

Molecular characterization of breeds and its use in conservation [☆]

Miguel A. Toro ^{a,*}, Jesús Fernández ^a, Armando Caballero ^b

^a *Departamento de Mejora Genética Animal, INIA, Carretera de la Coruña Km. 7, 28040 Madrid, Spain*

^b *Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Universidad de Vigo, 36310 Vigo, Spain*

Abstract

The conservation of farm animal resources is important for coping with future breeding needs and for facilitating the sustainable use of marginal areas. The increasing availability of molecular markers for most farm animal species and the development of techniques to analyse molecular variation is widening our capacity to characterise the genetic variation of breeds. In this paper we review the most popular molecular markers used in conservation and animal breeding studies, the different measures of genetic diversity that they provide, and their application for managing within-breed genetic diversity and for setting between-breed conservation priorities. We also address the relationship between genomic and marker heterozygosity, the relationship between molecular and quantitative measures of genetic diversity, and the characterization of breeds based on non-neutral markers.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Molecular markers; Heterozygosity; Gene diversity; Allele diversity; Within- and between-breed genetic diversity; Conservation priorities

1. Introduction

The development of tools for the analysis of DNA that has taken place in the last few decades has increased enormously our capacity to characterise variation within and between breeds. The restricted traditional characterisation by means of phenotypic attributes can now be complemented by an increasing available number of molecular markers and the development of sophisticated statistical techniques for their analysis. We review these tools and their potential for helping in the maintenance of genetic variability both within and between breeds and

for contributing to establishing conservation priorities. More specifically we deal with the following topics: a) the types of molecular markers and the appropriate measures of molecular diversity; b) the relationship between genomic and marker heterozygosity; c) the optimization of within-breed genetic diversity; d) the different ways of analysing between-breed gene diversity: partition of total gene diversity, phylogenetic reconstruction and cluster analysis; e) the combination of genetic and other sources of information for setting priorities in conservation; f) the search for the hidden structure of a population; g) the relationship between molecular and quantitative measures of genetic diversity; and h) the characterization of breeds based on non-neutral markers.

2. Types of molecular markers

Molecular markers are DNA locations presenting different detectable variants. The first markers used in

[☆] This paper is part of a special issue entitled Animal Genetic Resources, Guest Edited by Ricardo Cardellino.

* Corresponding author. Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Carretera de la Coruña Km. 7, 28040 Madrid, Spain. Tel.: +34 913476807; fax: +34 913572293.

E-mail address: toro@inia.es (M.A. Toro).

livestock were the gene's coding protein polymorphisms (allozymes), but they present the limitation that both the number of loci that can be assayed and their polymorphism level are low. For this reason, as soon as DNA technologies were developed, markers at the DNA chain level (especially on nuclear material) quickly replaced allozymes as a tool to assess genetic diversity. Mitochondrial DNA (mtDNA) markers are very useful in phylogenetic analysis for several reasons. First, mtDNA is maternally inherited with no recombination and, therefore, the number of nucleotide differences between mitochondrial genomes reflects directly the genetic distance that separates them. Second, it mutates 5–10 times more frequently than nuclear DNA, thus allowing the study of the divergence between wild and domestic populations under the short time scale of domestication. For example, in sheep and goats there is one major geographical mtDNA lineage that probably represents initial domestication in the Fertile Crescent, with two more restricted lineages representing later independent domestication events (Bruford et al., 2003). Analogously to mtDNA, sequences of the Y chromosome provide similar information on paternal lineages. In horses, there is extensive matrilineal diversity seen in mtDNA variation that it does not match with the patrilineal diversity of the male specific Y chromosome. This is consistent with a strong sex-bias in the domestication process with only a few stallions contributing genetically to the domestic horses (Lindgren et al., 2004).

There are several types of nuclear DNA markers. The AFLPs are dominant biallelic markers that provide an easy way to carry out a genome-wide screening of variation. They have the disadvantage of a reduced power to analyse within-breed diversity due to the dominant mode of inheritance, but could be very useful in analysing between-breed variation. One potential problem of AFLPs is associated to size fragment homoplasy, i.e. that different loci with the same fragment length will be contributing to the same band, thus biasing upwards the estimates of allele frequencies (Vekemans et al., 1990; Caballero et al., 2008). This problem, however, is minimised if the number of loci analysed per primer combination is low (Gort et al., 2006).

Microsatellites have been the markers of choice to study genetic variation in the recent years. Based upon sites in which the same short sequence is repeated multiple times, they present a high mutation rate and codominant nature, making them appropriate for the study of both within- and between-breed genetic diversity.

Single Nucleotide Polymorphisms (SNPs) are point mutations in the genome sequence, predominantly biallelic and highly abundant throughout the genome. They have the potential to detect both neutral and functional genetic variation because, although most of them are located in non-coding regions, some correspond to mutations inducing changes in expressed genes. Large numbers of SNPs, however, are required for precision (it is said as a rule of thumb that about six SNPs are equivalent to one microsatellite). In addition, another critical aspect is their discovery, usually by sequencing techniques. Nevertheless, it seems that they are becoming the markers of choice (except perhaps for pedigree-assignment, but see Andersson and Garza, 2006) because of increasing automation coupled with low cost, and several large-scale projects are currently carried out to identify SNPs in, for example, chickens (Wong et al., 2004), pigs (Chen et al., 2007), and cattle (Van Tassell and Wiggans, 2007).

Sample collection is an important step in any molecular diversity study. The FAO Secondary Guidelines (FAO, 1998) required that at least 25 non-related animals per breed should be sampled and genotyped for at least 25 microsatellites with an effective allele number of two. Baumung et al. (2004) surveyed 87 projects on conservation with 50% of them studying more than eight breeds. In 96% of the projects, the sample size was larger than 25 individuals. The preferred tissue sample was blood, and in 90% of the projects microsatellites were used as genetic markers, with an average of 18 loci genotyped per project.

The comparative analysis of separate studies, although desirable, is highly problematic because most studies are limited to a small number of breeds from the same country and using different subsets of the FAO recommended markers. An exception to the large-scale analyses are the studies of the European Cattle Genetic Diversity Consortium (2006) in cattle, Hillel et al. (2003) in chickens, Peter et al. (2007) in sheep, Ollivier et al. (2005) and SanCristobal et al. (2006) in pigs, and Cañón et al. (2006) in goats. Recently, there have been some attempts to develop methods of meta-analysis of data sets that have only few breeds or few markers in common (Freeman et al., 2006).

3. Measures of molecular diversity

At the level of molecular markers, genetic diversity is usually measured by the frequencies of genotypes and alleles, the proportion of polymorphic loci, the observed and expected heterozygosity, or the allelic diversity. The most widely used parameter to measure diversity within

populations is the expected heterozygosity (H), or gene diversity, defined by Nei (1973) as the probability that two alleles chosen at random from the population are different. This can be calculated for a particular locus with k alleles as

$$H = 1 - \sum_{i=1}^k p_i^2,$$

where p_i is the population frequency of allele i . The sampling variance of the heterozygosity has two components: among individuals and among loci. Nei (1987) provided variances for these components and suggested that, for a fixed number of genotypes, it is preferable to sample more loci than more individuals, because variation among loci is generally larger than among individuals. It should be noticed that the expected heterozygosity is also equal to

$$1 - \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N f_{Mxy},$$

where f_{Mxy} is the molecular pairwise coancestry, or kinship, between individuals x and y and N is the number of individuals. Following Malécot (1948), this is the probability that two alleles at the locus taken at random from each individual (with replacement within individuals) are alike in state. Analogously, the observed heterozygosity (the actual frequency of heterozygotes in the population) is equal to

$$1 - \frac{1}{N} \sum_{x=1}^N F_{Mx},$$

where F_{Mx} is the molecular inbreeding of individual x , that is the probability that the two genes carried out by this individual at a given locus are alike in state.

A measure of within-individual genetic diversity specific to microsatellites is the statistic

$$d^2 = \frac{1}{L} \sum_{l=1}^L (n_{l1} - n_{l2})^2$$

(Coulson et al., 1998), where L is the number of microsatellite loci and n_{l1} and n_{l2} represent the number of repetitions of the two alleles of locus l carried by each individual. This statistic is related to the inbreeding coefficient (although it can be extended to the distance between individuals) because it gives an idea of the genetic distance between the two gametes producing each individual, assuming the stepwise mutation model.

The allelic diversity (number of alleles segregating in the population) is an alternative criterion to measure genetic diversity, and some authors (Petit et al., 1998; Notter, 1999; Barker, 2001; Simianer, 2005a; Foulley and Ollivier, 2006) consider that this parameter is of key relevance in conservation programmes. A high number of alleles imply a source of single-locus variation for important traits such as the major histocompatibility complex, which is responsible for the recognition of pathogens. Allelic diversity is also important from a long-term perspective, because the limit of selection response is determined by the initial number of alleles (Hill and Rasbash, 1986) and, because it is more sensitive to bottlenecks than expected heterozygosity, it reflects better past fluctuations in population size. In addition, allelic diversity and gene diversity can behave rather differently, for example, a higher differentiation between populations can be observed for allelic diversity than for gene diversity (Foulley and Ollivier, 2006).

When sequences of DNA are available, the variation can be estimated by the proportion of nucleotide sites that differ in the population, $\hat{p}_S = S/T$, where S is the number of nucleotide sites that are different (segregating) and T the total number of sites analysed. A second measure is the nucleotide diversity, the proportion of nucleotide differences between pairs of sequences weighting these differences by the frequencies of the sequences, i.e.

$$\pi = \sum_i \sum_j p_i p_j \pi_{ij},$$

where p_i is the frequency of sequence i and π_{ij} the proportion of nucleotides on which the two sequences differ. For example, in human populations, two randomly chosen individuals differ at about 1 SNP per kilobase, whilst for cattle and sheep the figure is 2–2.5, and for chicken is 4–5.5. Similar measures to those for nucleotide diversity can be calculated for the amount of amino acid variation in coded proteins.

4. Relationship between genealogical and marker estimates of diversity

With pedigrees, the usual way to estimate diversity is to calculate $1-F$ and $1-f$, where F (inbreeding coefficient) and f (coancestry or kinship coefficient) are the probabilities that two genes taken at random from the same or different individuals, respectively, are identical by descent (Malécot, 1948). As F and f values are dependent on an assumed reference population, a

key parameter in monitoring conservation programmes is the rate of change in coancestry or inbreeding (Δf or ΔF), the effective population size (N_e) being inversely related to them (reviewed by Caballero, 1994).

The molecular inbreeding of an individual, F_{Mi} , equals the genealogical inbreeding coefficient, F_i , for a model where all the alleles in the base or reference population are assumed to be different. However, they can differ otherwise, because identity by descent may not coincide with identity in state. The relationship between both parameters is given by the well-known expression

$$(1 - F_{Mi}) = (1 - F_i) \left(1 - \sum_{j=1}^k p_j^2 \right),$$

where the last term of the right hand side is the expected heterozygosity of the founder population.

