

## The use of molecular markers in conservation programmes of live animals

Miguel Toro \*, Luis Silió, Jaime Rodríguez, Carmen Rodríguez

Departamento de Mejora Genética y Biotecnología, INIA,  
Ctra. de La Coruña km 7, 28040 Madrid, Spain

(Received 6 January 1998; accepted 11 September 1998)

**Abstract** – Monte Carlo simulation has been carried out to study the benefits of using molecular markers in a conservation programme to minimize the homozygosity by descent in the overall genome. Selection of the breeding individuals was either at random or based on two alternative criteria: overall heterozygosity of the markers or frequency-dependent selection. Even molecular information was available for all the 1 900 simulated loci, a conventional tactic such as restriction in the variance of the family size is the most important strategy for maintaining genetic variability. In this context: a) frequency-dependent selection seems to be a more efficient criterion than selection for heterozygosity; and b) the value of marker information increases as the selection intensity increases. Results from more realistic cases (1, 2, 3, 4, 6 or 10 markers per chromosome and 2, 4, 6 or 10 alleles per marker) confirm the above conclusions. This is an expensive strategy with respect to the number of candidates and the number of markers required in order to obtain substantial benefits, the usefulness of a marker being related to the number of alleles. The minimum coancestry mating system was also compared with random mating and it is concluded that it is advantageous at least for many generations. © Inra/Elsevier, Paris

**molecular markers / conservation genetics / frequency-dependent selection / minimum coancestry mating**

**Résumé** – Utilisation de marqueurs moléculaires dans les programmes de conservation des animaux. Des simulations Monte Carlo ont été effectuées pour étudier l'intérêt de l'utilisation des marqueurs moléculaires dans un programme de conservation avec  $N_s$  (= 4, 8 ou 16) mâles et  $N_d = 3 N_s$  femelles, choisis parmi 3  $N_d$  candidats de chaque sexe. Le génome a été simulé avec 1 900 locus distribués sur 19 chromosomes d'une longueur de 100 cM chacun. L'objectif était de minimiser le taux d'homozygotie chez la descendance pour l'ensemble du génome, le choix des

---

\* Correspondence and reprints  
E-mail: toro@inia.es

reproducteurs s'effectuant au hasard ou sur la base d'un critère calculé à l'aide de l'information aux marqueurs : sélection pour le taux global d'hétérozygotie des marqueurs ou sélection en faveur des allèles rares. Dans la situation optimale, où l'information moléculaire est disponible pour l'ensemble des locus, les résultats montrent que l'emploi de stratégies conventionnelles telles que la restriction de la variance des tailles de famille demeure le facteur le plus important. Dans ce contexte : a) la sélection en faveur des allèles rares semble être un critère plus efficace que la sélection pour l'hétérozygotie ; b) la valeur de l'information des marqueurs augmente lorsque l'intensité de sélection augmente. Ces conclusions sont confirmées dans des situations plus réalistes en ce qui concerne le nombre de marqueurs par chromosome (1, 2, 3, 4, 6 ou 10) et le nombre d'allèles par marqueur (2, 4, 6 ou 10). On remarque que, pour obtenir des bénéfices substantiels, on a besoin d'une stratégie coûteuse en termes de nombres de candidats et de marqueurs, l'utilité d'un marqueur dépendant du nombre d'allèles. Finalement, l'effet d'un système d'accouplement minimisant la parenté a été trouvé avantageux à moyen terme. © Inra/Elsevier, Paris

**marqueurs moléculaires / génétique de la conservation / sélection dépendant de la fréquence / accouplement pour le minimum de parenté**

## 1. INTRODUCTION

The interest in conserving different breeds and strains of farm livestock has arisen owing to the awareness of dangers created by the continuous decrease in the number of commercially exploited breeds and/or by the reduction of genetic variability imposed in modern breeding programmes [14].

The limited size of conserved populations of domestic strains causes inbreeding and loss of genetic variance, which lowers the performance of animals for at least some traits and increases the risk of extinction [12]. There are several ways to measure genetic variation and its loss but there is a consensus that in populations with genealogical records, calculation of inbreeding and coancestry coefficients are the most common tools for monitoring conservation schemes and for designing strategies to minimize inbreeding [3, 4].

The application of new technologies in molecular biology provides information on genotypes of several polymorphic loci and therefore allows one to quantify the genetic variability by a list of alleles and their joint distribution of frequencies at many loci. A summary of this information is given by the observed genetic heterozygosity (homozygosity) defined as the proportion of loci heterozygous (homozygous) either at individual or at population level. Other measures are the effective number of alleles or the expected genetic heterozygosity, both related to the squares of allele frequencies [1, 2].

The use of molecular markers allows one to increase the efficiency of conservation methods. Chevalet and Rochambeau [8] proposed a selection using an index equal to the inverse of the product of the frequencies of the alleles and more recently Chevalet [7] proposed a selection using an index equal to the heterozygosity measured at several marker loci.

In this paper, we present Monte Carlo simulation results on the benefits of using molecular information in a small conservation nucleus, considering different alternatives: individual or within-family selection, heterozygosity or frequency-dependent selection and random or minimum coancestry mating.

## 2. SIMULATION

The breeding population consisted of  $N_s$  ( $= 4, 8$  or  $16$ ) sires and  $N_d = 3 N_s$  dams. Each dam produced three progeny of each sex. These three  $N_d$  offspring of each sex were the maximum possible number of candidates for selection to form the breeding individuals of the next generation.

The genome was simulated as 19 chromosomes, each with 100 loci placed at 1 cM intervals. All the loci of the founder population,  $2(N_s + N_d)$ , were considered different by descent. For selection purposes, a variable number of marker loci with a variable number of alleles were also situated in the chromosomes in an equally spaced manner. These marker loci were generated in linkage equilibrium in the base population.

