Livestock Production Science* **20**, 317–323.

Toro, M. A., Silió, L., Rodrigañez, J., Rodríguez, C. & Fernández, J. (1999). Optimal use of genetic markers in conservation programmes. *Genetics, Selection, Evolution* **31**, 255–261.

Villanueva, B., Woolliams, J. A. & Simm, G. (1994). Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics, Selection, Evolution* **26**, 517–535.

Wang, J. (1997*a*). Effect of excluding sib matings on inbreeding coefficient and effective size of finite diploid populations. *Biometrics* **53**, 1354–1365.

Wang, J. (1997*b*). More efficient breeding systems for controlling inbreeding and effective size in animal populations. *Heredity* **79**, 591–599.

Wang, J. & Caballero, A. (1999). Developments in predicting the effective size of subdivided populations. *Heredity* **82**, 212–226.

Wang, J. & Hill, W. G. (2000). Marker assisted selection to increase effective population size by reducing Mendelian segregation variance. *Genetics* **154**, 475–489.

Wang, J., Xu, J. C., Song, J. Z. & Tang, W. S. (1994). Study of mating systems for controlling inbreeding levels in domestic animal populations with overlapping generations. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **21**, 536–539.

Woolliams, J. A. (1989). Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Production* **49**, 1–14.

Woolliams, J. A. & Mäntysaari, E. A. (1995). Genetic contributions of Finnish Ayrshire bulls over four generations. *Animal Science* **61**, 177–187.

Woolliams, J. A. & Thompson, R. (1994). A theory of genetic contributions. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **19**, 127–134.

Wray, N. R. (1990). Accounting for mutation effects in the additive genetic variance–covariance matrix and its inverse. *Biometrics* **46**, 177–186.

Wray, N. R. & Goddard, M. E. (1994). Increasing long-term response to selection. *Genetics, Selection, Evolution* **26**, 431–451.

Wray, N. R. & Thompson, R. (1990). Predictions of rates of inbreeding in selected populations. *Genetical Research* **55**, 41–54.

Wray, N. R., Woolliams, J. A. & Thompson, R. (1990). Methods for predicting rates of inbreeding in selected populations. *Theoretical and Applied Genetics* **80**, 503–512.

Wright, S. (1921). Systems of mating. *Genetics* **6**, 111–178.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.

Wright, S. (1969). *Evolution and the Genetics of Populations*, vol. 2, *The Theory of Gene Frequencies*. Chicago: University of Chicago Press.