It is clear that the genealogical inbreeding coefficient is directly related to the average genomic homozygosity. That means that the genealogical inbreeding (or coancestry) is a very good predictor of the molecular inbreeding (or coancestry) of the actual loci of the genome. Advantage is taken of this fact in genetic evaluations of a trait by using the additive relationship matrix as a surrogate of the proportion of genes with effect on the trait shared by relatives. However, the reverse is not true: the molecular inbreeding (or coancestry) measured with a handful of molecular markers is not necessarily a good predictor of the genealogical (or genomic) inbreeding (or coancestry). Furthermore, it is well known (M. Nei, quoted by Chakraborty, 1981) that there are problems in estimating genomic heterozygosity using only a few molecular markers, because the expected correlation between the heterozygosity of the genome and of a sample is approximately $\sqrt{r/n}$, where n is the number of loci in the genome and r the number of loci assayed. For example, 20 random markers from a genome of 20,000 genes would give an expected correlation of 0.03. This simple relationship seems to have been overlooked until the recent empirical reviews and theoretical updates of Slate et al. (2004), Balloux et al. (2004), Pemberton (2004) and DeWoody and DeWoody (2005).

To illustrate the above point, consider the following simple simulation. From an infinite base population 10 breeds were derived with census numbers 10, 10, 10, 10, 25, 25, 25, 50, 50, and 100 and allowed to diverge randomly for 10 generations. Twenty independent markers with six alleles/marker (in the base population) were evaluated at generation 10. Table 1 shows the value of the regression coefficient of molecular

Table 1

Regressions and correlation between genealogical and molecular coancestry or inbreeding from simulated data on ten breeds

	Regression molecular- genealogical	Regression genealogical- molecular	Correlation
Coancestry	0.84	0.39	0.57
Inbreeding	0.71	0.16	0.33

coancestry on genealogical coancestry, the regression coefficient of genealogical coancestry on molecular coancestry and the correlation between both, and analogously for the inbreeding coefficient. It is clear that the prediction of coancestry is better than the prediction of inbreeding, and that the prediction of molecular coancestry from genealogical coancestry is rather good (0.84) but the prediction of genealogical coancestry from molecular coancestry is not (0.39).

There are also empirical results on the above issues. Toro et al. (2002) compared the coancestry between 62 Iberian pigs from two related strains, using a pedigree going back 20 generations, with molecular coancestries estimated from 49 microsatellites. The correlation was high (about 0.90) when all animals were considered, but it decreased substantially to 0.60 and 0.40 when the two strains were analysed separately. Furthermore, the attempt to infer coancestries from molecular markers gave severely biased results, as the inference required information on the true allelic frequencies of markers in the base population that are usually unknown. Slate et al. (2004) examined 590 of the Coopworth sheep with seven generations of known pedigree and genotyped at 101 microsatellites. The correlation between inbreeding calculated from pedigree and homozygosity calculated from markers was remarkably low (0.17). However, Daetwyler (quoted by Woolliams and Toro, 2007), using 10,000 SNPs, found a very strong correlation between the heterozygosity calculated from pedigree and from markers in dairy cattle. The general message is that, for the correlation between the genealogical and the molecular inbreeding (coancestry) to be substantial, a considerable number of loci is required and, more importantly, a high variance of the genealogical inbreeding values should be present. This casts serious doubts about the methods proposed to estimate heritability using relationship inferred from markers in the absence of pedigree data (Ritland, 2000; Rodríguez-Ramilo et al., 2007). A second message is that it should be preferable to use pedigree information whenever available, and limiting the use of markers to verify, correct, complete or even implement pedigree recording (Fernández et al., 2005).

Recently, Flury et al. (2006a) have proposed the concept of *epistatic kinship*, later called (more appropriately) *haplotype kinship* by Flury et al. (2007). This is defined, in parallel to the Malecot's definition of coancestry, as the probability of identity by descent for a chromosome segment of a given length (within an individual or between pairs of individuals). Assuming that the number of crossovers follows a Poisson distribution, the probability of receiving the complete strand of length x (i.e., no crossover) is e^{-x} in the framework of Haldane's mapping function. From this model, and applying the same rules used in the calculation of pedigree-based coancestries and inbreeding coefficients, the haplotype inbreeding or kinship of any individual (pair) in a genealogy can be calculated. For example, the self-coancestry (K_j^x) is related to the coancestry of parents (s and d) by the expression $K_j^x = 0.5 (1 + e^{-2x} K_{sd}^x)$. However, in real-life data, pedigree of animals is rarely available and the identity-by-descent status of a chromosome segment must be assessed using markers sets spanning along this segment.

5. Optimization of within-breed genetic diversity

5.1. General considerations

Managing the rate of inbreeding (ΔF), or equivalently the effective population size, provides a general framework for managing genetic resources. The control of ΔF would restrict inbreeding depression, the probability of losing beneficial rare alleles, and the risk of extinction (Frankham et al., 2002). The operational aspects of managing a conservation scheme have been recently reviewed by Meuwissen (2007). The optimal solution is inspired by the considerable amount of past research that dealt with the development of selection and mating strategies for restricting the increase of inbreeding (e.g. Toro and Pérez-Enciso, 1990; Villanueva et al., 1994). There is not a sharp distinction between selection and conservation programmes in farmed populations. In selection programmes the aim is to not only maximise the increase in performance for economically valuable traits but also to impose restrictions on ΔF . In conservation programmes the aim is to minimise ΔF but generally some selection is imposed to avoid decreased performance in traits that make the breed valuable. In both cases two main decisions need to be taken, one referring to which animals should be used for breeding and how widely they should be used (i.e. selection decisions), the other referring to how the selected animals should be mated (i.e. mating decisions).

There is a consensus on the optimal way to face the first decision when the pedigree of the population is available (Caballero and Toro, 2000; Fernández et al., 2003). In this scenario, the best strategy is to optimize contributions of parents (number of offspring that each individual leaves to the next generation) by minimizing the global coancestry weighted by those contributions, i.e.

$$\frac{1}{4} \sum_{i=1}^{N_m} \sum_{j=1}^{N_m} \frac{c_i c_j f_{ij}}{N_m^2} + \frac{1}{2} \sum_{i=1}^{N_m} \sum_{j=1}^{N_f} \frac{c_i c_j f_{ij}}{N_m N_f} + \frac{1}{4} \sum_{i=1}^{N_f} \sum_{j=1}^{N_f} \frac{c_i c_j f_{ij}}{N_f^2},$$

where c_i is the contribution from individual i , f_{ij} is the coancestry between individuals i and j , and N_m and N_f are, respectively, the number of males and females evaluated. Optimal solutions for contributions are usually obtained via non exact algorithms (e.g. *simulated annealing*) that allow for the inclusion of restrictions frequent in practical breeding programs, such as a minimum number of sires, a fixed contribution for a particular set of males, or a maximum contribution for females. Fernández et al. (2008) have extended the minimum coancestry methodology to the management of subdivided populations integrating the control of individuals' contributions with the migration design.

The procedure of contributions of minimum coancestry maximizes the genetic diversity of the population in terms of expected heterozygosity and effective population size, it is very flexible and robust against departures from the ideal conditions (Fernández et al., 2003), and it is very effective in preserving the original distribution of allelic frequencies in conservation programmes (Saura et al., in press). The minimum coancestry mating system can be applied subsequently to selected animals resulting from the optimisation (Caballero et al., 1996; Sonesson and Meuwissen, 2000; Meuwissen, 2007), but the impact of the mating system on the maintenance of gene diversity is generally much lower than that from the selection decisions.

5.2. Incorporation of molecular genetic information for the management of conservation programmes

When only molecular markers (rather than genealogical) information is available, the optimal strategy for managing a conservation programme is to minimize the global molecular coancestry, as defined above, or to obtain an estimation of the genealogical coancestry from the markers and use it for the optimisation (Oliehoek

et al., 2006). When both genealogical and molecular information is available, it can be combined to calculate the coancestry conditional on markers (Toro et al., 1999; Wang, 2001; Fernández et al., 2005). In this way, markers can help to ascertain the global ‘realized’ coancestry from the ‘expected’ coancestry provided by the pedigree.

Earlier studies (Toro et al., 1999; Wang, 2001) showed that the use of coancestry conditional on markers to decide the selected offspring to be kept as breeders could yield effective population sizes up to 40% larger than those obtained using only pedigree coancestry. More recent analyses (Fernández et al., 2005) have shown that genealogical information, when properly used, proves to be very useful for arranging individuals’ contributions via the minimization of global coancestry. In fact, the levels of expected heterozygosity after 10 generations yielded by this strategy were 88–100% of the maximum possible improvement obtained if the genotype for all loci were known. Marker information was of very limited value if used alone. The amount and degree of polymorphism of markers to be used to compute molecular coancestry had to be high to mimic the performance of the strategy relying on pedigree, especially in the short-term (for example, more than 10 markers per Morgan with 10 alleles each were needed if only the parents’ genotypes were available). When both sources of information were combined to calculate the coancestry conditional on markers, clear increases in effective population size were found, but observed diversity levels (either gene or allelic diversity) in the early generations were quite similar to the ones obtained with pedigree alone. The advantage of including molecular information is greater when information can be available for a greater number of individuals (offspring and parents versus parents only). However, for realistic situations (i.e., large genomes) the benefits of using information on offspring are small. The same conclusions were reached when comparing the use of the different types of information (genealogical and/or molecular) to perform minimum coancestry matings.

The management of genetic diversity through molecular markers can be focused on the allelic diversity rather than on the gene diversity. In this context, it is known that, on the one hand, the population size required to maintain several alleles is much larger than that needed to keep inbreeding at acceptable rates (Denniston, 1978). But on the other hand, because the *effective number of alleles* is, by definition, the inverse of the mean coancestry (Crow and Kimura, 1970, p. 324), the strategy of maximising gene diversity keeps levels of allelic diversity as high as strategies

maximising allelic diversity itself, but with better control of inbreeding (Fernández et al., 2004).