Selection was either at random or based on two alternative criteria based on genetic markers.

a) Selection for overall heterozygosity of the markers (HET), where the value of the genotype at each locus was computed as 1 if it was heterozygous, or 0 if it was homozygous, the value of an individual being the sum over loci.

b) Frequency-dependent selection (FD), where the value assigned to the genotype increased as the population frequency of the alleles that make this genotype decreased. There are many possible schemes of frequency-dependent selection but perhaps the simplest one is that proposed by Crow [9] in his basic textbook on population genetics. In this particular scheme, the value of the genotype  $A_i A_j$  at each locus is  $(1 - p_i/2)(1 - p_j/2)$ ,  $p_i$  and  $p_j$  being the frequencies of the  $A_i$  and  $A_j$  alleles, respectively, and therefore the homozygote for the rare allele is favoured over the heterozygote, which is favoured over the homozygote for the more common allele (except when the allelic frequencies are equal, where heterozygotes are favoured). For biallelic dominant markers, the equivalent method is to assign to the genotypes  $A_2 A_2$  and  $A_1 A_1$  the values  $(1 - p_2/2)^2$  and  $(1 - p_1/2)^2$ , respectively. The value of an individual is the sum over all the marker loci. In a small number of additional simulations, the effective number of alleles of the selected individuals as a group was used as selection criterion. By analogy with the concept defined by Crow and Kimura [10], this parameter was calculated as  $n_a = L / \sum_i \sum_j p_{ij}^2$  where  $p_{ij}$  is the average frequency, in the selected population, of the allele  $i$  at locus  $j$ , and  $L$  is the number of marker loci.

Two types of selection were also considered: a) within-family selection (WFS), where each dam family contributed one dam and each sire family contributed one sire to the next generation; b) individual selection (IND) where no restriction was imposed on the number of breeding animals that each family contributed to the next generation.

Two types of matings were implemented: a) random mating, and b) minimum coancestry mating where the average pairwise coancestry coefficient in the selected group was minimized. Minimum coancestry mating was implemented using linear programming techniques [20].

The selection scheme was carried out for 15 generations. In each generation, several parameters were calculated: a) the proportion of the genome identical by descent calculated over the 1900 loci that describe the genome; b) the proportion of homozygosity for the marker loci used in the selection criterion;

c) the average inbreeding and coancestry coefficients of selected individuals calculated from the pedigrees; and d) the effective number of alleles calculated as previously indicated.

### 3. RESULTS

#### 3.1. Complete molecular information

For different population structures, criteria and types of selection (including the situation of no selection due to the lack of molecular information) and random mating, the average homozygosity by descent of the population and the inbreeding coefficient calculated through the pedigree are shown in *table I*. The average coancestry coefficient of all possible mates between the sires and dams of the previous generation was also calculated but is not included in the table because it gives values almost identical to those of inbreeding, as expected due to random mating.

**Table I.** Homozygosity (Ho, %) and inbreeding coefficient (F, %) at generation 15, for different population size ( $N_s$  sires,  $3 N_s = N_d$  dams) and criteria of choice of breeding animals, with random mating and complete molecular information from the genotyped candidates<sup>a</sup>.

| $N_s$   | Selection criterion <sup>b</sup> | Type of selection <sup>c</sup> | Number of genotyped candidates of each sex |                 |                  |                 |                  |                 |                  |                 |
|---------|----------------------------------|--------------------------------|--|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|
|         |                                  |                                | 0  |                 | $N_d$            |                 | $2 N_d$          |                 | $3 N_d$          |                 |
|         |                                  |                                | Ho <sup>15</sup>                           | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> |
| Four    | HET                              | IND                            | 43.85                                      | 43.85           | 38.19            | 41.39           | 38.39            | 44.67           | 39.09            | 46.92           |
| Four    | HET                              | WFS                            | 35.10                                      | 35.09           | 32.98            | 34.75           | 31.07            | 34.75           | 30.32            | 34.78           |
| Four    | FD                               | IND                            | 43.85                                      | 43.85           | 35.03            | 39.87           | 37.12            | 45.57           | 41.74            | 51.79           |
| Four    | FD                               | WFS                            | 35.10                                      | 35.09           | 28.44            | 31.51           | 24.69            | 32.97           | 23.35            | 32.97           |
| Eight   | HET                              | IND                            | 25.04                                      | 25.09           | 22.07            | 23.96           | 22.93            | 26.71           | 23.75            | 28.47           |
| Eight   | HET                              | WFS                            | 18.90                                      | 18.90           | 17.91            | 18.77           | 17.02            | 18.73           | 16.63            | 17.72           |
| Eight   | FD                               | IND                            | 25.04                                      | 25.09           | 19.19            | 22.16           | 21.81            | 27.08           | 25.85            | 32.61           |
| Eight   | FD                               | WFS                            | 18.90                                      | 18.90           | 14.98            | 17.89           | 12.85            | 17.65           | 12.10            | 17.60           |
| Sixteen | HET                              | IND                            | 13.45                                      | 13.45           | 12.09            | 13.03           | 12.86            | 14.73           | 13.69            | 16.11           |
| Sixteen | HET                              | WFS                            | 9.82                                       | 9.81            | 9.40             | 9.77            | 9.00             | 9.73            | 8.90             | 9.75            |
| Sixteen | FD                               | IND                            | 13.45                                      | 13.45           | 10.04            | 11.66           | 12.01            | 14.92           | 14.98            | 18.90           |
| Sixteen | FD                               | WFS                            | 9.82                                       | 9.81            | 7.69             | 9.26            | 6.54             | 9.13            | 6.17             | 9.11            |

<sup>a</sup> 4 000, 2 000 and 1 000 runs for  $N_s = 4, 8$  and  $16$ , respectively; standard errors of both parameters range from 0.03 to 0.05. <sup>b</sup> HET = selection for heterozygosity of markers or FD = frequency-dependent selection. <sup>c</sup> IND = individual selection or WFS = within-family selection.

With random choice of breeding animals (no molecular information available), the true values of genomic homozygosity at generation 15 were almost identical to the values of inbreeding calculated from pedigree records. On the other hand, the inverse of the effective number of alleles coincided

with the mean coancestry (including self-coancestries and reciprocals) since  $1/n_a = \sum_i \sum_j p_{ij}^2 / L$  can be interpreted as the probability that two alleles taken at random from the pool of gametes produced by the current population are identical by descent. From *table I*, it is clear that, besides the obvious effect of the number of breeding individuals, the most important factor lowering the rate of homozygosity was restriction on the variance of family size (i.e. ensuring that each sire family leaves a sire and each dam family leaves a dam to the next generation), which resulted in decreasing this rate by about 25 %.

When selection using complete molecular information was practised, the inbreeding coefficient did not reflect the true homozygosity and the discrepancy increased as selection intensity increased. The criterion of restricted family size was of paramount importance. When the maximum molecular information was used but no restriction was placed on family size, the homozygosity was always greater than when molecular information was ignored but within-family selection was practised. With individual selection, from the maximum number of candidates available ( $3 N_d$ ), a variable number ( $N_d$ ,  $2 N_d$  or  $3 N_d$ ) was chosen at random to be genotyped and then the best individuals were selected. The efficiency of the use of markers decreased as selection intensity increased. That implies that a selection intensity lower than those tested could have been optimal for this number of generations. Although there is no guarantee that these results will be maintained in the long term, they are rather paradoxical and can be attributed to the fact that as selection intensity increases there is a tendency to coselect full- or half-sibs. This is essentially the same effect that was first considered by Robertson [15] in the context of truncation selection and more recently analysed by Woolliams et al. [22] and Santiago and Caballero [17]. Within-family selection involves a restriction on the family size and, with this type of selection and for both criteria, the efficiency increased as selection intensity increased.

