
Managing Individuals' Contributions to Maximize the Allelic Diversity Maintained in Small, Conserved Populations

J. FERNÁNDEZ,* †‡ M. A. TORO,* AND A. CABALLERO†

*Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

†Departamento de Bioquímica, Genética e Inmunología, Facultad de Ciencias, Universidad de Vigo, Vigo, Spain

Abstract: *The maintenance of diversity is, from a genetic perspective, one of the key aims in a conservation program. Because the most widely used measure of genetic diversity is the expected heterozygosity of the population, or gene diversity (GD), most research has been devoted to finding optimal strategies for maximizing this parameter. Little attention has been paid, however, to the development of strategies to manage allelic diversity (AD), the number of alleles maintained in the population. Using computer simulations, we show that the strategies that maximize GD, by managing contributions from parents, keep levels of AD as high as strategies maximizing AD itself, for a wide range of situations including different numbers of molecular markers used and the possibility of evaluating a number of offspring per parent to make decisions. Because maximization of GD also minimizes levels of inbreeding, this should be the strategy of choice in any conservation program.*

Key Words: expected heterozygosity, management strategies, molecular markers

Manejo de Contribuciones de Individuos para Maximizar la Diversidad Alélica Mantenido en Poblaciones Conservadas Pequeñas

Resumen: *El mantenimiento de la diversidad es, desde una perspectiva genética, una de las metas clave de un programa de conservación. Debido a que la medida de diversidad genética más utilizada es la heterocigosidad esperada en la población, o diversidad genética (DG), la mayor parte de la investigación se ha enfocado a encontrar estrategias óptimas para maximizar el parámetro. Sin embargo, no se ha puesto mucha atención al desarrollo de estrategias para manejar la diversidad alélica (DA), (i.e., el número de alelos en la población). Utilizando simulaciones en computadora mostramos que estrategias que maximizan DG, mediante el manejo de contribuciones de padres, mantienen niveles de DA tan altos como con estrategias para maximizar DA mismo, para un amplio rango de situaciones incluyendo el uso de diferentes números de marcadores moleculares y la posibilidad de evaluar a un número de descendientes por padre para tomar las decisiones. Debido a que la maximización de DG también minimiza los niveles de endogamia, esta debe ser la estrategia elegida en cualquier programa de conservación.*

Palabras Clave: estrategias de manejo, heterocigosidad esperada, marcadores moleculares

‡email jmj@inia.es

Paper submitted July 28, 2003; revised manuscript accepted January 6, 2004.

Introduction

The maintenance of high levels of genetic variability is a major objective in conservation programs. Genetic variation is a prerequisite for the population to be able to face future environmental changes and to ensure long-term response to selection, either natural or artificial, for traits of economic or cultural interest (Ballou & Lacy 1995; Oldenbroek 1999; Barker 2001).

To understand to what extent a population is threatened and to monitor the performance of a conservation program, we must be able to measure the amount of a species' available genetic variability. The classical method used to quantify genetic variability is expected heterozygosity (Nei 1973), usually called gene diversity (GD). This represents the expected proportion of heterozygotes if the population were in Hardy-Weinberg equilibrium. Usually, the only available information with which to calculate GD is that from neutral molecular markers. High levels of heterozygosity also imply high levels of additive genetic variance and, thus, greater potential responses to selection (Falconer & Mackay 1996).

From an evolutionary perspective, a straightforward measure of variability is allelic diversity (AD)—or allelic richness—the number of different alleles at a locus or the average over loci—present in the population. Some authors believe that this parameter is the most relevant criterion for measuring diversity because high levels of AD imply a source of single-locus variation for important qualitative traits (Schoen & Brown 1993; Bataillon et al. 1996; Petit et al. 1998). It is also essential for the long-term evolutionary potential of populations because the limit of selection response is determined by the initial number of alleles (assuming that mutation is negligible), regardless of the allelic frequencies (James 1971; Hill & Rasbash 1986).

Two well-known cases illustrate the direct relationship between high levels of allelic variation and higher probabilities of population survival. One is the self-incompatibility system of many flowering plants, avoiding self-fertilization. Individuals from populations with few alleles have reduced opportunity to find a suitable mate with which to reproduce (Young et al. 2000 and references therein). The other is the major histocompatibility complex, which is responsible for the recognition of pathogens. The lower the number of alleles at each gene of the complex, the lower the number of combinations and the lower the probability of populations surviving infections, because if an individual is susceptible to a pathogen then most members of the population will be susceptible too. Populations with reduced genetic variability often have a higher load of parasites and experience higher mortality than populations with greater diversity (Frankham et al. 2002 and references therein). As in the case of GD, there is usually only information on neutral molecular markers with which to calculate AD.

Different strategies have been proposed to maximize GD in managed populations. Chevalet (1992) suggested determining the contributions to the next generation by maximizing an index equal to the heterozygosity of individuals measured at several marker loci. Toro et al. (1998) proposed the implementation of frequency-dependent selection, in which homozygotes at rare alleles are favored until frequencies are equalized. In the absence of molecular information, the average pedigree coancestry measures the identity-by-descent, and the minimization of global pedigree coancestry (\bar{f}), weighted by the product of individuals' contributions, is the optimal strategy with which to maximize GD (Ballou & Lacy 1995; Fernández & Toro 1999; Caballero & Toro 2000). Furthermore, this strategy implies the equalization of contributions from all previous generations to that of the actual one; therefore, it also implies the maximization of effective population size (Caballero & Toro 2000, 2002). With the availability of markers, GD can be obtained as 1 minus the global molecular coancestry (\bar{f}_M), and the optimal strategy is, analogously, based on the minimization of \bar{f}_M weighted by individuals' contributions (Toro et al. 1999).