As previously stated, optimisation of contributions of parents by minimising the average coancestry of the progeny is expected to maintain the largest possible levels of gene diversity. Because the maximum gene diversity occurs when the allele frequencies are equal, the effect of the optimisation method when applied to marker data is, in fact, homogenising the allelic frequencies at each locus (Fernández et al., 2004). This distortion in the distribution of allele frequencies of the original population may not have consequences for neutral genes, but could be relevant for genes maintained at low or high frequencies by selection (e.g. genes under artificial selection or natural adaptation). Saura et al. (in press) have developed an optimisation method aimed at minimising the distance (actually the *Kullback–Leibler distance*) between the original allele frequencies for a number of markers and their observed frequencies every generation. The method is compared with contributions of minimum coancestry using either genealogical or marker data. The results indicate that the proposed method is effective in maintaining the original distribution of allele frequencies, particularly when these are initially very unequal (e.g. under strong selection and linkage), and that also maintains low levels of average coancestry in the population. As expected, contributions of minimum coancestry based on markers maintain a lower average coancestry but at the cost of changing the allele frequency distributions of the conserved population. Thus, for example, traits that have been artificially selected in a previous breeding programme, lose part of their gain when they are conserved under contributions of minimum marker coancestry and relaxed selection, whereas they do not suffer losses in gain under the proposed method. Minimum coancestry based on genealogical data, however, is shown to be more effective in preserving the original distribution of allelic frequencies than the proposed method. The reason is that minimum coancestry based on pedigrees does not perform selection on marker frequencies, as it occurs with minimum coancestry based on markers, so it retains all the advantages of the minimum coancestry method without its negative consequences.

6. Partition of gene diversity within and between breeds

6.1. Total gene diversity

In a structured population with n breeds, the total gene diversity or expected heterozygosity (H_T) is

Table 2

Allele frequencies for a locus with nine alleles in a population consisting of four breeds

Allele	Breed I	Breed II	Breed III	Breed IV
1	0.5	0.9	0.125	0
2	0.1	0.1	0.125	0
3	0.1	0	0.125	0
4	0.1	0	0.125	0
5	0.1	0	0.125	0
6	0	0	0.125	0.2
7	0	0	0.125	0.3
8	0	0	0.125	0.5
9	0.1	0	0	0

partitioned into a component within breeds (H_S) and another ($H_T - H_S$) between breeds,

$$H_S = 1 - \frac{1}{n} \sum_i \left(\sum_k P_{i,k}^2 \right),$$

$$H_T = 1 - \sum_k \left(\sum_i \frac{P_{i,k}}{n} \right)^2,$$

where $p_{i,k}$ is the frequency of allele k in breed i . The between-breed component of genetic diversity ($H_T - H_S$) is also the average Nei's minimum distance between populations (\bar{D} ; see Caballero and Toro, 2002). In terms of the molecular coancestry of breeds i and j ,

$$f_{Mij} = \sum_k P_{i,k} P_{j,k},$$

with the distance between breeds i and j ,

$$D_{Mij} = \left(\frac{f_{Mii} + f_{Mij}}{2} \right) - f_{Mij}.$$

Defining the averages $\bar{f}_M = \frac{1}{n} \sum_i f_{Mii}$, $\bar{f}_M = \frac{1}{n^2} \sum_{i,j} f_{Mij}$, and $\bar{D} = \frac{1}{n^2} \sum_{i,j} D_{Mij}$, then $H_S = 1 - \bar{f}_M$, $\bar{D} = \bar{f}_M - \bar{f}_M$, and $H_T = 1 - \bar{f}_M$.

Wright's (1969) fixation index is simply the proportion of diversity between breeds relative to the total diversity,

$$F_{ST} = (H_T - H_S)/H_T = \bar{D}/H_T.$$

Because the value of F_{ST} cannot exceed the value of H_S , it is difficult to compare F_{ST} values across different species or studies or between different types of markers. For example, for highly variable loci, even when no alleles are shared between populations, F_{ST} may be low. Recently, Hedrick (2005) has proposed to standardise it

by the maximum level that can be obtained, $F_{ST(\max)} = (1 - H_S)/(1 + H_S)$, given the observed heterozygosity within breeds. Thus, $F'_{ST} = F_{ST}/F_{ST(\max)}$ is a measure of population differentiation relative to the maximum possible and, therefore, it allows the comparison of loci with different levels of variation such as allozymes and microsatellites, or mtDNA and Y-chromosome genes. Meirmans (2006) has extended the calculation of the standardised F_{ST} into an analysis of molecular variance (AMOVA) framework.

In order to illustrate the partition of gene diversity into components consider the following example of one marker locus with nine alleles, which frequencies in a group of four breeds are presented in Table 2 (a more practical application can be found in Fabuel et al., 2004 for pigs and Marletta et al., 2006 for horses).

The total heterozygosity of the population is $H_T = 0.796$, with components within breeds $H_S = 0.594$ and between breeds $\bar{D} = 0.202$. The amount of differentiation is $F_{ST} = 0.254$, the maximum possible being $F_{ST(\max)} = 0.255$, so that the standardised value is very close to one in this case ($F'_{ST} = 0.997$).

The contribution of each breed to the whole population diversity can be specified (see Caballero and Toro, 2002), and these contributions are shown in Table 3. For example, Breed III contributes the most to the within-breed diversity component but relatively little to the between-breed component, as it does not present much genetic distance with Breeds I and II. In contrast, Breed II contributes the least to the within-breed variation but substantially to the averaged between-breed variation.

Another way of studying the relevance of the different breeds to the overall diversity as a tool for establishing conservation priorities is, following Petit et al. (1998), to calculate the loss or gain of diversity if one breed is removed from the set (Table 4). For example, the removal of Breed III causes a decrease of 11.8% in the within-breed variation, but an increase of 3.7% in the between-breed variation. In contrast, the removal of Breed II involves a substantial gain in the

Table 3

Relative contribution (in %) of each breed to the within, between and total gene diversity

Breed	Within-breed diversity	Between-breed diversity	Total gene diversity
I	29	17	26
II	8	32	14
III	37	18	32
IV	26	33	28
	$H_S = 0.594$	$\bar{D} = 0.202$	$H_T = 0.796$

Table 4
Loss(-)/gain(+) of diversity (in %) when each breed is removed from the set

Breed removed	Within-breed diversity	Between-breed diversity	Total gene diversity	Order of priority for conservation	Weitzman loss of diversity	Weitzman order of priority for conservation
I	-4.4	+4.6	+0.2	3	-12	4
II	+17.3	-9.5	+7.8	4	-47	1
III	-11.8	+3.7	-8.1	2	-15	3
IV	-1.1	-10.1	-11.2	1	-46	2

within-breed variation (17.3%), but a loss (9.5%) in the between-breed component. The most valuable breed for conservation (priority order) turns out to be Breed IV, the removal of which implies a loss of both within- and between-breed gene diversity.

6.2. The Weitzman approach

Thaon d'Arnoldi et al. (1998) proposed to set conservation priorities for livestock breeds through the analysis of genetic distances by the Weitzman (1992) approach to measure the global diversity and the marginal loss of diversity attached to each breed. From a genetic distance matrix, Weitzman (1992) proposed a computationally intensive method of constructing hierarchical trees based on a form of maximum likelihood phylogeny conditional on the model. Thus, the contribution of an element to group diversity is proportional to the reduction in tree length caused by the removal of the element from the group. Since the first work of Laval et al. (2000) in pigs and Cañón et al. (2001) in beef cattle the Weitzman method has been widely used, being some of the recent applications those of Reist-Marti et al. (2003) to 49 African cattle breeds, Bennewitz et al. (2006) to 44 North Eurasian cattle breeds, Tapio et al. (2006) to 35 Northern Europe cattle breeds, and the European Cattle Genetic Diversity Consortium (2006) to 69 European cattle breeds.

Several authors have criticised the application of the Weitzman approach in the context of within-species diversity (Caballero and Toro, 2002; Eding et al., 2002; see also Chevalet et al., 2006), mainly on the grounds that the method ignores within-population variability, which is a crucial component of global diversity. Ignoring the within-group variability is a characteristic not only of the Weitzman method but also of all methods based only on genetic distances. In fact, one of the properties of the method (monotonocity in distance) is that the diversity in a set of populations should increase if the distance between populations increases. Thus, it will favour inbred populations with extreme allele frequencies, whereas the coancestry approach would favour non-inbred populations with an even distribution

of gene frequencies. Interestingly, in the European Cattle Genetic Diversity Consortium (2006) work the authors conclude that prioritization based on Weitzman diversity differs only slightly from prioritization based on the most homozygous breeds. It should be noted, however, that small inbred populations can be useful if they harbour unusual alleles. An application of the Weitzman method to set priorities for conservation is also shown in Table 4. Because the method only considers diversity between breeds, the Weitzman loss of diversity column correlates almost perfectly with the between-breed diversity one. Removal of any breed always produces a loss in genetic diversity, and Breeds II and IV are those producing the largest loss in diversity as they are those contributing more to the between-breed diversity (Table 3). Note that the order of priorities for conservation with the Weitzman method changes relative to the total gene diversity approach. For example, Breed II has the largest priority for conservation according to the Weitzman method, in spite of making the lowest contribution to within-breed variation (Table 3).

6.3. Phylogenetic reconstruction based on genetic distances and cluster analysis

Genetic distances estimated from polymorphic microsatellite markers have been the most popular method of choice to assess genetic diversity among populations. But genetic distances were developed with the concept of species in mind and ignore migration, which is an important feature in livestock populations. On the other hand, livestock populations have been domesticated and improved by man and, therefore, the divergence period is short and the role of mutation in creating differences must be small. Another important difference is that, when applied to breeds, genetic distance is a measure of *distinctiveness* at a given time, without reference to any model having generated the differences but, in contrast, in the population genetics approach, genetic distance is an estimate of parameters of the model underlying the generation of differences observed.

The behaviour of the different measures of genetic distances in the livestock context has been reviewed by Laval et al. (2002). They conclude that all distances strongly depend on the number of generations since the divergence and on the effective population size of the breeds and, therefore, no phylogeny can be inferred from the tree in the case of closely related breeds exhibiting different effective sizes. For this reason, it is generally admitted that, in dealing with breeds of farm animals, the interpretation of trees in terms of phylogeny can be misleading. However, some authors (e.g. Barker, 1999) have argued that phylogenetic diversity will provide the best objective criterion for making conservation decisions, i.e. breeds that are taxonomically distinct should be favoured for conservation. This approach presents several problems. First, genetic variation within populations is completely ignored. Second, it fails to take into account the fact that genetic distances vary greatly according to the marker used and the recent demographic history of the breed (e.g. whether it has passed through a population bottleneck). And third, construction of trees using admixed populations, as often happens in livestock, contradicts the principles of phylogeny reconstruction (Felsenstein, 1982). In fact, the phylogenetic trees are being substituted by Net graphs that do not assume that phylogeny is hierarchical (Hudson and Bryant, 2006).

The haplotype kinship has also been suggested as a parameter to differentiate phylogenetically close populations (Flury et al., 2006b). The hypothesis is that this metric is more sensitive to small phylogenetic distances caused by short-time since separation than conventional distance metrics, which are based on mutation and/or genetic drift as the diversity generating process. In a certain way, the haplotype kinship will measure the number of crossovers separating two populations. Flury et al. (2007) have recently carried out the first practical evaluation of the haplotype kinship to three populations of the Goettingen Minipig, obtaining microsatellite-based haplotype kinships within and between the three populations.