In the framework of individual selection, frequency-dependent selection (FD) is more efficient for controlling the homozygosity than selection for overall heterozygosity of the markers (HET), except for the highest selection intensity which is also due to an increased importance of Robertson's effect. But with restricted family size, frequency-dependent selection is more efficient in controlling homozygosity than selection for overall heterozygosity in all the analysed cases. An indication of the genetic similarity among the selected individuals is given by the effective number of alleles ( $n_a$ ), inversely related to their coancestry. In the nucleus of eight sires and 24 dams, the values of  $n_a$  in generation 15 are 3.82 (HET) and 3.52 (FD) for the more intense individual selection, but 5.37 (HET) and 7.23 (FD) for the more intense within-family selection.

The effect of minimum coancestry mating was also considered. With this mating system, the average value of the coancestry coefficient between pairs of selected sires and dams was greater (from 5 to 29 %) than the inbreeding coefficient of the progeny. It induced in all cases a delay in the appearance of inbreeding. *Table II* is equivalent to *table I* but with minimum coancestry mating (mCM) instead of random mating (RM). At generation 15, the values of the homozygosity attained were considerably lower with the use of mCM. The advantage of mCM over RM ranged from 6 to 33 %.

**Table II.** Homozygosity (Ho, %) and inbreeding coefficient (F, %) at generation 15, for different population size ( $N_s$  sires,  $3 N_s = N_d$  dams) and criteria of choice of breeding animals, with minimum coancestry mating and complete molecular information from the genotyped candidates<sup>a</sup>.

| $N_s$   | Selection criterion <sup>b</sup> | Type of selection <sup>c</sup> | Number of genotyped candidates of each sex |                 |                  |                 |                  |                 |                  |                 |
|---------|----------------------------------|--------------------------------|--|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|
|         |                                  |                                | 0  |                 | $N_d$            |                 | $2 N_d$          |                 | $3 N_d$          |                 |
|         |                                  |                                | Ho <sup>15</sup>                           | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> | Ho <sup>15</sup> | F <sup>15</sup> |
| Four    | HET                              | IND                            | 40.31                                      | 40.32           | 35.76            | 40.03           | 34.90            | 42.55           | 35.39            | 44.50           |
| Four    | HET                              | WFS                            | 32.26                                      | 32.23           | 29.55            | 31.97           | 27.73            | 31.97           | 27.03            | 31.97           |
| Four    | FD                               | IND                            | 40.31                                      | 40.32           | 32.35            | 38.54           | 31.60            | 42.42           | 35.00            | 47.88           |
| Four    | FD                               | WFS                            | 32.26                                      | 32.23           | 25.34            | 30.97           | 21.33            | 30.69           | 19.80            | 30.56           |
| Eight   | HET                              | IND                            | 21.90                                      | 21.90           | 19.60            | 22.00           | 19.68            | 24.11           | 20.21            | 25.56           |
| Eight   | HET                              | WFS                            | 16.11                                      | 16.11           | 14.98            | 16.04           | 14.21            | 16.04           | 13.90            | 16.03           |
| Eight   | FD                               | IND                            | 21.90                                      | 21.90           | 16.66            | 20.37           | 17.92            | 25.05           | 20.73            | 30.09           |
| Eight   | FD                               | WFS                            | 16.11                                      | 16.11           | 12.13            | 15.35           | 9.96             | 15.19           | 9.13             | 15.11           |
| Sixteen | HET                              | IND                            | 11.10                                      | 11.08           | 10.13            | 11.25           | 10.20            | 12.58           | 10.79            | 13.35           |
| Sixteen | HET                              | WFS                            | 7.73                                       | 7.74            | 7.31             | 7.71            | 7.02             | 7.72            | 6.91             | 7.72            |
| Sixteen | FD                               | IND                            | 11.07                                      | 11.08           | 8.09             | 9.99            | 9.59             | 13.58           | 11.95            | 17.53           |
| Sixteen | FD                               | WFS                            | 7.73                                       | 7.74            | 5.65             | 7.31            | 4.55             | 7.21            | 4.14             | 7.18            |

<sup>a</sup> 4 000, 2 000 and 1 000 runs for  $N_s = 4, 8$  and  $16$ , respectively; standard errors of both parameters range from 0.02 to 0.03. <sup>b</sup> HET = selection for heterozygosity of markers or FD = frequency-dependent selection. <sup>c</sup> IND = individual selection or WFS = within-family selection.

The diverse situations analysed were also compared according to their rate of homozygosity per generation. This parameter was calculated from generation 6 to 15 as  $\Delta Ho = (Ho^t - Ho^{t-1}) / (1 - Ho^{t-1})$ , where  $Ho^t$  was the average homozygosity by descent of individuals in generation  $t$  (averaged over replicates). In the absence of molecular information, the rate of homozygosity per generation was higher for mCM than for RM, when the variance of family size was restricted. The opposite occurred with individual random choice of breeding animals. This indicates that with restriction on family size RM would be superior in the long term. Some simulation results indicated that the RM superiority will be attained very late, mCM being advantageous for more than 50 generations. In the nucleus of eight sires and 24 dams, the values of homozygosity in generation 50 were  $Ho^{50} = 61.64$  (RM) and 59.30 (mCM), for individual random choice, and  $Ho^{50} = 49.15$  (RM) and 48.20 (mCM) for within-family choice of breeding animals.

The rate of homozygosity summarizes the evolution of genetic variability during the period involved, but when molecular information is used for selection, it does not have an asymptotic meaning and, therefore, it will not necessarily give a good prediction of the increase of homozygosity in later generations. In this case, the disadvantage of the combination of mCM and restricted family size for controlling the homozygosity rate is attenuated. Additional simulation results for a longer term horizon indicated that, in the situations considered, mCM was also superior to RM for more than 50 generations. In the nucleus of

eight sires and 24 dams with the more intense frequency-dependent selection, the values of  $H_o^{50}$  were 51.35 (RM) and 44.38 (mCM) for individual selection, and 26.59 (RM) and 24.32 (mCM) for within-family selection.