The dynamics of allelic diversity have been studied under different models. Ewens (1972) developed a formula (Ewens's sampling formula) that allows one to determine the distribution of the frequency of neutral alleles at a locus in a population when an equilibrium between drift and mutation (infinite-alleles model) has been reached. The effects of bottlenecks or processes of extinction and recolonization on genetic diversity (both allelic and gene diversity) have also been studied (e.g., Allendorf 1986; Cornuet & Luikart 1996; Luikart et al. 1998; Comps et al. 2001; England et al. 2003). Little attention has been paid, however, to strategies directed at maintaining AD. Engels (1980) developed predictions for the number and frequencies of neutral alleles under different mating regimes, but most efforts have been focused on the use of AD as a criterion by which to choose among conservation units (Petit et al. 1998) or as an indicator of population history (Richards & Leberg 1996; Luikart et al. 1998) because it is more sensitive to bottlenecks than GD. Because the loss of alleles in small populations is mainly driven by genetic drift (Falconer & Mackay 1996), any strategy directed at minimizing genetic drift should, a priori, keep the largest number of alleles. However, the relationship between GD and the actual number of alleles, AD, may not be straightforward. For example, large values of GD can be compatible with a low number of alleles because GD also depends on the allele frequencies. Therefore, the effectiveness of strategies developed to maintain GD should be assessed with respect to the maintenance of AD.

The increasing availability of neutral genetic markers in a number of species makes it possible to develop conservation strategies based on the parameters AD or GD. Our objective was to test, through computer simulations,

the efficiency of strategies, based either on AD itself or on GD, for the maintenance of genomic allele diversity of a small population under different scenarios.

Conservation Strategies Based on Molecular Markers

Neutral molecular markers can be used to make decisions about individuals' contributions to the next generation in order to maximize gene or allele diversity. However, the use of molecular markers introduces another factor in the design of conservation strategies. Genotyping markers is an expensive procedure, and a balance between the possible increase in accuracy and the rise of costs should be found when one has to decide the number of individuals and markers to be studied.

The particular characteristics of the species considered also influences the conservation scheme. In a multiparous species, several offspring per parent can be obtained with the same effort, and then breeding decisions can be made based on offspring data by selecting a determined number of genotyped individuals and discarding the rest. This procedure is not possible with most uniparous species, however, because the cost (in time and money) of getting individuals that will not be used is not affordable. In general, there are two groups of strategies to adopt when we want to use the information from molecular markers in the management of populations: genotyping parents or genotyping offspring.

Genotyping Parents

In the simplest situation, a manager deals with a group of individuals that have been typed at a number of molecular markers. Then the manager selects which of those individuals will contribute to the next generation and determines how many offspring they will leave. Depending on the parameter used to monitor the process, AD or GD, the strategy varies.

ALLELIC DIVERSITY

One direct approach to taking AD into account in management decisions would be choosing as parents a group of individuals in which all alleles of the population are represented. Therefore, if two individuals share the same alleles, one of them could be excluded from the breeding group. Moreover, this strategy identifies the parents that will generate offspring but tells us nothing about the particular contribution of each of them, and thus individuals carrying rare alleles are not favored. Therefore, the probability of effectively transmitting the alleles to the descendants is not controlled.

A more efficient approach proposed by Vales et al. (2003) is to calculate the set of parents' contributions that

Table 1. Example calculations of the expected number of alleles lost for a particular combination of offspring number obtained from each parent.^a

| Allele (i) | Parent (j) ^b | | | | Allele loss probabilities |
|---------------------------------|-------------------------|--------------|--------------|--------------|---------------------------|
| | 1 (6,6) 2 | 2 (1,2) 2 | 3 (1,3) 3 | 4 (4,5) 1 | |
| 1 | 1 | 0.25 | 0.125 | 1 | 0.031 |
| 2 | 1 | 0.25 | 1 | 1 | 0.250 |
| 3 | 1 | 1 | 0.125 | 1 | 0.125 |
| 4 | 1 | 1 | 1 | 0.5 | 0.500 |
| 5 | 1 | 1 | 1 | 0.5 | 0.500 |
| 6 | 0 | 1 | 1 | 1 | 0.000 |
| Expected number of alleles lost | | | | | 1.406 |

^aPopulation with $n = 4$ parents and a locus with 6 alleles.

^bNumber of parent, alleles (in parentheses), followed by the contributions from each of them.

maximizes the expected number of alleles, at a particular locus, that will be transmitted to the next generation. Because this strategy demands a huge number of calculations to evaluate each solution, an alternative procedure, less computationally demanding, is to minimize the expected number of alleles lost in the offspring (Vales et al. 2003).