An alternative to genetic distances that is increasing in popularity is the cluster analysis. Traditional estimators of population structure rely on the *a priori* definition of populations. Recently, methods have been developed that try to divide the total sample of genotypes of a population into an unknown number of subpopulations (clusters). This allows the data themselves to define the population structure. The individuals are assigned (probabilistically) to populations, or jointly to two or more populations if their genotypes indicate that they are admixed. The methods also

estimate, for each individual, the fraction of its genome that belongs to each cluster without any prior information on the structure of the population. The algorithms are based on multi-locus genotypes and are solved by adopting a Bayesian approach computed using Markov Chain Monte Carlo methods and assuming multi-locus genotypes in Hardy–Weinberg and linkage equilibrium within each randomly mating subpopulation. Three programs are available: STRUCTURE (Pritchard et al., 2000), PARTITION (Dawson and Belkhir, 2001), and BAPS (Corander et al., 2004), and their differences are summarised by Pearse and Crandall (2004). This approach was first applied by Rosenberg et al. (2001) to 20 chicken breeds and later to pigs (Fabuel et al., 2004; García et al., 2006; Alves et al., 2006b), cattle (Kumar et al., 2003; Li et al., 2005), sheep (Álvarez et al., 2004; Tapio et al., 2005), and horses (Marletta et al., 2006). Interestingly, cluster analysis provides an alternative method of estimating genetic distances that does not assume isolation (Cañón et al., 2006). If $q_k(i)$ is the proportion of the genome of breed i that comes from cluster k , then the distance between breeds i and j can be defined as

$$d_S(i,j) = \sum_{k=1}^K |q_k(i) - q_k(j)| \frac{q_k(i) + q_k(j)}{2},$$

where K is the total number of clusters.

Rosenberg et al. (2001) have argued that genetically distinctive populations can be identified on the basis of how difficult it is to separate them from others. That is, if some populations were easier to separate into clusters than others with only a small number of markers, this could indicate the presence of distinctive multilocus genetic combinations in those populations that were easier to separate. Therefore, they suggest that the relative number of loci required for the correct clustering of several populations can be used as a way of identifying those that are genetically distinctive with respect to a collection. As an example, Alves et al. (2006a) applied cluster analysis to 24 Spanish wild boars and 170 Iberian pigs from diverse varieties genotyped for 36 microsatellites, and found that, for a correct clustering of Duroc pigs and wild boars, at least 14–16 markers were required.

7. Analysis of allelic diversity

Because allelic diversity is highly dependent on sample size, techniques have to be used to compensate for the different census size of populations (allelic richness). The most common method is rarefaction (El Mousadik and Petit, 1996). If N_{ik} represents the number

of copies of the k th allele from the sample of breed i and N_i represents the total number of genes sampled from that breed, the allelic richness at one locus is denoted as the expected number of different alleles that a sample had if the sample size had been g genes (usually the smallest sample size) instead of N_i . The expected number of different alleles in a sample of genes taken at random is then equal to

$$a_i = \sum_k (1 - P_{ik}),$$

where

$$P_{ik} = \binom{N_i - N_{ik}}{g} / \binom{N_i}{g}$$

is the probability that allele k does not occur in a sample of g genes chosen at random. An extension of the rarefaction method to count private alleles (alleles present in a given breed but absent from all the others) was defined by Kalinowski (2004) as

$$PR_i = \sum_k (1 - P_{ik}) \prod_{j \neq i} P_{jk}.$$

Foulley and Ollivier (2006) have recently proposed an alternative method, based on extrapolation, that consists of adding to the number of alleles actually observed in a sampled breed the expected number of alleles missing, given the number of genes examined in the sample and the allelic frequencies observed in the whole set of breeds. If K_i is the number of alleles in the sample of breed i , K is the total number of alleles sampled in the whole set of breeds and π_k the frequency of the k allele in the whole population, the allelic richness of breed i with a sample size N_i would be

$$R_i = K_i + \sum_k (1 - \pi_k)^{N_i},$$

where the summation is over the subset of alleles actually missing in the sample. Therefore, the allelic richness for each breed obtained with this method will be a value between the actual number of different alleles sampled in the breed, K_i , and the total number of alleles in the whole set of breeds, K . Foulley and Ollivier (2006) compared the two methods with microsatellite data of European pigs. They concluded that rarefaction may lack sensitivity to rare alleles when the sample size of reference is small and, therefore, they recommended extrapolation when the sample sizes of populations are either low or highly unbalanced.

Note that the rarefaction method reduces the observed richness and can be applied only when there are differences in sample sizes. However, the extrapolation method, that exceeds the observed richness, can be applied even when all samples have the same size. For example, consider the allele frequencies of the four breeds of Table 2 and assume that these are the true numbers of alleles and the true allelic frequencies. Simulations were run (100 replicates) in which samples of s genes were taken from each breed assuming the true allelic frequencies of Table 2 and the extrapolation allelic richness was calculated for each sample. The average number of alleles sampled and the corresponding extrapolation allelic richnesses are shown in Table 5 for different samples sizes. Note that the allelic richnesses of breeds become closer to one another as the sample size gets smaller. It may also be assumed that there is unbalanced sample sizes so that extrapolation should only be applied to the samples with the lowest sample sizes. For example, assume that the sample size for Breed II is $s=10$ whereas for the other breeds is $s=30$. The simulation results for this case gives an average number of alleles sampled in Breed II of 1.73 ± 0.04 and the corresponding extrapolation allelic richness is 4.87 ± 0.03 . An assumption of this approach is that all breeds are drawn at random from the same founder population (Foulley and Ollivier, 2006), so that alleles can be potentially present in all samples. Thus, if one allele is missing in a given sampled population, the reason is the low sampling size, ignoring the possibility that it could be truly lost by genetic drift or other reasons.

A partition of diversity within and between breeds can also be made considering allelic diversity instead of gene diversity. Foulley and Ollivier (2006) advocate the methodology proposed by Petit et al. (1998) for gene diversity, to be possibly adapted to allelic diversity. The

Table 5

Average number of alleles (A_i) found in a sample of size s and corresponding allelic richness (R_i) estimated by the extrapolation method of Foulley and Ollivier (2006), assuming true breed allelic numbers and frequencies from Table 2

Breed	$s=10$		$s=20$		$s=30$		$s=50$	
	A_i	R_i	A_i	R_i	A_i	R_i	A_i	R_i
I	4.34	6.02	5.33	5.98	5.83	6.05	5.99	6.02
II	1.55	4.78	1.86	3.89	1.94	3.18	1.99	2.56
III	5.85	7.20	7.44	8.11	7.86	8.35	8.00	8.31
IV	2.87	5.39	2.98	4.80	3.00	4.16	3.00	3.57

Averages were obtained from 100 replicates. Standard errors lower than 0.1.

contribution of each breed i to the total allelic richness would be

$$CTR_i = K - R(S/i),$$

where $R(S/i)$ is the allelic richness of the whole set of breeds excluding breed i . CTR_i lies between zero, when $R(S/i)=K$, i.e. when Breed i has no private alleles, and an integer equal to the number of private alleles of the breed. Therefore, only those breeds containing private alleles contribute to the total allelic richness with this definition. In the case of the example of Table 2, the only breed contributing to the allelic richness would be Breed I (see Table 6).

The contribution of breed i to the within-breed allelic richness is defined as

$$CWR_i = [R_i - \bar{R}]/(n - 1)$$

(Foulley and Ollivier, 2006), where R_i is the allelic richness of breed i , and \bar{R} is the average allelic richness of the whole set of breeds. The values of CWR_i are also shown in Table 6, showing that Breeds I and III contribute positively to the within-breed allelic richness, whereas Breeds II and IV contribute negatively. Analogously, the between-breed contribution can be obtained as $CBR_i = CTR_i - CWR_i$, but the meaning of this parameter is unclear. Obviously, CWR_i equals CBR_i with opposite sign in the frequent case where $CTR_i=0$ (i.e. the breed has no private alleles), and this seems to lack an intuitive justification. Finally, El Mousadik and Petit (1996) have also proposed a coefficient of allelic richness differentiation,

$$\rho_{ST} = 1 - \frac{(\bar{R} - 1)}{(K - 1)},$$

which would take a value of 0.531 in the example of Table 2 assuming that sample sizes are large.

Because, using the above definitions, only breeds with private alleles contribute to the allelic diversity of the system, the question arises as to whether other procedures could be defined that take into account the

Table 6

Contribution of each breed to the within (CWR), between (CBR) and total (CTR) allelic richness (assuming a large sample size) for the example of Table 2, using the proposal of Foulley and Ollivier (2006)

Breed	CWR	CBR	CTR
I	+0.42	+0.58	1.00
II	-0.92	+0.92	0.00
III	+1.08	-1.08	0.00
IV	-0.58	+0.58	0.00

Table 7

Relative contribution (in %) of the breeds defined in Table 2 to the within, between and total gene diversity, assuming that allele frequencies are equal within each breed

Breed	Within-breed diversity (H_S)	Between-breed diversity (\bar{D})	Total gene diversity (H_T)
I	29	20	27
II	17	33	20
III	31	14	28
IV	23	33	25
	$H_S=0.719$	$\bar{D}=0.146$	$H_T=0.865$

allelic diversity of breeds even if they do not have private alleles. Consider the following alternative. Because the larger the number of alleles the larger is the potential diversity of a breed, and because the maximal diversity occurs when alleles are at equal frequencies, a hypothetical situation can be assumed where all alleles present in a breed have identical frequencies. An estimate of gene diversity under this situation may take account of the allelic diversity, by considering the potentiality of each breed according to the number and type of alleles that it carries. For example, in the case presented in Table 2, assume that Breed I carries alleles 1, 2, 3, 4, 5 and 9 with frequencies 1/6 each, Breed II carries alleles 1 and 2 with frequencies 0.5 each, etc. Applying the same calculations as above, the total heterozygosity of the whole set of breeds is $H_T=0.865$, with components within breeds $H_S=0.719$ and between breeds $\bar{D}=0.146$, being the amount of differentiation $F_{ST}=0.169$. The proportional contributions of each breed to the whole population diversity and the loss or gain of diversity if one breed is removed are shown in Tables 7 and 8, respectively. Note that, according to this criterion, the prioritization of the breeds for conservation changes with respect to the gene diversity criterion (Tables 3 and 4). In particular, Breed I, whose removal implied a gain in total gene diversity (Table 4) and was third in priority, now produces a loss in diversity when removed (Table 8), becoming second in priority. This occurs in part because Breed I has the interesting feature of carrying the unique (private) allele 9.