### 3.2. Limited number of markers and alleles per marker

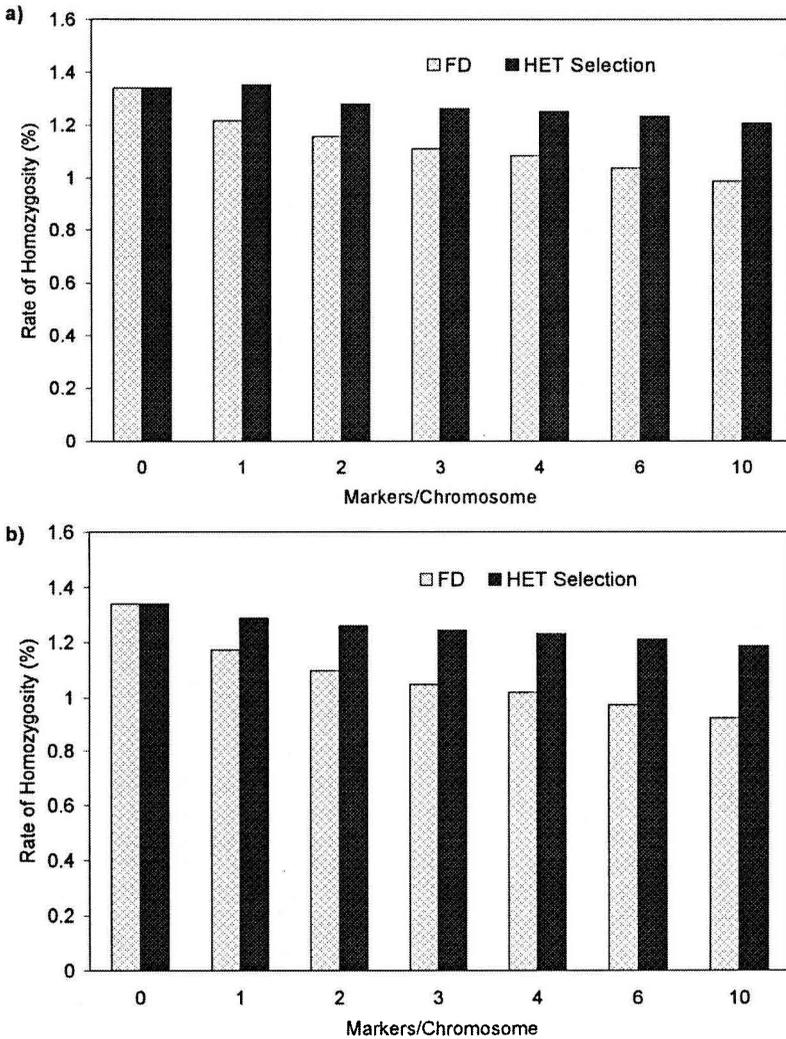
The relative value of the number of markers and the number of alleles per marker has been analysed only for the breeding structure of eight sires, 24 dams and two offspring of each sex per family using RM and WFS in a variety of situations. The homozygosity rate per generation was calculated for both the marker loci and the whole genome.

Two extreme situations were initially considered: a) maximum number of alleles (64, in this particular case) at a limited number of markers per chromosome; and b) maximum number of markers (100 per chromosome) with a limited number of alleles per marker. With totally informative markers, the benefits of using an increasing number of them followed the law of diminishing returns. The use of one marker per chromosome reduced by 5.85 (HET) or 21.00 % (FD) the rate of homozygosity attained without molecular information, while the corresponding values when two markers are genotyped were 8.47 (HET) and 27.16 % (FD). Six markers per chromosome could be enough to achieve similar homozygosity rates to those obtained with 100 markers. On the other hand, if the maximum number of markers is available, then 6–8 alleles per marker allow for the maximum efficiency to be attained.

In a more realistic situation, the joint effect of variable numbers of candidates, markers per chromosome and alleles per marker are shown in *figures 1* and *2*. The results of *figure 1* confirm that frequency-dependent selection was a better method than selection for heterozygosity and that the advantage increased as molecular information increased. The relative value of increasing the number of candidates was also greater with more markers per chromosome although the effect followed the law of diminishing returns as shown in *figure 2*. Finally, the relative advantage of higher number of alleles also increased as both the number of candidates and the number of markers increased (*figure 2*). In summary, these results emphasize that an expensive strategy with respect to the number of candidates and the number of markers is required to obtain appreciable benefits.

More detailed results for both the rate of homozygosity in the whole genome and at the marker loci in a breeding population of eight sires and 24 dams chosen from 48 candidates of each sex, using within-family selection with two selection criteria (HET and FD) and two types of matings (mCM and RM) are given in *tables III* and *IV*. Contrary to the genomic homozygosity rate, homozygosity rate of markers increased as the number of alleles and/or markers increased owing to decreasing level of homozygosity in the initial base population.

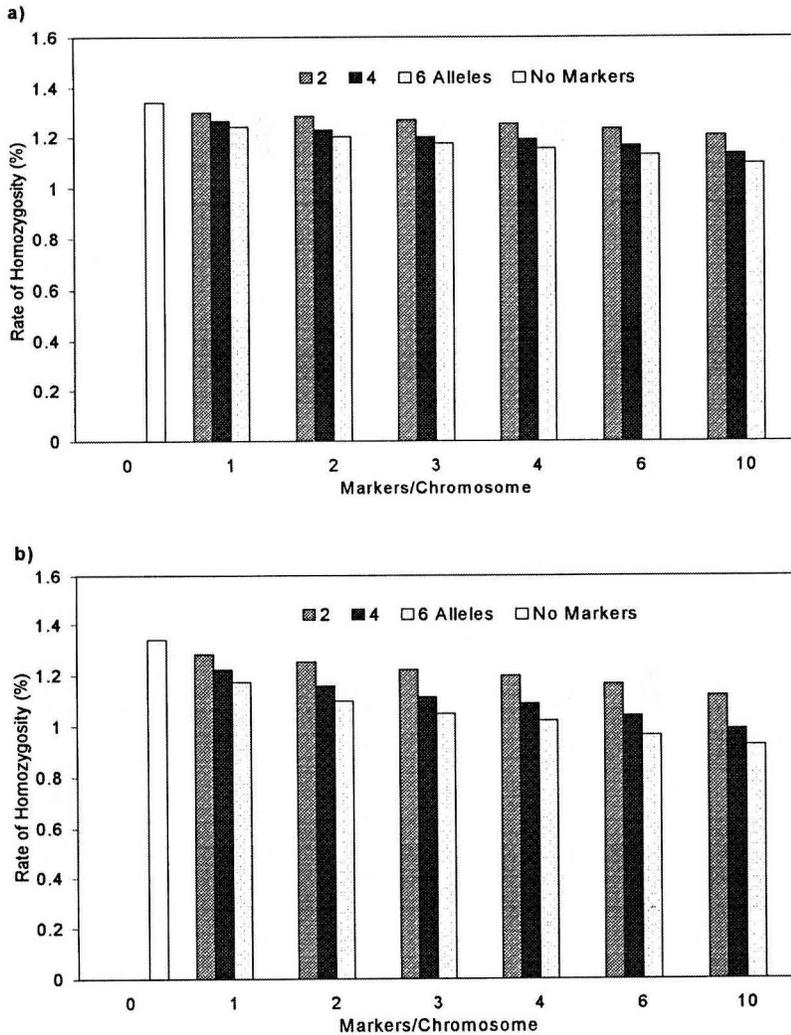
It was confirmed that the value of a marker is related to the number of alleles, especially for FD selection. For example, two markers with six alleles were equally as valuable as (HET) or more valuable than (FD) three markers with two alleles (HET). The greater efficiency of frequency-dependent selection over selection for heterozygosity was more marked for maintaining marker heterozygosity than for maintaining genome heterozygosity and, for example,