An application example is given in Table 1 for $n = 4$ parents contributing 2, 2, 3, and 1 offspring, respectively. Parent 1 will necessarily contribute allele 6, and therefore this allele will not be lost (probability of loss equal to zero). With respect to this parent, however, all the remaining alleles will be lost (probability of loss equal to one). A similar argument can be used for the other parents. Multiplying all elements in the same row, we obtain the probability of no individual transmitting a particular allele. Thus, summing these marginal probabilities, we calculate that, for the set of contributions in the example, the expected number of alleles lost is 1.40625. One would operate in the same way for other combinations of contributions until the solution that minimizes the expected number of alleles lost is found.

Generalizing, the objective function we have to minimize is

$$\sum_{i=1}^A \prod_{j=1}^N p_{ij}, \quad (1)$$

where N is the population size, p_{ij} is the probability of individual j not transmitting allele i , and A is the total number of different alleles. If parent j carries a unique type of allele (k) and leaves descendants, p_{ij} is 0 if $i = k$ and 1 if $i \neq k$. If it carries two different alleles (k and m), the probability is $p_{ij} = (0.5)^{x_j}$ if $i = k$ or m and 1 if $i \neq k$ and m , where x_j is the variable of interest that determines the number of offspring contributed by parent j . If we use information from several loci, the value of the objective function for a specific solution is the average of Eq. 1

over all loci. To complete the optimization model, we must impose some constraints: only non-negative integer contributions are allowed ($x_j \in \mathbb{N}$), and the sum of all contributions must be $2N$ (N from each of the sexes).

GENE DIVERSITY

When the only available information is the genotype of parents and the focus is the control of GD, the expected heterozygosity in the offspring can be estimated by weighting the squared allelic frequencies of parents by the individual contributions. The optimal number of offspring per parent can be obtained by maximizing the following expression:

$$1 - \sum_{i=1}^L \sum_{j=1}^{A_i} \left[\frac{\sum_{k=1}^N n_{ijk} x_k}{(2N)^2} \right]^2, \quad (2)$$

where N is the population size, L is the number of loci considered, A_i the number of different alleles in locus i , n_{ijk} the number of alleles j carried by individual k in locus i (0, 1, or 2), and x_k the number of offspring contributed by individual k . Feasible solutions must fit the same kind of restrictions as in the previous case. This procedure is exactly the same as searching for the set of contributions that minimizes the weighted global molecular coancestry.

Genotyping Offspring

For genotyping offspring, every possible parent has an equal number of offspring that are genotyped for the markers (evaluated individuals). Then, the manager decides which individuals are used as parents in the following generation according to the chosen criterion.

There is no need for calculating probabilities for allelic diversity because the offspring have been generated before the contributions are decided. Thus, managers only have to select the group of individuals that collectively have the highest AD. Contributions from parents are optimized because different numbers of individuals can be taken from each sib family.

Accordingly, the group of individuals, from the available offspring, yielding the highest GD (calculated from the allelic frequencies on the typed loci) should be selected.

Computer Simulations

We considered populations with $N = 8$ or 16 parents, half of each sex, constant along generations. When genotyping and selection were done on offspring, the number of offspring tested per parent was either two or eight. The genome of individuals consisted of 1 or 20 chromosomes, with 100 evenly spaced multiallelic loci per chromosome.

When gametes were obtained, a Poisson-distributed random number of crossovers with a mean of one were generated in randomly chosen places on each chromosome without interference (simulated chromosomes 1 Morgan in length). In the initial populations we assumed that all individuals were unrelated and not inbred, so they carried two different alleles at each locus. Therefore, both AD and GD were at their maximum value ($2N$ and $1 - 1/2N$, respectively). However, to compare the effect of the amount of diversity present when the conservation program starts on the relative performance of the strategies and to allow for a more realistic situation where different relationships between individuals exist, 5 or 10 unmanaged generations were performed prior to the application of any management strategy. Thus, starting values of AD ranged from 18% to 30% of the initial 16 or 32 alleles per locus for $N = 8$ or 16, respectively, and values of GD ranged from 0.51 to 0.69.

Information from 0.5, 1, 2, 10, or all (100) loci per chromosome was used for the management decisions, simulating the genotyping of molecular markers. In the case of 0.5, 1, and 2 markers per chromosome—a total of 10, 20, or 40 markers in a genome of 20 chromosomes—markers were randomly distributed among and within chromosomes. Thus, when 20 chromosomes were simulated, there were 0.5, 1, or 2 markers per chromosome only on average. In the case of 10 markers, these were evenly spaced within each chromosome. The simulation using 100 markers per chromosome corresponds to the case of knowing the genotype for all loci of the genome; therefore, it establishes the upper theoretical limit of efficiency for the different strategies. Fifteen discrete generations were run under each of the different combinations of factors. Every generation, the AD and the expected GD and observed heterozygosity were calculated for the whole population with all loci (markers included) and averaged over 100 replicates.

Management Procedures Tested

Minimum pedigree coancestry: Every generation, the set of contributions from every available parent was chosen in a way that it yielded the minimum global pedigree coancestry weighted by the individual contributions (e.g., Fernández & Toro 1999). The minimization of the pedigree coancestry was simulated to represent a point of reference.

Allelic diversity in parents: In this strategy, the group of individuals selected from the available parents to generate the offspring was that with the highest AD, calculated from the corresponding markers. Particular contributions to the next generation were randomly assigned within this group.