Table 8

Loss(-)/gain(+) of diversity (in %) when each breed is removed from the set (assuming that allele frequencies are equal within each breed)

Breed removed	Within-breed diversity	Between-breed diversity	Total gene diversity	Order or priority for conservation
I	-4.4	+1.3	-3.1	2
II	+8.4	-6.7	+1.7	4
III	-6.0	+4.5	-1.5	3
IV	+2.0	-6.7	-4.7	1

8. Importance of within- vs. between-breed genetic diversity

The results obtained either using between-breed diversity or total diversity will produce different and sometimes opposite conservation priorities. An over-emphasis on between-breed variation may result in favouring inbred populations even though they might not contain specific interesting alleles, but an over-emphasis on within-breed variation will favour the largest breeds, which are commercially more valuable and, therefore, less endangered. Thus, some compromise should be attempted.

Caballero and Toro (2002) and Eding et al. (2002) suggested that a way to prioritise breeds is to calculate the contribution of each one to a pool of animals or a germplasm bank that would maximise its genetic diversity. These optimal contributions can also be applied considering a weighted combination of the within- and between-breed components of gene diversity, $\lambda H_S + \overline{D}$. Ollivier and Foulley (2005) considered a similar aggregate diversity index that makes use of the Weitzman approach as a measure of the between-breed diversity. They showed that, when the between-breed component of this aggregate is weighted by F_{ST} and the within-breed component by $(1 - F_{ST})$, it produces an outcome highly correlated with that using $\lambda = 1$. In some contexts, however, other weights would be desirable. For example, Piyasatian and Kinghorn (2003) suggested giving five times more weight to the variation between breeds than to that within breeds ($\lambda = 0.2$), reflecting the speed by which genetic change can be made across populations compared with selection within one large mixed population. In practice, the weights will depend on the scenario imagined for the medium term use of the genetic diversity. Variation between populations may be more important because the most valuable characteristics are likely to be those for which genes are fixed or at high frequencies within the breed displaying these characteristics (Reist-Marti et al., 2003). It will be also more important in the context of animal breeding, because it plays an essential role in the benefits derived from heterosis and complementarity (Ollivier and Foulley, 2005). Bennewitz and Meuwissen (2005) also proposed an automatic weight based on maximizing the total genetic variance of a hypothetical quantitative trait, which is equivalent to using a weighting factor of $\lambda = 0.5$.

An illustration of the optimal contributions of the four breeds of Table 2 to a germplasm bank using different weights for within- and between-breed gene diversity is shown in Fig. 1A. Breeds II and IV contribute more when the weight given to within-breed gene diversity is lower,

i.e. when the objective is to maximise the genetic distance, as these were those breeds contributing more to the between-breed gene diversity (see Table 3). In contrast, Breed III contributes more if the aim is to maximise the within-breed variation (larger λ), as this breed was that presenting the largest contribution to the within-breed diversity (Table 3). Interestingly, Breed I contributes more for intermediate values of λ than for low or large values of λ , as this breed was not a main contributor to neither the within-breed nor the between-breed gene diversity (Table 3). An analogous analysis of the contributions to gene diversity can be made under the assumption that allele frequencies are identical for all alleles carried by each breed (Tables 6 and 7). The results are similar to the previous ones, except that the contribution of Breed I is relatively larger and non-zero for a wider intermediate range of values of λ , at the expense of Breed III (Fig. 1B).

An attractive way of prioritizing breed is the ‘safe + 1’ approach proposed by Thaon d’Arnoldi et al. (1998). It consists of calculating the gene diversity of a safe core set formed by breeds that are considered to be safe

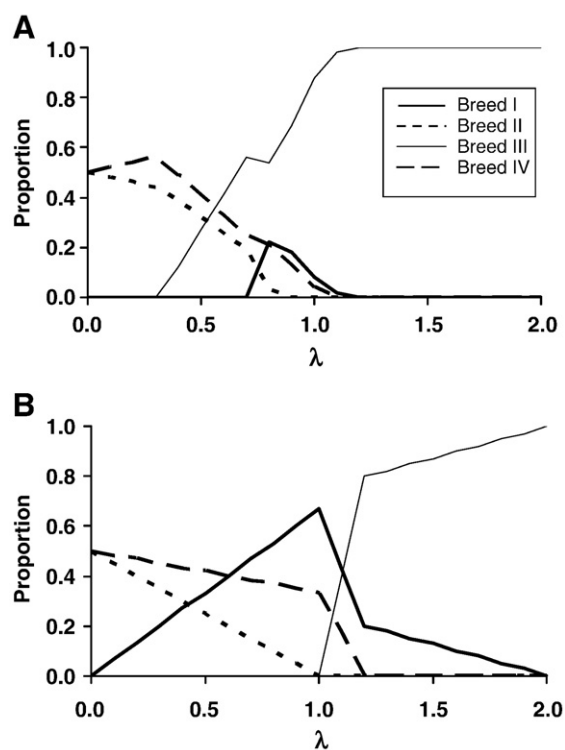


Fig. 1. Optimal contributions of the four breeds of Table 2 to a germplasm bank using different weights (λ) to the within (H_S) and between (\overline{D}) breed genetic diversity, such that $\lambda H_S + \overline{D}$. (A) Using the actual allele frequencies. (B) Assuming equal allele frequencies in each breed.

because they are commercial lines or are subject to conservation programmes. Then, the non-safe breeds are added one by one and the breeds that cause the largest increase in diversity obtain higher priority in the conservation plan. This approach has been applied to 69 European cattle breeds by the [European Cattle Genetic Diversity Consortium \(2006\)](#) using both the Weitzman and the optimal contributions methods to calculate the gain in diversity if a breed is added to a set of nine non-endangered breeds.

9. Probability of extinction

The genetic information of a set of breeds can be refined by introducing extinction probabilities. Note, however, that including extinction probabilities in the Weitzman method (which only considers between-breed diversity) will give an extra weight to the inbred populations, exacerbating their problems ([Eding et al., 2002](#)). [Gandini et al. \(2004\)](#) used different criteria to assess the degree of endangerment of a population and proposed to estimate the number of years needed to reach a critical population size, taking into account both demographic and genetic factors. A quantitative method inspired in conservation biology was used by [Bennewitz and Meuwissen \(2005\)](#). It involves a random process to predict likely future population size based on recent census data. However, when they applied the method to five German cattle breeds, the extinction probabilities were at the bounds of the parameter range (i.e. either close to zero or close to one) and the confidence intervals covered almost the whole parameter space.

[Simianer \(2005a\)](#) suggested the use of the criterion of the expected number of alleles segregating in the population after a given time period (usually a planning horizon of 50 years). He argued that this quantity would account for both the extinction probability of the populations and genetic drift. The extinction probability is assumed to be directly proportional to the rate of increase in inbreeding ($\Delta F=1/2N_e$), and the drift was calculated from the probability of loss of alleles. For 13 European cattle breeds (64 individuals per breed on average genotyped for 26 microsatellites) he assessed the effect of reducing genetic drift and changing extinction probabilities (for example paying a premium to keepers of animals of endangered breeds). Interestingly, he noticed that reducing drift within breeds is much more important than preventing breed extinction. Therefore, it might be better to focus on the management of diversity within a reasonably safe breed than to “rescue” breeds with a high risk of extinction. Similar results were obtained by [Simianer et al. \(2003\)](#) for

African cattle breeds, where the optimal allocation of funds to a larger set of breeds was 60% more efficient in terms of conserving diversity than spending the entire sum to save the three most endangered breeds from extinction after 50 generations. [Ollivier et al. \(2005\)](#) estimated the priority for cryopreservation of 29 local breeds of pigs as $CP=CB(i) P(i)$ where CB is the marginal between-breed diversity (according to Weitzman) and $P(i)$ the probability of extinction calculated by the expression $P(i)=1-\exp(-50/N_e)$.

Extinction probabilities could also be estimated in a more elaborated way. For example, in the analysis of 49 African cattle breeds, [Reist-Marti et al. \(2003\)](#) calculated extinction probabilities using four variables related with the population (population size, change over time, distribution of the breed and risk of indiscriminate crossing), four related with the environment (organisation among farmers, existence of a conservation scheme, political situation and reliability of the information), and two related with the value of the breed (presence of special traits and cultural value).

10. Setting priorities in livestock conservation

One of the most difficult issues in conservation of animal genetic resources is to give criteria for establishing priorities in conservation. The explosion of projects measuring genetic diversity does not explain much of the socio-cultural context in which the breeds exist. The conservation of animal genetic resources is not just one of the chapters of biological conservation but it must be justified by its present or future contribution to human livelihoods ([Rege and Gibson, 2003](#)). Two general scenarios can be considered. In the developed world the economic value lies in productivity, and the usefulness of conserved populations is far from obvious ([Hill and Zhang, 2004](#)) because introducing genetic variability from a conserved population into a commercial one is a lengthy process (and requires that the performance of the conserved population is not far behind the commercial one). If new selection objectives are to be introduced (such as adaptation to increasing temperatures) applying new selection indices to the actual commercial populations should be less costly. However, the last point is controversial and, as it was said before, crucial information of within versus between variance for economically valuable traits is lacking. For these authors, cultural reasons are the main justification for breed conservation. On the other hand, in low input systems in the developing world value lies on adaptation, survival, disease resistance, traction and nutrient recycling ([Rege and Gibson, 2003](#)).

The topic has recently been formally discussed by Bennewitz et al. (2007) and we follow their presentation. There are basically three strategies: the maximum-risk strategy, the maximum-diversity strategy and the maximum-utility strategy. In the first, the risk status is deduced mainly from the degree of endangerment classifying the breed in the categories given by FAO or EAAP (Gandini et al., 2004). This is the simplest approach and it has been criticised because there is not a formal way to include the specific value of a breed or its contribution to genetic diversity. In the maximum-diversity strategy, breeds are ranked according to the expected future diversity at the end of a time horizon (Simianer, 2005b; Bennewitz and Meuwissen, 2006). If there are N breeds with a known probability z_i of going extinct within a defined time horizon (e.g. 50 years), the effect of a breed could be assessed by how much the expected diversity would be changed if the extinction probability of the breed, z_i , would be changed by one unit by a conservation effort (md_i , marginal diversity; Weitzman, 1993). The marginal diversities can be multiplied by the extinction probabilities in order to obtain the conservation potentials of the breeds, i.e. $CP_i = md_i \times z_i$ (Weitzman, 1993). The conservation potential gives an idea of how much diversity can be conserved additionally if a particular breed would be made completely safe. The maximum-diversity strategy using the conservation potentials for prioritising breeds has been applied to several conservation plans and seems to work efficiently if diversity is the objective of the conservation plan (Reist-Marti et al., 2003; Simianer et al., 2003).