**Figure 1.** Effect of the number of markers per chromosome on the rate homozygosity per generation, using frequency-dependent selection (FD) or selection for heterozygosity of markers (HET), with: a) four alleles per marker, b) six alleles per marker. (Population structure:  $N_s = 8$  sires and  $N_d = 24$  dams chosen from  $3 N_d = 72$  candidates of each sex. Random mating.)

in the case of one marker with two alleles, all the initial marker heterozygosity was maintained after 15 generations. This advantageous characteristic could be relevant if the objective were to maintain the heterozygosity of a specific chromosomal region.

The rate of genomic homozygosity was higher for mCM matings owing to the balanced family structure but, as indicated before, the advantage of



**Figure 2.** Effect of the number of markers per chromosome on the rate homozygosity per generation, using frequency-dependent selection, with different numbers of alleles per marker in a population of  $N_s = 8$  sires and  $N_d = 24$  dams chosen from: a) 2  $N_d = 48$ , b) 3  $N_d = 72$  candidates of each sex. (Random mating.)

RM appeared very late (after more than 50 generations in all the situations considered). On the other hand, the rate of marker homozygosity was lower for mCM in all cases of selection for heterozygosity considered or was equal in the cases of low number of markers (one, two or three per chromosome) and frequency-dependent selection. The effective number of alleles retained (results not shown), in contrast to homozygosity, was higher for strategies maintaining more heterozygosity. However, as expected, the loss of alleles was greater when the initial number was higher. For example, with one marker per

**Table III.** Rate of homozygosity per generation of the genome ( $\Delta H_o$ , %), and of the markers ( $\Delta H_{om}$ , %), in a population of family structure  $N_s = 8$  sires and  $N_d = 24$  dams, chosen from 2  $N_d$  candidates of each sex, using within-family selection with two alternative criteria, different numbers of markers per chromosome and alleles per marker and random mating (5 000 runs).

| Selection criterion <sup>a</sup> | Markers per chromosome | Alleles per Marker |                 |              |                 |              |                 |              |                 |
|----------------------------------|------------------------|--------------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|
|                                  |                        | 2                  |                 | 4            |                 | 6            |                 | 10           |                 |
|                                  |                        | $\Delta H_o$       | $\Delta H_{om}$ | $\Delta H_o$ | $\Delta H_{om}$ | $\Delta H_o$ | $\Delta H_{om}$ | $\Delta H_o$ | $\Delta H_{om}$ |
| HET                              | 1                      | 1.32               | 0.08            | 1.31         | 0.38            | 1.30         | 0.53            | 1.28         | 0.69            |
| FD                               | 1                      | 1.29               | -0.01           | 1.23         | 0.00            | 1.19         | 0.03            | 1.14         | 0.16            |
| HET                              | 2                      | 1.31               | 0.19            | 1.28         | 0.53            | 1.27         | 0.68            | 1.26         | 0.80            |
| FD                               | 2                      | 1.25               | 0.00            | 1.17         | 0.02            | 1.12         | 0.13            | 1.07         | 0.28            |
| HET                              | 3                      | 1.30               | 0.29            | 1.27         | 0.63            | 1.26         | 0.76            | 1.24         | 0.87            |
| FD                               | 3                      | 1.23               | 0.01            | 1.14         | 0.08            | 1.08         | 0.21            | 1.03         | 0.37            |
| HET                              | 4                      | 1.29               | 0.37            | 1.26         | 0.70            | 1.25         | 0.81            | 1.23         | 0.91            |
| FD                               | 4                      | 1.21               | -0.01           | 1.11         | 0.13            | 1.05         | 0.27            | 1.00         | 0.42            |
| HET                              | 6                      | 1.28               | 0.49            | 1.25         | 0.78            | 1.23         | 0.88            | 1.21         | 0.97            |
| FD                               | 6                      | 1.18               | 0.01            | 1.07         | 0.22            | 1.02         | 0.36            | 0.96         | 0.50            |
| HET                              | 10                     | 1.27               | 0.62            | 1.23         | 0.87            | 1.21         | 0.95            | 1.19         | 1.02            |
| FD                               | 10                     | 1.14               | 0.07            | 1.03         | 0.34            | 0.97         | 0.47            | 0.92         | 0.59            |

<sup>a</sup> HET = selection for heterozygosity of markers; FD = frequency-dependent selection.

chromosome, RM and HET, if the number of initial alleles was ten, only half of them ( $n_a = 4.62$ ) were retained at generation 15, whereas if the number of initial alleles was two, both of them were retained ( $n_a = 1.91$ ).

A way of diminishing genotyping costs is to use dominant markers such as RAPD or AFLP. In *table V*, dominant and codominant markers are compared considering bi-allelic loci with either equal or unequal frequencies of the two alleles. For the codominant markers, the results with equal and unequal frequencies were similar although the situation of equal frequencies was advantageous especially as the number of markers increased. The use of frequency-dependent selection with dominant markers caused only a small reduction in efficiency compared with codominant bi-allelic markers, although the reduction was greater if the objective was to maintain heterozygosity at markers. The effectiveness of dominant markers was greater if the two phenotypes of each locus were at intermediate frequencies, which implied that the dominant alleles were at low frequencies. Although this comparison with bi-allelic codominant markers is satisfactory, the usual microsatellites are multi-allelic. According to the results of *tables III* and *IV*, obtaining similar homozygosity rates with microsatellites and dominant markers would require, for the second one, a greater number of individuals and/or markers to be genotyped. The first tactic would be adequate for RAPD markers and the second one for AFLP, which produces many markers per analysed sample.