Expected allelic diversity from parents (EAP): Contributions from individuals in the population were arranged to minimize the objective function (Eq. 1), the expected

number of alleles not transmitted to the offspring in the marker loci.

Gene diversity from parents (GP): The number of offspring per parent was obtained from the maximization of Eq. 2. This is the function that calculates the expected heterozygosity for the marker loci in the offspring by weighting allelic frequencies of parents by their contributions.

For the above three strategies, once the contributions were decided, a random mating design was implemented, allowing individuals to mate several times independently.

Allelic diversity in offspring (AO): In this strategy, the first step consisted of generating an equal number of offspring per parent (two or eight). The couples were randomly assigned, allowing for a design in which every parent could mate with more than one individual. Then the group of selected offspring was the one with the highest AD for the marker loci.

Gene diversity in offspring (GO): This strategy is identical to the previous one except that the selection criterion was the expected heterozygosity calculated with the frequencies of the selected offspring at the corresponding markers.

Minimum pedigree coancestry was also used as a reference for strategies based on genetic information from the offspring (allelic diversity in offspring and gene diversity in offspring). Because mating of parents was arranged at random and pedigree relationships for all sibs from a particular couple are the same, setting the parents' contributions to yield the minimum pedigree coancestry implies the minimization of pedigree coancestry in their offspring.

We performed all optimizations with the simulated annealing algorithm (Kirkpatrick et al. 1983; Press et al. 1989). Restrictions were as follows: (1) only integer non-negative solutions were allowed; (2) the sum of all contributions (or all selected offspring) equaled $2N$ to maintain a constant population size; (3) half the contributions came from each of the sexes.

To explain the performance of the different strategies, we investigated their effect on allelic frequencies. The average variance of allele frequencies is

$$S_q^2 = \sum_i^L \frac{1}{N_{Ai}} \sum_j^{N_{Ai}} \left(q_{ij} - \frac{1}{N_{Ai}} \right)^2,$$

where L is the number of loci considered, q_{ij} is the frequency of allele j in locus i and N_{Ai} is the actual number of alleles in locus i . Using the definition given by Crow and Kimura (1970) of the effective number of alleles (N_{Ei}) as $1/\sum q^2$, we get the following expression:

$$S_q^2 = \sum_i^L \left(\frac{1}{N_{Ai}N_{Ei}} - \frac{1}{(N_{Ai})^2} \right).$$

To take into account the fact that different loci may have different numbers of alleles, we calculated the average

coefficient of variation,

$$CV = \sum_i^L \left(\sqrt{\frac{N_{Ai}}{N_{Ei}}} - 1 \right) = \sum_i^L \left(\sqrt{\frac{N_{Ai} - N_{Ei}}{N_{Ei}}} \right).$$

When all alleles have the same frequency, $N_{Ai} = N_{Ei}$, and the coefficient of variation equals zero.

Results

The mean levels of GD and AD maintained in the population after 15 generations under any of the considered management strategies were very different depending on the number of markers used and on the initial diversity level, when a population size of eight individuals was considered (Tables 2 & 3). Results arising from simulations with a population size of 16 individuals showed a similar pattern (data not shown). Because we were interested in the global genome-wide diversity, the calculations were made using all loci in the genome (100 loci per chromosome). Thus, data from markers were also included in the calculations of AD and GD. Nevertheless, values obtained by discarding this information were similar (data not shown).

As a rule, when the number of markers was large (10 or 100), all strategies were less efficient in preserving genetic diversity (both AD and GD) for large genomes (larger number of chromosomes). However, when few markers per chromosome were used (one or two), the efficiency was either not affected or increased with the number of chromosomes.

The performances of the strategies for populations with different degrees of initial genetic variability (5 or 10 generations prior to the application of the strategy) differed for AD and GD. The former (Table 2) consistently decreased with decreasing genetic variability in the starting population. However, efficiency in maintaining GD (Table 3) decreased with lower levels of initial diversity when information was available from two markers or less, but the trend was reversed when many markers were available.

When only genotypic information from parents was used, improvements on the levels of genetic diversity relative to those reached by applying minimum pedigree coancestry were only obtained when information from many markers was used. In some situations neither expected allelic diversity from parents nor gene diversity from parents' methods outperformed the strategy of minimizing the pedigree coancestry, even with 10 markers per chromosome (Tables 2 & 3). There was an improvement in performance when data on offspring was used, even when only two individuals per parent were available for selection. In fact, the magnitude of the increase, either in AD or GD, when information from two offspring per

Table 2. Mean allelic diversity (percent relative to the levels achieved by minimizing the global pedigree coancestry) at generation 15 under different management strategies.^a