In the maximum-utility strategy the objective of the conservation plan includes, besides neutral diversity and special traits, other features related with the sustainable use of rural areas. In this case the utility conserved by a set of K non-extinct breeds at a defined future time horizon can be written as

$$U_K = w_D D_K + \sum_{j \in K} w_{F_j} + \sum_i k_i w_{B_i}$$

(Simianer et al., 2003), where U_K is the utility of the breed set K , w_D is the relative value of a unit of neutral diversity, D_K is the neutral diversity of the breed set K , w_{F_j} is the relative value of feature j (e.g. a special trait), $j \in K$ denotes that feature j is present in at least one of the non-extinct breeds, i.e. present in the set K , w_{B_i} is the relative value of breed i (the sustainable utilisation value of the breed) and, finally, k_i is an indicator variable that is 1 if breed i is present in the set K (i.e. not extinct) or zero otherwise (i.e. extinct at the end of the

time horizon), depending on its extinction probability. Here, a marginal utility is defined as the change in conserved utility at the end of the defined time horizon if the extinction probability would be lowered by one unit by a conservation effort. Similarly, a conservation potential of a breed with respect to the utility can be estimated as the product of the marginal utilities and the extinction probabilities and these can be used to select breeds for conservation.

Although the maximum-utility strategy seems to be the most sophisticated method for the selection of breeds, the problem is that estimates for the relative economic values of neutral diversity, for the special features (e.g. special traits) and for the breed specific values (e.g. the historical value of a breed) are needed (w_D , w_{F_i} and w_{B_i} , respectively) and there is no obvious way to obtain these relative weights. Some attempts have been made by Gandini and Villa (2003) to determine the cultural values of livestock populations as historical witnesses and custodians of local traditions. Definitely, more research is needed to obtain the economic weights for getting the full benefit out of the maximum-utility approach.

11. Traceability: assignment of animals and products to breeds

Molecular genetic information may be used to assign individuals correctly to a breed, especially when the phenotypic differentiation between breeds is difficult to detect or when genealogical information is absent. On the other hand, breed conservation usually emphasizes the maintenance of a pure breed and, consequently, there are cases where regulations exist against breed substitution for conservation and trade reasons. In this context molecular markers can detect whether introgression or crossbreeding has occurred. The assignment tests are also relevant to animal traceability from birth to market and that is increasingly requested as an element of a food safety assurance system and for improving the profitability of local breeds establishing a marketing link between product and breed.

The classical setting in assignment problems (often called supervised method) is to have a reference set of populations (in Hardy–Weinberg and linkage equilibrium) and the objective is to establish which is the most likely origin of a particular individual. It is important that these references are carefully identified and characterised with an adequate number of individuals. Following Baudoin et al. (2004) the simplest approach follows from the Bayes theorem: the probability that a given individual belongs to breed i , $P(b_i | g, p_i)$ is the

conditional probability of finding the genotype (g) of this individual in breed i (characterized by the gene frequencies p_i), divided by the sum of all the conditional probabilities for all n breeds,

$$P(b_i|g, p_i) = \frac{P(g|p_i)}{\sum_{j=1}^n P(g|p_j)}.$$

There are two possible types of errors: the type I error (α), or the proportion of individuals of one breed that are allocated to another breed, and the type II error (β), or the proportion of individuals that are allocated to one breed but are really of another breed. Blott et al. (1999) have found that an error rate smaller than 5% requires 11–18 microsatellites (or about 65–100 SNPs). The most discriminatory markers are those with the highest expected global heterozygosity and the largest number of alleles, especially the private alleles that are fixed for single breeds. The above approach has been applied in several settings: cattle breeds (Cañón et al., 2001), horses (Bjørnstad and Røed, 2001), and sheep (Bau-mung et al., 2006). The assignment of raw or dry-cured pig products has also been dealt by Alves et al. (2002) and García et al. (2006). The most widely used supervised method is that implemented in GENE-CLASS 2 (Piry et al., 2004), available at <http://www.ensam.inra.fr/URLB>.

The clustering methods (e.g. STRUCTURE above) also permit the assignment of individuals to populations or mixtures and are usually called unsupervised methods. The advantage of clustering is that it makes possible to take into account complex genetic situations, such as admixtures. However, besides being computationally demanding, when the number of populations is left as a parameter to be estimated, there are situations where the produced clusters do not represent real populations in the field.

12. Relationship between molecular and quantitative measures of genetic diversity

Quantitative genetic variation is the basis of productive and reproductive traits and, therefore, monitoring quantitative genetic diversity may reveal variation more closely related to fitness, yielding more interesting information (Lynch, 1996). An important question, for example, is the proportion of total genetic variation corresponding to variation between breeds (σ_B^2). For a particular trait, σ_B^2 gives an indication of how much progress in a trait may be obtained by selection among breeds, and its utilization has ultimately

led to breed substitution with many breeds considered unprofitable and consequently at risk of extinction. However, if crises were to occur that required livestock production to adapt quickly to new challenges, then conserving breeds with diversity of characteristics is a rational and important strategic response. Woolliams and Toro (2007) reviewed the few experiments of multibreed comparisons and concluded that the range of values for $\sigma_B^2/(\sigma_B^2 + \sigma_W^2)$ remains poorly documented but provide justification for the broad statement that breed variation accounts for approximately half of the total genetic variation for most traits.

When using molecular markers, estimates of gene diversity (expected heterozygosity, H) and allele frequency differentiation (F_{ST}) are usually intended as indirect ways to measure variation for adaptive polygenic traits (additive genetic variance, V_A , or heritability, h^2 , and population differentiation, Q_{ST} , respectively). The last parameter is a dimensionless measure of the quantitative genetic variance among populations, defined by Spitze (1993) as $Q_{ST} = V_B/(V_B + 2V_W)$, where V_W and V_B are, respectively, the additive within- and between-population components of the genetic variance for the trait considered. The relationship between the degree of divergence in neutral markers and the degree of divergence for quantitative traits can be addressed comparing F_{ST} and Q_{ST} indexes. For genes that are neutral for fitness, with additive action between and within loci for some trait, heterozygosity and additive variance behave in parallel. Accordingly, it is expected that $F_{ST} = Q_{ST}$. For traits under divergent selection pressure between populations, Q_{ST} is expected to be greater than F_{ST} whereas $Q_{ST} < F_{ST}$ would indicate that selection acts on the trait towards the same optimal phenotype. The above predictions are based exclusively on additive gene action for the quantitative trait and Hardy–Weinberg equilibrium. Non-additive genetic components and uncontrolled maternal and common environmental effects can potentially modify the expectations (López-Fanjul et al., 2003).

Notice also that the precision of estimates obtained from molecular markers and quantitative traits can be different (see Carvajal-Rodríguez et al., 2005). A single biallelic molecular marker gives estimates of F_{ST} with the same precision as Q_{ST} for a neutral polygenic quantitative trait, but the precision of the marker increases with the number of alleles. Using computer simulations, Le Corre and Kremer (2003) observed that, contrary to what might be expected *a priori*, estimates of variation obtained by directly studying the loci controlling the quantitative trait (QTL), are not necessarily closer to direct estimates obtained from the

quantitative measures than to those from neutral variation. This is a consequence of covariances of allelic effects generated by linkage disequilibrium among selected loci and contributing to differentiation for the quantitative trait that are not expressed in single-locus estimates. Thus, differentiation for QTLs might be not more informative than differentiation for neutral markers.

On the other hand, the empirical relationship between molecular variability and morphological, behavioural or life-history variability seems to be generally low. For example, Reed and Frankham (2001) carried out a meta-analysis based on 71 data sets (60 of allozymes) of molecular heterozygosities and genetic distances and quantitative measures of genetic variation. The mean correlation between molecular and quantitative estimates was weak (0.217 ± 0.05), indicating that molecular measures only explain 4% of the variation in quantitative traits. Furthermore, the correlation did not differ significantly from zero for life-history traits (-0.110) but was higher and significant for morphological traits (0.311 ± 0.052). Finally, there was no significant relationship with heritability (-0.08), considered the best indicator of adaptive potential. Merilä and Crnokrak (2001), McKay and Latta (2002), and Leinonen et al. (2008) have carried out meta-analyses involving studies on a variety of plant and animal species, showing that Q_{ST} was generally larger than F_{ST} . This result has been interpreted in the sense that a considerable part of the observed population divergence for quantitative traits should be attributed to differential selection pressures imposed by local environmental conditions.

13. Neutral vs. adaptive or deleterious variation: characterization of breeds with non-neutral markers

From a functional point of view DNA markers are grouped in two types: type I markers are DNA segments encoding for expressed DNA sequences which pose a relatively low degree of polymorphism but high evolutionary conservation; Type II markers have no identifiable biological function, are highly polymorphic and not well conserved between species. Microsatellites are typically type II markers although there is some evidence that they can act on gene regulation, transcription and protein function of the genes with which they are associated.

In population genetics, neutral variants are those whose selection coefficients are smaller than $1/(2N_e)$. This means that, in endangered populations, genetic variants are more likely to be effectively neutral because of their reduced population size. Strictly neutral variation would be of prime interest in order to carry out genetic analysis of population structure or history. It

allows for the identification of ancestral populations holding a possible source of alleles that are of economic value and which still may have been lost by chance during domestication. Neutral variants have also been used as a surrogate of non-neutral variation, either deleterious or adaptive. Although neutral and non-neutral diversity are expected to be correlated because of the disequilibrium generated by random drift or hitchhiking effects, their comparison could be also misleading because adaptive variants might differ in mutation rates and selective regimes (Hedrick, 2001).

As a large number of highly polymorphic markers are now available, the statistical power to detect differentiation between groups is very high. The problem is to know whether such differentiation reflects meaningful differences. In parallel, there could be no significant differences based on molecular markers but some important loci might be highly differentiated because selective forces are strong enough at such loci to overcome the effects of low effective size, gene flow or short divergence time.

Adaptive variation can provide new criteria and measurements to back-up conservation decisions. Differences between populations that are functional rather than neutral can be required, either for individual loci or genome regions. One way of approaching the problem is to use the existing type I markers (markers associated to known functional genes) to characterise the populations, as it is planned in recent biodiversity projects.

The second way of finding adaptive variation is to identify regions that have been subject to selection or, in other words, identifying signatures of selection among molecular markers. Unlike demographic processes, which affect the entire genome, selection affects specific important loci but will also affect closely linked sites, leaving its signature in the region: the level of variability will be reduced, the level of linkage disequilibrium will be increased and the genetic differentiation between populations will also be increased. This approach is particularly promising when we are able to compare the wild ancestor, local 'unimproved' breeds and highly selected lines, as is the case in the pig.

A matter of the utmost importance is how to distinguish the selection footprint from other demographic (neutral) processes. To do that, it is important to analyze several genome regions and characterize the expected variability under plausible models using coalescence based techniques (Hein et al., 2005). Livestock species have not been studied with as much detail as humans, although hapmap projects are being launched in the dog and in ruminants. In the latter, a set of non synonymous mutations is being genotyped in a

sample of sheep and cattle to look for SNPs with an outlier behaviour in terms of F_{ST} (Pariset et al., 2006).