**Table IV.** Rate of homozygosity per generation of the genome ( $\Delta H_o$ , %), and of the markers ( $\Delta H_{o_m}$ , %), in a population of family structure  $N_s = 8$  sires and  $N_d = 24$  dams, chosen from  $2 N_d$  candidates of each sex, using within-family selection with two alternative criteria, different numbers of markers per chromosome and alleles per marker and minimum coancestry mating (5 000 runs).

| Selection criterion <sup>a</sup> | Markers per chromosome | Alleles per Marker |                  |              |                  |              |                  |              |                  |
|----------------------------------|------------------------|--------------------|------------------|--------------|------------------|--------------|------------------|--------------|------------------|
|                                  |                        | 2                  |                  | 4            |                  | 6            |                  | 10           |                  |
|                                  |                        | $\Delta H_o$       | $\Delta H_{o_m}$ | $\Delta H_o$ | $\Delta H_{o_m}$ | $\Delta H_o$ | $\Delta H_{o_m}$ | $\Delta H_o$ | $\Delta H_{o_m}$ |
| HET                              | 1                      | 1.35               | 0.06             | 1.33         | 0.35             | 1.32         | 0.52             | 1.31         | 0.67             |
| FD                               | 1                      | 1.30               | -0.02            | 1.25         | -0.00            | 1.21         | 0.03             | 1.17         | 0.16             |
| HET                              | 2                      | 1.34               | 0.17             | 1.31         | 0.50             | 1.30         | 0.65             | 1.28         | 0.79             |
| FD                               | 2                      | 1.28               | -0.00            | 1.20         | 0.03             | 1.15         | 0.13             | 1.09         | 0.29             |
| HET                              | 3                      | 1.33               | 0.26             | 1.30         | 0.60             | 1.28         | 0.72             | 1.27         | 0.85             |
| FD                               | 3                      | 1.25               | -0.01            | 1.16         | 0.08             | 1.11         | 0.21             | 1.05         | 0.37             |
| HET                              | 4                      | 1.32               | 0.32             | 1.29         | 0.66             | 1.27         | 0.78             | 1.25         | 0.89             |
| FD                               | 4                      | 1.23               | -0.01            | 1.13         | 0.14             | 1.08         | 0.28             | 1.02         | 0.42             |
| HET                              | 6                      | 1.30               | 0.44             | 1.28         | 0.74             | 1.26         | 0.84             | 1.24         | 0.95             |
| FD                               | 6                      | 1.20               | 0.01             | 1.09         | 0.22             | 1.03         | 0.36             | 0.98         | 0.50             |
| HET                              | 10                     | 1.30               | 0.57             | 1.26         | 0.84             | 1.23         | 0.93             | 1.21         | 1.01             |
| FD                               | 10                     | 1.17               | 0.07             | 1.04         | 0.34             | 0.99         | 0.47             | 0.94         | 0.58             |

<sup>a</sup> HET = selection for heterozygosity of markers; FD = frequency-dependent selection.

#### 4. DISCUSSION

Molecular markers have received considerable attention in recent years as a tool to aid conservation of genetic variability in both captive and natural populations [2]. Amplification of DNA sequences by the polymerase chain reaction (PCR) offers a non-destructive means for genotyping endangered species. With this technique, microsatellite DNA markers have been considered the most useful for conservation programmes because they are highly informative and because of their codominant nature. Other markers such as RAPD and AFLP are also very promising owing to their simplicity and low cost, although generally they are dominant markers which are not yet included in the gene maps of domestic animal species. Until now, genetic markers have been used to calculate genetic distances between breeds, to resolve taxonomic uncertainties and to determine paternity. However, their application in practical conservation programmes of strains of domestic species is only beginning, and there is no example of conservation units where markers are routinely scored and utilized.

Probably the clearer and less controversial application of molecular markers in conservation genetics will be to identify distinct populations that need to be conserved and to infer the genetic relationships among the possible founders so that the initial animals that constitute the conserved population carry most of the genetic variability present in the population. A less studied issue is the usefulness of markers in delaying the inevitable loss of genetic variability in a population of limited size in the generations following its foundation.

**Table V.** Rate of homozygosity per generation of the genome ( $\Delta H_o$ , %), and of the markers ( $\Delta H_{o_m}$ , %), in a population of family structure  $N_s = 8$  sires and  $N_d = 24$  dams, chosen from a variable number of candidates of each sex using frequency-dependent selection, random mating (RM) and codominant or dominant bi-allelic markers with diverse initial frequencies [ $\text{fr}(M)$ ,  $1 - \text{fr}(M)$ ] (5 000 runs).