| No. of prior unmanaged generations | No. of chromosomes in genome | Markers per chromosome | Management strategies ^b | | | | | | | |
|------------------------------------|------------------------------|------------------------|------------------------------------|-----|-----|-------|-------|-------|-------|----|
| | | | AP | EAP | GP | AO(2) | GO(2) | AO(8) | GO(8) | |
| 5 | 1 | 1 | -20 | -24 | -19 | -6 | 1 | -4 | 1 | |
| | | 2 | -21 | -11 | -10 | 3 | 18 | 11 | 17 | |
| | | 10 | -20 | 18 | 17 | 41 | 54 | 73 | 80 | |
| | | 100 | -20 | 26 | 23 | 58 | 61 | 100 | 96 | |
| | 20 | 0.5 | 1 | -21 | -19 | -20 | -6 | -2 | -3 | 0 |
| | | | 1 | -20 | -11 | -14 | -2 | -1 | 5 | 7 |
| | | | 2 | -19 | -5 | -9 | 2 | 3 | 15 | 13 |
| | | | 10 | -18 | 0 | 0 | 11 | 12 | 33 | 30 |
| | | 100 | 1 | -18 | 1 | 0 | 12 | 12 | 36 | 36 |
| | | | 2 | -19 | -5 | -9 | 2 | 3 | 15 | 13 |
| | | | 10 | -18 | 0 | 0 | 11 | 12 | 33 | 30 |
| | | | 100 | -18 | 1 | 0 | 12 | 12 | 36 | 36 |
| 10 | 1 | 1 | -17 | -22 | -14 | -7 | -1 | -9 | -7 | |
| | | 2 | -15 | -17 | -11 | -2 | 4 | -3 | 2 | |
| | | 10 | -14 | 18 | 15 | 20 | 35 | 35 | 42 | |
| | | 100 | -16 | 29 | 22 | 37 | 42 | 56 | 62 | |
| | 20 | 0.5 | 1 | -18 | -22 | -21 | -7 | -4 | -8 | -4 |
| | | | 1 | -16 | -13 | -16 | -3 | -1 | -3 | 1 |
| | | | 2 | -15 | -7 | -10 | 0 | -2 | 5 | 7 |
| | | | 10 | -14 | -1 | -1 | 7 | 10 | 24 | 22 |
| | | 100 | 1 | -18 | -22 | -21 | -7 | -4 | -8 | -4 |
| | | | 1 | -16 | -13 | -16 | -3 | -1 | -3 | 1 |
| | | | 2 | -15 | -7 | -10 | 0 | -2 | 5 | 7 |
| | | | 10 | -14 | -1 | -1 | 7 | 10 | 24 | 22 |
| 100 | -14 | 0 | 0 | 10 | 11 | 29 | 28 | | | |

^aPopulation size is 8.

^bManagement strategy: AP, allelic diversity in parents; EAP, expected allelic diversity from parents; GP, gene diversity from parents; AO, allelic diversity in offspring; GO, gene diversity in offspring. Numbers in parentheses are the number of offspring per individual. Standard deviations range from 0.3 to 3.0.

parent was used instead of parental data only was similar to the gain obtained by using eight individuals per parent instead of two. Moreover, for cases with few markers per chromosome, increasing the number of offspring per parent did not make any difference or even lead to poorer results (Tables 2 & 3).

The allelic diversity in parents' strategy performed poorly because it only accounted for selected individuals, not for their contributions. Thus, parents carrying rare alleles were selected but not forced to reproduce more, so the probability of losing those alleles by chance was high. Consequently, the use of more markers did not significantly improve the performance of the strategy.

When only the genotypes of the parents were considered, the most favorable strategy was gene diversity from parents. This strategy outperformed expected allelic diversity from parents in most situations with regard to the preservation of GD (Table 3), and it was roughly equivalent with regard to AD (Table 2). The same pattern was observed for strategies relying on offspring information, irrespective of the number of individuals obtained from each parent. Gene diversity in offspring never yielded significantly lower levels of gene diversity than allelic diversity in offspring (Table 3), and only in a few cases for allelic diversity (Table 2).

The coefficient of variation of the allelic frequencies evolved differentially over time under the different management strategies (Fig. 1; the evolution of the coefficient of variation under the allelic diversity in parents'

strategy is not shown because of its poor performance). Both gene diversity from parents and gene diversity in offspring strategies reduced the coefficient of variation of allele frequencies across generations, the latter being more effective because it deals with real gene-diversity values and not with predictions. The consequences of strategies based on allelic diversity differed depending on the source of information (parents or offspring genotype). With the strategy of expected allelic diversity from parents, the coefficient of variation was also reduced because it takes into account the probability of allele loss, so optimum values should be obtained with even frequencies. However, the reduction was smaller than under gene diversity from parents because, as the number of alleles is the criterion for expected allelic diversity from parents, solutions where allele frequencies are not equilibrated may exist in this latter case. When the strategy of allelic diversity in offspring was implemented, the coefficients of variation remained at their initial level or increased. The reason is that this strategy attempts to maintain as many alleles as possible, including those at very low frequencies.

Discussion

From a genetic point of view, managers could be advised to concentrate on the maintenance of the genetic diversity