When a locus shows extraordinary levels of differentiation between populations, measured for example by F_{ST} compared with other loci, this may be interpreted as evidence for selection of an allele in one of the populations. A classical test of neutrality (Lewontin and Krakauer, 1973) exploits this fact. The idea behind this is to compare the observed distribution of F_{ST} values or other population parameters from markers with that expected under the neutral hypothesis for different demographic scenarios, so as to identify those loci that significantly deviate from neutrality (Beaumont and Balding, 2004). Although the detection of outliers that show overdispersion of the F_{ST} values are usually considered, it could also be interesting to detect loci whose allelic frequencies are under-differentiated with respect to the neutral expectations, which would indicate adaptive differences important in all environments.

Outliers or departures from neutral expectations can also be assessed looking at the sequence diversity. The statistic Tajima's D measures the difference between two estimators of nucleotide diversity (the average number of nucleotide differences between pairs of sequences and the total number of segregating sites) that, under neutrality, should be approximately equal and, therefore, its expected value is zero. Positive values of this statistic arise from an excess of intermediate frequency alleles and can result from moderate population bottlenecks, structure and/or balancing selection. Negative values indicate an excess of low frequency alleles and can result from population expansion, strong population bottlenecks, or positive selection. Finally the levels of linkage disequilibrium will increase in selected regions although the pattern could be complex.

Mutation on the myostatin gene causing extreme muscle mass (double muscling) has been actively selected in several cattle breeds. Wiener et al. (2003) analysed three double-muscled and six non-double-muscled cattle breeds for this gene and 18 markers on the same chromosome. They found that there was a correlation between heterozygosity and the distance from the gene and that linkage disequilibrium was greater in double-muscled breeds. The results were not as consistent as would be expected and the authors considered the age and history of selection to be distorting the patterns. More recently, in sheep, three of eleven SNPs genotyped for about 1700 animals from 57 breeds lie outside the 95% confidence region of F_{ST} distribution (Pariset et al., 2006). Li et al. (2006) found slightly higher differentiation across candidate genes for production traits in cattle than across microsatellites. Finally, in Atlantic salmon, genetic signatures of divergent selection have been searched by screening

95 microsatellites in eight natural populations living in different habitats (salt-, brackish and freshwater). Nine of them exhibited highly significant deviations from neutral expectations becoming candidate genes associated with divergent adaptation (Vasemägi et al., 2005).

Putative adaptive markers could be removed from the computation of neutral differentiation and used as indicators of adaptive differentiation. It seems likely that the characterisation of diversity in future works will include an increasingly high use of adaptive variation, through the analysis of specific genes, quantitative traits or outlier markers, in combination with neutral variation.

Adaptive variation can provide new criteria and measurements to back-up conservation decisions. In the context of natural populations, Bonin et al. (2007) have developed a Population Adaptive Index (PAI) to give priority to populations for conservation. The PAI is calculated as the percentage of adaptive loci (loci detected as outliers with respect to the neutral prediction) with allelic frequencies significantly different from those in the other populations. They apply the methodology to six populations of the common frog and seven populations of dragonhead genotyped for 392 and 87 AFLPs respectively. Fourteen loci in each species were detected as potentially adaptive. They considered which set of populations had to be selected to maximize the fraction of total diversity (neutral or adaptive) protected. Because the neutral and adaptive diversity within and among populations were not correlated, conservation strategies based on both types of loci would not select the same populations for conservation. Furthermore, the protection of just one population was enough to preserve almost 75% of the neutral variation whereas three populations had to be protected to preserve 75% of the adaptive diversity.

Bioinformatics can also be used to predict the deleterious effects of SNPs on proteins. For example, analysis of more than 30,000 human SNPs show that about 0.01% change the subcellular localization of proteins and are rarely found in human populations, indicating that they are detrimental. On the contrary, about 50% affect protein stability and are comparatively benign. It has been suggested that populations with elevated SNP levels at sites which bioinformatics predicted detrimental consequences, would be considered genetically most at risk (Kohn et al., 2006).

14. New directions in future studies

From now on, studies dealing with the conservation and management of breeds will have to take careful account of the kind of information on which to base

decisions (genealogies, molecular markers, quantitative traits, demography, etc.), the measures of diversity that really make sense and the importance of preserving particular breeds, in terms of the utility to human livelihood.

We predict that SNPs will quickly become the marker of choice because SNPs are well suited to the high-throughput genotyping required for large studies, and because SNP genotyping error rates are low. Furthermore, the advantages of SNPs are particularly relevant in situations where several laboratories share the genotyping effort, as standardization of microsatellite allele calls across all the labs is costly and frequently infeasible.

Besides the classical measures of diversity, such as expected heterozygosity or molecular coancestry, other estimators associated with the increasing prevalence of sequencing techniques, such as nucleotide or haplotypic diversity, will gain in popularity. Recently proposed tools such as the *haplotype kinship* or the analysis of multi-locus inbreeding coefficients (Hernández-Sánchez et al., 2004), dependent on the number of crossovers and thus containing historical information on the short-term separation of breeds, are expected to be more widely applied if high density marker maps are available.

Molecular markers such as microsatellites are very well suited for solving individual or breed assignment problems, but their use on the management of within-breed conservation programmes is arguable. There is a consensus that pedigree information is the main tool for such a management when the aim is to control the increase in inbreeding. The general strategy is to minimize global coancestry among parents weighted by the optimized contributions. The main use of molecular markers will be, at least in the near future, to correct, complete or implement pedigree recording. However, it is expected that, in the future with thousands of markers genotyped, the molecular coancestry coefficient will be in itself the most relevant measure.

The classical approaches of using genetic distances and phylogenetic reconstruction to analyse genetic diversity between breeds may present several problems because they were developed with the concept of species in mind. In fact, phylogenetic trees are being substituted by Net graphs (Hudson and Bryant, 2006). Another approach that is gaining popularity is the cluster analysis, which allows us to verify the correctness of the assumed structure of the breed or to suggest an alternative one. However, it should be emphasized that this type of analysis is highly dependent on the number of markers used.

Although population genetics offers standard techniques to partition gene diversity within and between

breeds and to calculate the contribution of each breed to the whole population diversity (e.g. Toro et al., 2006), these techniques have been utilized rarely. However, other techniques such as the Weitzman approach (rarely mentioned in conservation of wild life) have been widely utilised in spite of being questioned because it ignores within-breed variability. Although it is generally accepted that between-breed diversity is more important than within-breed diversity (precisely this is one of the main reasons for conservation), the last could be important in some settings and, therefore, recent developments of the Weitzman approach do include this component. Perhaps the best way of assessing the genetic value of a breed is to consider how much diversity the breed adds to a core set constituted by commercial lines or breeds that are already subject to successful conservation programmes.

Livestock breeds, being a human construction, must be preserved because of their contribution to the human livelihood, now or in the future, or because of their cultural value as a historical witness. Therefore, there is a variety of objectives for conserving them but only just recently are economic arguments being used. The utility of a conserved set of breeds should include the relative features (e.g. special traits) of the breeds, their relative value for sustainable use in rural areas and the relative value of neutral genetic diversity. A formal approach to the problem is possible, although it is hard to calculate the relative economic weight of the above three components. Especially controversial is, in the context of the present review, the question of whether we should only consider neutral genetic variation. Although neutral and non-neutral diversity are expected to be correlated, this claim has not been verified in empirical studies. Adaptive variation, based on functional rather than neutral differences between populations, can provide new criteria to back-up conservation decisions. There are two ways of approaching the problem. The first is examining known genes such as those affecting growth and reproduction. The second way to picture the adaptive variation is to identify regions that have been subject to artificial or natural selection. Comparing, for example, the observed distribution of F_{ST} (population differentiation) for different loci with the values expected for neutral loci, to identify the outlier markers which deviate significantly from neutrality. These putative adaptive markers can be removed from quantifying the neutral differentiation and used as indicators of adaptive differentiation.

It seems likely that the characterisation of genetic diversity in future research will increasingly use adaptive variation, through the analysis of specific genes or outlier markers, as well as through quantitative traits in

combination with neutral variation. Differentiation between populations for quantitative traits (Q_{ST}) as compared to differentiation for neutral molecular markers (F_{ST}) is one of the standard tools for detecting diversifying selection in evolutionary studies and conservation of natural populations (see, e.g., [Toro and Caballero, 2005](#)). This is a tool that may be also considered for setting conservation strategies for farm animal breeds, particularly if differentiation for traits of economic or cultural value is considered. The problem is that estimation of genetic components of variation requires randomisation of breed members over farms in order to avoid environmental components of bias attached to farm environments.

The most recent technologies address the investigation of the transcriptome (the set of all expressed sequences in a tissue) and the proteome (the set of all proteins) by different methods. Microarrays, for example, allow us to compare the levels of mRNA expression of thousands of genes between two types of animals, or the same animal at different times, or exposed to different treatments. They have already been used in the animal breeding context, particularly for the detection of variation related to the immune system and susceptibility to pathogens (e.g. [Moser et al., 2004](#); [Donaldson et al., 2005](#)). The power of these techniques is limited by the difficulties of sampling, because RNA is much more sensitive to degradation than DNA, the high cost of equipment, the false positives, and the problems for statistical analysis of a huge amount of data. However, some authors think that they will be relevant in the next conservation genetic agenda based on functional rather than neutral genetic variation ([Kohn et al., 2006](#)).

Acknowledgments

The authors thank Louis Ollivier, Jean-Louis Foulley, Luis Silió, María Saura, and Humberto Quesada for helpful comments on the manuscript, and to Jeffrey G. Beasley for the English corrections. This work was supported by grants CPE03-004-C2 (Plan Estratégico del INIA, Spain) and CGL2006-13445-C02/BOS (Ministerio de Educación y Ciencia, Spain).