| Candidates of each sex per family | Markers per chromosome some | Codominant markers   |                  |                      |                  | Dominant markers     |                  |                      |                  |                      |                  |
|-----------------------------------|-----------------------------|----------------------|------------------|----------------------|------------------|----------------------|------------------|----------------------|------------------|----------------------|------------------|
|                                   |                             | $\text{fr}(M) = 0.5$ |                  | $\text{fr}(M) = 0.2$ |                  | $\text{fr}(M) = 0.5$ |                  | $\text{fr}(M) = 0.2$ |                  | $\text{fr}(M) = 0.8$ |                  |
|                                   |                             | $\Delta H_o$         | $\Delta H_{o_m}$ |
| 1                                 | 1                           | 1.30                 | -0.01            | 1.30                 | 0.04             | 1.32                 | 0.04             | 1.31                 | 0.02             | 1.33                 | 1.73             |
| 2                                 | 1                           | 1.29                 | -0.01            | 1.29                 | 0.00             | 1.29                 | -0.01            | 1.29                 | -0.01            | 1.31                 | 0.90             |
| 3                                 | 1                           | 1.28                 | 0.00             | 1.29                 | 0.00             | 1.29                 | -0.01            | 1.29                 | -0.01            | 1.30                 | 0.57             |
| 1                                 | 2                           | 1.28                 | 0.02             | 1.28                 | 0.42             | 1.29                 | 0.18             | 1.29                 | 0.29             | 1.32                 | 1.96             |
| 2                                 | 2                           | 1.25                 | 0.00             | 1.26                 | -0.02            | 1.26                 | 0.01             | 1.26                 | -0.02            | 1.30                 | 1.26             |
| 3                                 | 2                           | 1.25                 | 0.01             | 1.26                 | -0.01            | 1.26                 | -0.01            | 1.25                 | 0.00             | 1.29                 | 0.88             |
| 1                                 | 4                           | 1.25                 | 0.15             | 1.26                 | 0.87             | 1.27                 | 0.33             | 1.26                 | 0.64             | 1.31                 | 1.96             |
| 2                                 | 4                           | 1.21                 | -0.01            | 1.21                 | 0.08             | 1.23                 | 0.06             | 1.22                 | 0.06             | 1.28                 | 1.76             |
| 3                                 | 4                           | 1.20                 | -0.01            | 1.21                 | -0.01            | 1.22                 | 0.01             | 1.21                 | -0.01            | 1.27                 | 1.40             |
| 1                                 | 6                           | 1.23                 | 0.26             | 1.25                 | 1.05             | 1.26                 | 0.47             | 1.25                 | 0.78             | 1.30                 | 1.83             |
| 2                                 | 6                           | 1.18                 | 0.01             | 1.19                 | 0.28             | 1.21                 | 0.12             | 1.19                 | 0.19             | 1.28                 | 1.96             |
| 3                                 | 6                           | 1.17                 | -0.01            | 1.17                 | 0.09             | 1.19                 | 0.06             | 1.17                 | 0.05             | 1.26                 | 1.70             |
| 1                                 | 10                          | 1.20                 | 0.40             | 1.23                 | 1.15             | 1.23                 | 0.57             | 1.22                 | 0.90             | 1.28                 | 1.68             |
| 2                                 | 10                          | 1.14                 | 0.07             | 1.15                 | 0.60             | 1.18                 | 0.24             | 1.16                 | 0.43             | 1.26                 | 2.00             |
| 3                                 | 10                          | 1.12                 | 0.02             | 1.13                 | 0.33             | 1.16                 | 0.14             | 1.13                 | 0.21             | 1.24                 | 1.94             |

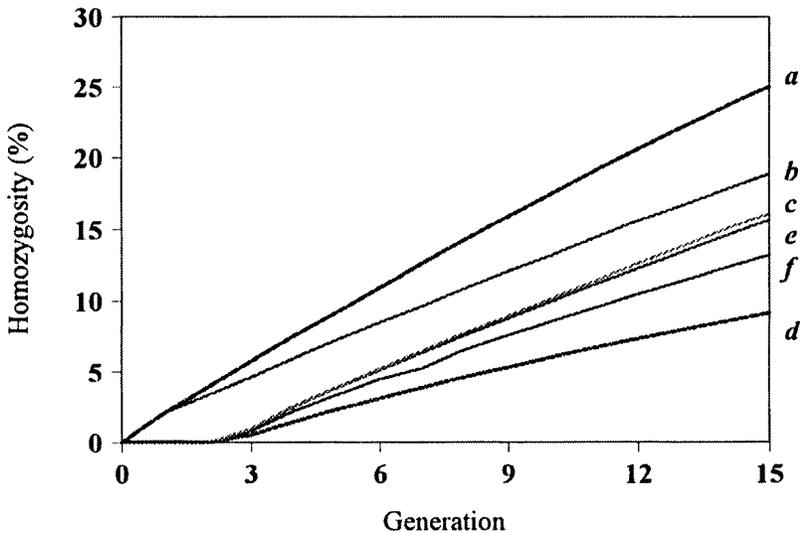
Monte Carlo simulation allows one to evaluate the gains expected with the use of these technologies. In the present work, we have studied a particular nucleus of small size mimicking the conservation programme carried out in strains of Iberian pig [16], but the conclusions could be generalized. Markers have been generated in linkage equilibrium, but this limitation is not very important: we have run some simulations with the parameters considered in *table III*, but assuming that base populations have undergone ten previous generations of random individual selection. As an example, with four markers and six alleles per marker and frequency-dependent selection, the rates of homozygosity (%) of the genome and of the markers were  $\Delta H_o = 1.02$  and  $\Delta H_{o_m} = 0.18$ , respectively, instead of the current values of  $\Delta H_o = 1.05$  and  $\Delta H_{o_m} = 0.27$  for a base population in linkage equilibrium, indicating that the efficiency of maintaining genetic variability will be improved, especially with respect to the markers.

The main measure of genetic variability that we have chosen is the global homozygosity by descent of all the genome calculated in all the candidates for selection. The homozygosity for the markers themselves would indicate the success of a conservation programme to maintain the variability at specific loci

of potential economic or biological interest. Another measure of the genetic variability used in conservation genetics is the effective number of alleles, which is inversely related to the expected homozygosity and therefore to the overall coancestry of the population. According to Allendorf [1], heterozygosity is a simple and accurate indicator of the loss in genetic variation and is a good measure of the ability of the population to respond to selection in the short term, whereas the effective number of alleles will be optimal for long term considerations and will be more affected by bottleneck effects.

When molecular information is used as a selection criterion, there is a disagreement between the true homozygosity by descent and the inbreeding coefficient calculated by pedigree analysis. Moreover, the rate of homozygosity, unlike the rate of inbreeding, does not attain an asymptotic value after the first generations but it will decrease as selection proceeds. Some theoretical work needs to be carried out on the prediction of homozygosity by descent under these circumstances.

Figure 3 summarizes the relative advantages of the diverse tactics analysed in this paper. When the molecular information is lacking, the first clear conclusion that appears in this study is that the use of conventional tactics such as restriction of family size is the most important criterion that should be considered in the genetic management of a conservation programme. Standardizing



**Figure 3.** Evolution of the homozygosity in a population of  $N_s = 8$  sires and  $N_d = 24$  dams chosen from 3  $N_d = 72$  candidates of each sex using different tactics: a) random choice of breeding animals, b) random choice with restriction on family size, c) random choice with restriction on family size plus minimum coancestry mating, d) frequency-dependent selection based on complete molecular information with restriction on family size plus minimum coancestry mating, e) the same as d but based on 24 candidates of each sex genotyped for two dominant markers, f) the same as d but based on 48 candidates of each sex genotyped for four microsatellites with six alleles.

family sizes is predicted to double effective population size and is widely recommended when breeding rare breeds [3, 11, 12, 18] and Brisbane and Gibson [5] proposed the minimization of the mean coancestry of individuals chosen for breeding as the optimal criterion for maintaining genetic variability. But the implementation of this criterion requires an iterative procedure which may be computationally expensive. However, if only full- and half-sib relationships are considered, this criterion would be the same as minimizing the variance of family sizes.