Table 3. Mean gene diversity (percent relative to the levels achieved by minimizing the global pedigree coancestry) at generation 15 under different management strategies.^a

| No. of prior unmanaged generations | No. of chromosomes in genome | Markers per chromosome | Management strategies ^b | | | | | | | |
|------------------------------------|------------------------------|------------------------|------------------------------------|-----|-----|-------|-------|-------|-------|--|
| | | | AP | EAP | GP | AO(2) | GO(2) | AO(8) | GO(8) | |
| 5 | 1 | 1 | -30 | -36 | -28 | -11 | 0 | -11 | 1 | |
| | | 2 | -30 | -16 | -13 | -2 | 20 | 3 | 18 | |
| | | 10 | -29 | 22 | 30 | 30 | 59 | 52 | 73 | |
| | | 100 | -29 | 30 | 39 | 44 | 67 | 68 | 85 | |
| | 20 | .5 | -30 | -26 | -27 | -9 | -3 | -6 | -1 | |
| | | 1 | -29 | -15 | -17 | -5 | 0 | 1 | 6 | |
| | | 2 | -27 | -6 | -9 | 0 | 6 | 12 | 16 | |
| | | 10 | -25 | 0 | 2 | 9 | 19 | 27 | 40 | |
| | | 100 | -24 | 1 | 2 | 10 | 21 | 30 | 46 | |
| | | | | | | | | | | |
| 10 | 1 | 1 | -31 | -39 | -26 | -13 | 0 | -19 | -9 | |
| | | 2 | -30 | -30 | -16 | -7 | 9 | -13 | 6 | |
| | | 10 | -28 | 33 | 41 | 23 | 70 | 41 | 79 | |
| | | 100 | -32 | 51 | 57 | 44 | 86 | 65 | 113 | |
| | 20 | 0.5 | -30 | -41 | -38 | -14 | -7 | -17 | -7 | |
| | | 1 | -30 | -24 | -28 | -8 | -1 | -8 | 2 | |
| | | 2 | -29 | -12 | -14 | -2 | -1 | 4 | 14 | |
| | | 10 | -26 | -1 | 2 | 7 | 24 | 29 | 47 | |
| | | 100 | -26 | 0 | 3 | 11 | 28 | 35 | 59 | |
| | | | | | | | | | | |

^aPopulation size is eight.

^bManagement strategy: AP, allelic diversity in parents; EAP, expected allelic diversity from parents; GP, gene diversity from parents; AO, allelic diversity in offspring; GO, gene diversity in offspring. Numbers in parentheses are the number of offspring per individual. Standard deviations range from 0.3 to 2.6.

of their populations. High levels of GD are desirable in any population because, at least in populations at equilibrium, GD is directly related to additive genetic variance and therefore to the short-term response to selection (either natural or artificial), and it is inversely related to the amount of inbreeding depression. On the other hand, preserving high levels of AD is also an important issue in conservation programs because, as pointed out by Hill & Rasbash (1986), the long-term response to selection depends on the number of different alleles. Moreover, the adaptation to specific new environments or habitats is conditioned by the existence of alternative genetic information (i.e., particular alleles). Therefore, a high AD level is expected to widen the range of situations a population can cope with. Measuring genetic diversity as AD is also useful when inferring population history because it is more sensitive to short-duration bottlenecks than GD (Allendorf 1986).

There is a consensus on the optimal way to manage GD (Ballou & Lacy 1995; Toro et al. 1999; Caballero & Toro 2000; Fernández & Caballero 2001). However, the ideal strategy to follow when the key parameter is AD is unclear, as is the relation between this and other diversity measures. In the long term, the ideal theoretical strategy would be to divide the population into as many subpopulations as possible (e.g., brother-by-sister lines). In such a scenario, the expectation is that different alleles would be fixed in different lines. Thus, the overall allelic frequencies (and also AD) would tend to remain at

initial values in the population if the number of subpopulations is high enough (Falconer & Mackay 1996), or at least they would reach a permanent stable situation. This management procedure would lead to a rapid increase in inbreeding, however, with its associated deleterious effects, hence lowering the viability of the population. On the contrary, it has been shown that the optimal strategies by which to maximize GD should also keep levels of inbreeding low as genetic drift is lowered (Caballero & Toro 2000).

Our simulations show that strategies aimed at maximizing GD also keep levels of AD high, at least for the range of scenarios simulated. At first thought, this result could be surprising. From the definition of GD, however, it can be seen that the maximum value for this parameter is obtained with intermediate allele frequencies, when all existing alleles at a particular locus have the same frequency. Therefore, maximization of GD should lead frequencies to intermediate values (Fig. 1). In this situation, alleles are, globally, as far as possible from zero frequencies, and thus the loss of alleles by genetic drift is minimized.

A further advantage of the use of GD as the management criterion is that it also retains higher levels of observed heterozygosity and thus lower levels of inbreeding (data not shown). This should be another key objective in a conservation program that seeks to avoid deleterious effects of inbreeding depression and the extinction of populations.

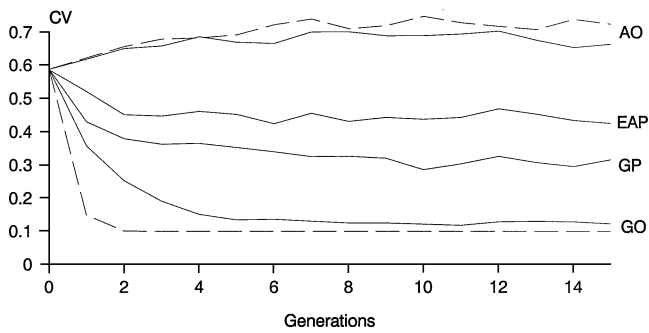


Figure 1. Average coefficient of variation (CV) of allele frequencies under different management procedures (abbreviations defined in Table 2). For AO and GO, solid and broken lines refer to two and eight offspring per parent evaluated, respectively. Results are from simulations with a population census size of eight individuals, with one chromosome, two markers genotyped, and five previous unmanaged generations.

As shown by our simulations, procedures based only on parent molecular information need to use a large number of markers to be advantageous (Tables 2 & 3). In fact, around 10 markers per Morgan are necessary to maintain the same level of genetic diversity as that maintained by using pedigree coancestry. If the level of polymorphism is low—for example, for allozymes, single-nucleotide polymorphisms, or amplified-fragment-length polymorphisms—this number should be much higher (J. F. B. Villanueva & M.A.T., unpublished data). When such a large number of markers is not available, as in many conservation programs, greater levels of diversity (and lower cost for the program) would be obtained by using only pedigree information to manage the population, assuming that pedigrees can be recorded. If molecular and pedigree information are available, they can be jointly used in two ways. On the one hand, markers can aid in the verification of unknown relations (e.g., parentage assignment). On the other, pedigree and markers can be used to calculate the coancestry conditional on markers, and this conditional coancestry (which is inversely related to GD) can be used to monitor the conservation program. Toro et al. (1999) and Wang (2000) found that the minimization of global coancestry conditional on markers is the optimal strategy by which to minimize inbreeding and genetic drift in a small population. Therefore, the maximization of genetic diversity should be optimal via the joint use of both types of information.

When the species under conservation presents no physiological or management restrictions, it seems advisable to base decisions on information from offspring instead of using data from parents. However, because the increase in efficiency is not linear with respect to the number of offspring (note, in Table 3, the less than two-

fold increase in diversity with a fourfold increase in the number of individuals genotyped), it is probably more practical to increase the number of markers per individual than to increase the number of offspring per parent. Moreover, if the number of markers used is low, increasing the number of offspring per parent could lead to worse results, meaning lower diversity levels. This counterintuitive result may be explained by the fact that when molecular information is scarce one cannot effectively determine relationships between offspring that are then selected largely by chance. If family size is large, there is a higher probability of selecting many full-sibs by chance than when few offspring are available per parent.

Many captive populations under conservation programs are kept with a skewed sex ratio because of management restrictions or conveniences. In this situation, minimizing coancestry is still the optimal way of maximizing GD and reducing genetic drift, provided the different numbers of males and females are taken into account by differentially weighting coancestries between males, between females, or between males and females (Caballero & Toro 2000, 2002). Therefore, the good performance of the levels of AD maintained by minimizing coancestry should also hold for unequal sex numbers.

In principle, the genetic diversity we would like to keep in the population is that of loci with effect on quantitative traits responsible for the evolvability of the population (fitness-related traits) or those valuable in the conservation program (economical, historical, or cultural values). In most situations, however, those genes are unknown, and the only information available is that from neutral molecular markers. Nevertheless, if the number of markers is large enough and if they effectively cover the entire genome, we expect selective genes to be more or less linked to any of the markers and therefore the fate of the former to be similar to that of the latter. Bataillon et al. (1996) showed, by computer simulation, that the use of genetic markers leads to the retention of allelic richness in neutral and non-neutral loci when a “core” is created from a larger germplasm collection. Seddon and Baverstock (1999) found a positive correlation between allelic diversity for the major histocompatibility complex (MHC) and for some neutral allozymes, indicating that drift acts even on strongly selected balanced polymorphisms. Therefore, any effort to lower the effects of genetic drift on neutral markers would also affect selective loci.

If loci with effect on traits of interest are known, they can be acted on directly. The criterion to be used in that situation is still GD because similar relative performances for the different strategies can be observed for the whole genome (Tables 2 & 3) and for only the genotyped markers (data not shown).

Some unconditionally deleterious alleles may exist in any population. The survival of the population will be threatened by the harmful effects of those alleles if we

maintain them as the result of a general conservation policy. A specific case is that of hereditary disorders or diseases caused by one allele in one locus of large effect (Laikre 1999). The simplest strategy for eliminating individuals with a high probability of carrying that allele is not advisable because it would lead to a great reduction in the effective population size and therefore to a great loss in genetic diversity for the whole genome (Lacy 2000). Sonesson et al. (2003) have tested the effect of different strategies in which the reduction of the frequency of undesirable alleles is balanced against the increase of inbreeding levels. Studies are required to test the effect of such strategies on the maintenance of genetic diversity at all loci in the genome but the one with the deleterious allele. The same argument can be used in the case of a particular allele known to have a beneficial or useful effect on the performance of individuals. The allele frequency should be increased, but the maintenance of genome-wide diversity should be guaranteed.

Acknowledgments

We thank B. Villanueva and two anonymous referees for helpful comments on the manuscript. This work was supported by grants EVK3-CT-2001-00048 from the European Union, BOS2000-0896 from the Ministerio de Ciencia y Tecnología and Fondos Feder, PGIDT02PXIC301014PM from the Xunta de Galicia, 64102C124 from the Universidade de Vigo, and RZ01-028-C2-1 from Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). J.F. was supported by a contract from Programa Ramón y Cajal.