The use of minimum coancestry matings is another important tool for delaying the loss of heterozygosity and is especially efficient for maintaining the heterozygosity of the markers themselves. The advantage will disappear in the long term if there is a balanced family structure, but only after a very large number of generations. Furthermore, as variance of family size increases, the advantage of random mating will disappear even in the long term (see Caballero [6] for a discussion on this point).

When the use of molecular markers is considered in the framework of the traditional strategies of minimizing the variance of family sizes, frequency-dependent selection seems to be a more efficient criterion than selection for heterozygosity to minimize the increase in homozygosity either of all the genome or of the markers themselves. An additional advantage of frequency-dependent selection is that it can be readily applied to dominant markers such as RAPD or AFLP. However, there are many possible ways of implementing frequency-dependent selection. In this paper, we have followed the model of Crow [9]. Chevalet and Rochambeau [8] proposed an index equal to the inverse of the product of the frequencies of the alleles carried by the individuals. Simulation results not shown here indicate that this criterion, at least in our schemes, is inferior to the one utilized here. As indicated before, the optimal criterion would be to minimize the mean coancestry of the selected animals or to maximize the corresponding effective number of alleles ( $n_a$ ), calculated using the complete molecular information. In a small simulation example (only 100 runs) for the intermediate nucleus size and the more intense selection, the homozygosity value attained at generation 15 with this criterion was 11.26, slightly lower than the value 12.10 obtained with FD selection (*table 1*).

The simulation results obtained considering a limited number of markers and alleles per marker indicate that substantial gain in control of the increase of homozygosity from molecular information required expensive strategies with high genotyping costs (with respect to both individuals and markers). In *figure 3*, line *f* represents an interesting strategy for reducing homozygosity, but implies genotyping 96 individuals for  $4 \times 19$  highly informative markers (microsatellites). However, the joint use of PCR multiplexing, new fragment analysis technology and automatic sequencing based on fluorescent detection methods can reduce the cost of microsatellite genotyping. The strategy represented by line *e* is considerably cheaper, because it requires only 48 individuals to be genotyped for two dominant markers (RAPD or AFLP) per chromosome, but the benefits obtained are disappointing.

In summary, the use of molecular markers in conservation programmes does not seem to be a feasible option with the current costs and future application of these technologies to conservation programmes will depend basically on much lower costs. Other ways of diminishing costs, such as genotyping only some of

the individuals or only in alternate generations, could be of some value. In the meantime, some methodological questions remain to be investigated, such as the appropriate method of combining marker and pedigree information, or the potential values of other strategies, such as the use of a variable contribution of breeding individuals (weighted selection) which has been proved to be efficient in more typical selection schemes [14, 19].

## ACKNOWLEDGEMENTS

This work was supported by the EC Network ERBCHRXCT 94-0508. We thank M. Pérez-Enciso (IRTA, Lleida) and J. Fernández (INIA, Madrid) for their computing advice.

## REFERENCES

- [1] Allendorf F.W., Genetic drift and the loss of alleles versus heterozygosity, *Zoo Biol.* 5 (1986) 181–190.
- [2] Avise J.C., *Molecular Markers, Natural History and Evolution*, Chapman and Hall, New York, 1994.
- [3] Ballou J.D., Lacy R.C., Identifying genetically important individuals for management of genetic variation in pedigreed population, in: Ballou J.D., Gilpin M., Foote T.J.R. (Eds.), *Population Management for Survival and Recovery. Analytical Methods and Strategies in Small Population Conservation*, Columbia University Press, 1995, pp. 76–111.
- [4] Boichard D., Maignel L., Verrier E., The value of using probabilities of gene origin to measure genetic variability in a population, *Genet. Sel. Evol.* 29 (1997) 5–23.
- [5] Brisbane J.R., Gibson J.P., Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions, *Theor. Appl. Genet.* 91 (1995) 421–431.
- [6] Caballero A., Developments in the prediction of effective population size, *Heredity* 73 (1994) 657–679.
- [7] Chevalet C., Utilisation de marqueurs pour la sauvegarde de la variabilité génétique des populations, in: *Éléments de génétique quantitative et application aux populations animales*, Inra Prod. Anim., Paris, 1992, pp. 295–297.
- [8] Chevalet C., Rochambeau H., Variabilité génétique et contrôle des souches non-consanguines, *Sci. Tech. Anim. Lab.* 11 (1986) 251–257.
- [9] Crow J.F., *Basic Concepts in Population, Quantitative and Evolutionary Genetics*, Freeman and Company, New York, 1986.
- [10] Crow J.F., Kimura M., *An Introduction to Population Genetics Theory*, Harper and Row, New York, 1970.
- [11] Folch P., Jordana J., Demographic characterization, inbreeding and maintenance of genetic diversity in the endangered Catalonian donkey breed, *Genet. Sel. Evol.* 30 (1998) 195–201.
- [12] Frankham R., Conservation genetics, *Annu. Rev. Genetics* 29 (1995) 305–327.
- [13] Hammond K., Conservation of genetic diversity for animal improvement, *Proc. 5th World Congr. Genet. Appl. Livest. Prod. Guelph*, University of Guelph, Canada, vol. 21, 1994, pp. 423–430.
- [14] Meuwissen T.H.E., Maximizing the response of selection with a predefined rate of inbreeding, *J. Anim. Sci.* 75 (1997) 934–940.
- [15] Robertson A., Inbreeding in artificial selection programmes, *Genet. Res.* 2 (1961) 189–194.

- [16] Rodrigañez J., Silió L., Toro M.A., Rodriguez C., Fifty years of conservation of a black hairless strain of Iberian pigs, in: Matassino D., Boyazoglou J., Cappuccio A. (Eds.), International Symposium on Mediterranean Animal Germplasm and Future Human Challenges, EAAP Publication no. 85, Wageningen Press, 1997, pp. 183–186.
- [17] Santiago E., Caballero A., Effective size of populations under selection, *Genetics* 139 (1995) 1013–1030.
- [18] Smith C., Alternative forms of genetic control, *Anim. Prod.* 23 (1976) 403–412.
- [19] Toro M.A., Nieto B., A simple method for increasing the response to artificial selection, *Genet. Res.* 44 (1984) 347–349.
- [20] Toro M.A., Nieto B., Salgado C., A note on minimization of inbreeding in small scale breeding programmes, *Livest. Prod. Sci.* 20 (1988) 317–323.
- [21] Weir B., *Genetic Data Analysis II*, Sinauer, Massachusetts, 1996.
- [22] Woolliams J.A.N., Wray N.R., Thompson R., Prediction of long term contributions and inbreeding in populations undergoing mass selection, *Genet. Res.* 62 (1990) 231–242.