Literature Cited

- Allendorf, F. W. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology* **5**:181-190.
- Ballou, J. D., and R. C. Lacy. 1995. Identifying genetically important individuals for management of genetic variation in pedigree populations. Pages 76-111 in J. D. Ballou, M. Gilpin, and T. J. Foose, editors. *Population management for survival and recovery*. Columbia University Press, New York.
- Barker, J. S. F. 2001. Conservation and management of genetic diversity: a domestic animal perspective. *Canadian Journal of Forestry Research* **31**:588-595.
- Bataillon, T. M., J. L. David, and D. J. Schoen. 1996. Neutral genetic markers and conservation genetics: simulated germplasm collections. *Genetics* **144**:409-417.
- Caballero, A., and M. A. Toro. 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genetical Research* **75**:331-343.
- Caballero, A., and M. A. Toro. 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conservation Genetics* **3**:289-299.
- Chevalet, C. 1992. Utilisation de marqueurs pour la sauvegarde de la variabilité génétique des populations. Pages 295-297 in *Éléments de génétique quantitative et application aux populations animales*. Institut National de la Recherche Agronomique et Production Animal, Paris.
- Comps, B., D. Gömöry, J. Letouzey, B. Thiébaud, and R. J. Petit. 2001. Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* **157**:389-397.
- Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**:2001-2014.
- Crow, J. F., and M. Kimura. 1970. *An introduction to population genetics theory*. Harper & Row, New York.
- Engels, W. R. 1980. Loss of selectively neutral alleles in small populations and regular mating systems. *Theoretical Population Biology* **17**:345-364.
- England, P. R., G. H. R. Osler, L. M. Woodworth, M. E. Montgomery, D. A. Briscoe, and R. Frankham. 2003. Effects of intense versus diffuse bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conservation Genetics* **4**:595-604.
- Ewens, W. J. 1972. The sampling theory of selectively neutral alleles. *Theoretical Population Biology* **3**:87-112.
- Falconer, D. S., and T. F. C. Mackay. 1996. *An introduction to quantitative genetics*. 4th edition. Longman, Harlow, United Kingdom.
- Fernández, J., and A. Caballero. 2001. A comparison of management strategies for conservation with regard to population fitness. *Conservation Genetics* **2**:121-131.
- Fernández, J., and M. A. Toro. 1999. The use of mathematical programming to control inbreeding in selection schemes. *Journal of Animal Breeding and Genetics* **116**:447-466.
- Frankham, R., J. D. Ballou, and D. A. Briscoe. 2002. *Introduction to conservation genetics*. Cambridge University Press, Cambridge, United Kingdom.
- Hill, W. G., and J. Rasbash. 1986. Models of long term artificial selection in finite populations. *Genetical Research* **48**:41-50.
- James, J. W. 1971. The founder effect and response to artificial selection. *Genetical Research* **12**:249-266.
- Kirpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* **220**:671-680.
- Lacy, R. C. 2000. Should we select genetic alleles in our conservation breeding programs? *Zoo Biology* **19**:279-282.
- Laikre, L. 1999. Hereditary defects and conservation genetic management of captive populations. *Zoo Biology* **18**:91-99.
- Luikart, G., W. B. Sherwin, B. M. Steele, and F. W. Allendorf. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* **7**:963-974.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* **70**:3321-3323.
- Oldenbroek, J. K. 1999. Genebanks and the conservation of farm animal genetic resources. Institute for Animal Sciences and Health, Lelystad, The Netherlands.
- Petit, R. J., A. El Mousadik, and O. Pons. 1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**:844-855.
- Press, W., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1989. *Numerical recipes*. Cambridge University Press, Cambridge, United Kingdom.
- Richards, C., and P. L. Leberg. 1996. Temporal changes in allele frequencies and a population's history of severe bottlenecks. *Conservation Biology* **10**:832-839.
- Schoen, D. J., and H. D. Brown. 1993. Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. *Proceedings of the National Academy of Sciences of the United States of America* **90**:10623-10627.
- Seddon, J. M., and P. R. Baverstock. 1999. Variation on island: major histocompatibility complex (*Mhc*) polymorphism in populations of the Australian bush rat. *Molecular Ecology* **8**:2071-2079.
- Sonesson, A. K., L. L. G. Janss, and T. H. E. Meuwissen. 2003. Selection against genetic defects in conservation schemes while controlling inbreeding. *Genetics Selection Evolution* **35**:353-368.
- Toro, M. A., L. Silió, M. C. Rodríguez, and J. Rodríguez. 1998. The

- use of molecular markers in conservation programmes of live animals. *Genetics Selection Evolution* **30**:585–600.
- Toro, M. A., L. Silió, M. C. Rodríguez, J. Rodríguez, and J. Fernández. 1999. Optimal use of genetic markers in conservation programmes. *Genetics Selection Evolution* **31**:255–261.
- Vales-Alonso, J., J. Fernández, F. J. González-Castaño, and A. Caballero. 2003. A parallel optimization approach for controlling allele diversity in conservation schemes. *Mathematical Biosciences* **183**: 161–173.
- Wang, J. 2000. Optimal marker-assisted selection to increase the effective size of small populations. *Genetics* **157**:867–874.
- Young, A. G., A. H. D. Brown, B. G. Murray, P. H. Thrall, and C. H. Miller. 2000. Genetic erosion, restricted mating and reduced viability in fragmented populations of endangered grassland herb *Rutidosis leptorrhynchoides*. Pages 335–359 in A. G. Young and G. M. Clarke, editors. *Genetics, demography and viability of fragmented populations*. Cambridge University Press, Cambridge, United Kingdom.