References

- Álvarez, I., Royo, L.J., Fernández, I., Gutiérrez, J.P., Gómez, E., Goyache, F., 2004. Genetic relationships and admixture among sheep breeds from Northern Spain assessed using microsatellites. *J. Anim. Sci.* 82, 2246–2252.
- Alves, E., Castellanos, C., Óvilo, C., Silió, L., Rodríguez, M.C., 2002. Differentiation of the raw material of the Iberian pig meat industry based on the use of amplified fragment length polymorphism. *Meat Sci.* 61, 157–162.
- Alves, E., Barragán, C., Fernández, A., Rodríguez, M.C., Silió, L., 2006a. Success rate of genetic clustering of domestic and wild pigs as a function of the number of markers. *Proceeding of the 8th World Congress on Genetics Applied to Livestock Production*, Belo Horizonte, Brasil.
- Alves, E., Fernández, A.I., Barragán, C., Óvilo, C., Rodríguez, C., Silió, L., 2006b. Inference of hidden population substructure of the Iberian pig breed using multilocus microsatellite data. *Span. J. Agri. Res.* 4, 37–46.
- Anderson, E.C., Garza, J.C., 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics* 172, 2567–2582.
- Balloux, F., Amos, W., Coulson, T., 2004. Does heterozygosity estimate inbreeding in real populations? *Mol. Ecol.* 13, 3021–3031.
- Barker, J.S.F., 1999. Conservation of livestock breed diversity. *AGRI* 25, 33–43.
- Barker, J.S.F., 2001. Conservation and management of genetic diversity: a domestic animal perspective. *Can. J. For. Res.* 31, 588–595.
- Baudoin, L., Piry, S., Cornuet, J.M., 2004. Analytical Bayesian approach for assigning individuals to populations. *J. Heredity* 95, 217–224.
- Baumung, R., Simianer, H., Hoffmann, I., 2004. Genetic diversity studies in farm animals—a survey. *J. Anim. Breed. Genet.* 121, 361–373.
- Baumung, R., Cubric-Curik, V., Schwend, K., Achmann, R., Sölkner, J., 2006. Genetic characterisation and breed assignment in Austrian sheep breeds using microsatellite marker information. *J. Anim. Breed. Genet.* 123, 265–271.
- Beaumont, M.A., Balding, R.A., 2004. Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13, 969–980.
- Bennewitz, J., Meuwissen, T.H.E., 2005. Estimation of extinction probabilities of five German cattle breeds by population viability analysis. *J. Dairy Sci.* 88, 2949–2961.
- Bennewitz, J., Meuwissen, T.H.E., 2006. Breed conservation priorities derived from contributions to the total future genetic variance. *Proc. 8th World Congress on Genetics Applied to Livestock Production*, CD-Rom Communication No. 33-06.
- Bennewitz, J., Kantanen, J., Tapio, I., Li, M.H., Kalm, E., Vilkkii, J., Ammosov, I., Ivanova, Z., Kiselyova, T., Popov, R., Meuwissen, T.H.E., 2006. Estimation of breed contributions to present and future genetic diversity of 44 North Eurasian cattle breeds using core set diversity measures. *Genet. Sel. Evol.* 38, 201–220.
- Bennewitz, J., Eding, H., Ruane, J., Simianer, H., 2007. Selection of breeds for conservation. In: Oldenbroek, K. (Ed.), *Utilisation and Conservation of Farm Animal Genetic Resources*. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 131–146.
- Bjørnstad, G., Røed, K.H., 2001. Breed demarcation and potential for breed allocation of horses assessed by microsatellite markers. *Anim. Genet.* 32, 59–65.
- Blott, S.C., Williams, J.L., Haley, C.S., 1999. Discriminating among cattle breeds using genetic markers. *Heredity* 82, 613–619.
- Bonin, A., Nicolè, F., Pompanon, F., Miaud, C., Taberlet, P., 2007. The Population Adaptive Index: a new index to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv. Biol.* 21, 697–708.
- Bruford, M.W., Bradley, D.G., Luikart, G., 2003. DNA markers reveal the complexity of livestock domestication. *Nat. Rev., Genet.* 4, 901–910.
- Caballero, A., 1994. Developments in the prediction of the effective population size. *Heredity* 73, 657–679.
- Caballero, A., Toro, M.A., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res.* 75, 331–343.

- Caballero, A., Toro, M.A., 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* 3, 289–299.
- Caballero, A., Quesada, H., Rolán-Alvarez, E., 2008. Impact of AFLP fragment size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics* 179, 539–554.
- Caballero, A., Santiago, E., Toro, M.A., 1996. Systems of mating to reduce inbreeding in selected populations. *Anim. Sci.* 62, 431–442.
- Cañón, J., Alexandrino, P., Bessa, I., Carleos, C., Carretero, Y., Dunner, S., Ferran, N., García, D., Jordana, J., Laloe, D., Pereira, A., Sánchez, A., Moazami-Goudarzi, K., 2001. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genet. Sel. Evol.* 33, 311–332.
- Cañón, J., García, D., García-Atance, M.A., Obexer-Ruff, G., Lenstra, J.A., Ajmone-Marsan, P., Dunner, S., the ECONOGENE Consortium, 2006. Geographical partitioning of goat diversity in Europe and the Middle East. *Anim. Genet.* 37, 327–334.
- Carvajal-Rodríguez, A., Rolán-Alvarez, E., Caballero, A., 2005. Quantitative variation as a tool for detecting human induced impacts on genetic diversity. *Biol. Conserv.* 124, 1–13.
- Chakraborty, R., 1981. The distribution of the number of heterozygous loci in an individual in natural populations. *Genetics* 98, 461–466.
- Chen, K., Baxter, T., Muir, W.M., Groen, M.A., Schook, L.B., 2007. Genetic resources, genome mapping and evolutionary genomics of the pig (*Sus scrofa*). *Int. J. Biol. Sci.* 3, 153–165.
- Chevalet, C., Nikolic, N., SanCristobal, M., 2006. Effects of genetic drift and mutations on some measures of genetic diversity in livestock populations. Proc. of the 8th World Congress on Genetics Applied to Livestock Production, CD-Rom Communication No. 33-07.
- Corander, J., Waldmann, P., Marttinen, P., Sillanpää, M.J., 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20, 2363–2369.
- Coulson, T.N., Pemberton, J.M., Albon, S.D., Beaumont, M., Marshall, T.C., Slate, L., Guinness, F.E., Clutton-Brock, T.H., 1998. Microsatellites reveal heterosis in red deer. *Proc. R. Soc. Lond., B* 265, 489–495.
- Crow, J.F., Kimura, M., 1970. *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- Dawson, K.J., Belkhir, K., 2001. A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genet. Res.* 78, 59–77.
- Denniston, C., 1978. Small population size and genetic diversity: implications for endangered species. In: Temple, S.A. (Ed.), *Endangered Birds: Management Techniques for Preserving Threatened Species*. University of Wisconsin Press, Madison, Wisconsin, pp. 281–289.
- DeWoody, Y.D., DeWoody, J.A., 2005. On the estimation of genome-wide heterozygosity using molecular markers. *J. Heredity* 96, 85–88.
- Donaldson, L., Vuocolo, T., Gray, C., Strandberg, Y., Reverter, A., McWilliam, S., Wang, Y.-H., Byrne, K., Tellam, R., 2005. Construction and validation of a Bovine innate immune microarray. *BMC Genomics* 6, 135.
- Eding, H., Crooijmans, P.M.A., Groenne, M.A.M., Meuwissen, T.H.E., 2002. Assessing the contribution of breeds to genetic diversity in conservation schemes. *Genet. Sel. Evol.* 34, 613–633.
- El Mousadik, A., Petit, R.J., 1996. Chloroplast DNA phylogeography of the argan tree of Morocco. *Mol. Ecol.* 5, 547–555.
- European Cattle Genetic Diversity Consortium, 2006. Marker-assisted conservation of European cattle breeds: an evaluation. *Anim. Genet.* 37, 475–481.
- Fabuel, E., Barragán, C., Silió, L., Rodríguez, M.C., Toro, M.A., 2004. Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity* 93, 104–113.
- FAO, 1998. *Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans—Measurement of Domestic Animal Diversity (MoDAD)*, FAO.
- Felsenstein, J., 1982. Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* 57, 379–404.
- Fernández, J., Toro, M.A., Caballero, A., 2003. Fixed contributions designs versus minimization of global coancestry to control inbreeding in small populations. *Genetics* 165, 885–894.
- Fernández, J., Toro, M.A., Caballero, A., 2004. Managing individuals' contributions to maximize the allelic diversity maintained in small, conserved populations. *Conserv. Biol.* 18, 1–10.
- Fernández, J., Toro, M.A., Caballero, A., 2008. Management of subdivided populations in conservation programs: Development of a novel dynamic system. *Genetics* 179, 683–692.
- Fernández, J., Villanueva, B., Pong-Wong, R., Toro, M.A., 2005. Efficiency of the use of molecular markers in conservation programmes. *Genetics* 170, 1313–1321.
- Flury, C., Täubert, H., Simianer, H., 2006a. Extension of the concept of kinship, relationship, and inbreeding to account for linked epistatic complexes. *Liv. Sci.* 103, 131–140.
- Flury, C., Tietze, M., Simianer, H., 2006b. Epistatic kinship a new measure of genetic diversity for short-term phylogenetic structures— theoretical investigations. *J. Anim. Breed. Genet.* 123, 159–171.
- Flury, C., Weigend, S., Ding, X., Täubert, H., Simianer, H., 2007. Haplotype kinship for three populations of the Goettingen minipig. *Genet. Sel. Evol.* 39, 159–179.
- Foulley, J.L., Ollivier, L., 2006. Estimating allelic richness and its diversity. *Liv. Sci.* 101, 150–158.
- Frankham, R., Ballou, J.D., Briscoe, D.A., 2002. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.
- Freeman, A.R., Bradley, D.G., Nagda, S., Gibson, J.P., Hanotte, O., 2006. Combination of multiple microsatellite data sets to investigate genetic diversity and admixture of domestic cattle. *Anim. Genet.* 37, 1–19.
- Gandini, G.C., Villa, E., 2003. Analysis of the cultural value of local livestock breeds: a methodology. *J. Anim. Breed. Genet.* 120, 1–11.
- Gandini, G.C., Ollivier, L., Danell, B., Distl, O., Georgoudis, A., Groeneveld, E., Martyniuk, E., Arendonk, J.A.M. van, Woolliams, J.A., 2004. Criteria to assess the degree of endangerment of livestock breeds in Europe. *Liv. Prod. Sci.* 91, 173–182.
- García, D., Martínez, A., Dunner, S., Vega-Pla, J.L., Fernández, C., Delgado, J.V., Cañón, J., 2006. Estimation of the genetic admixture composition of Iberian dry-cured ham samples using DNA multilocus genotypes. *Meat Sci.* 72, 560–566.
- Gort, G., Koopman, W.J.M., Stein, A., 2006. Fragment length distributions and collision probabilities for AFLP markers. *Biometrics* 62, 1107–1115.
- Hedrick, P.W., 2001. Conservation genetics: where are we now? *TREE* 16, 629–636.
- Hedrick, P.W., 2005. A standardized genetic differentiation measure. *Evolution* 59, 1633–1638.
- Hein, J., Schierup, M.K., Wiuf, C., 2005. *Gene Genealogies, Variation and Evolution*. Oxford University Press, Oxford.
- Hernández-Sánchez, J., Haley, C.S., Woolliams, J.A., 2004. On the prediction of simultaneous inbreeding coefficients at multiple loci. *Genet. Res.* 83, 113–120.
- Hill, W.G., Rasbash, J., 1986. Models of long term artificial selection in finite populations. *Genet. Res.* 48, 41–50.
- Hill, W.G., Zhang, X.-S., 2004. Genetic variation within and among animal populations. In: Simm, G., Villanueva, B., Sinclair, K.D.,

- Townsend, S. (Eds.), Farm Animal Genetic Resources. Nottingham University Press, Nottingham, UK, pp. 67–84.
- Hillel, J., Groenen, M.A.M., Tixier-Boichard, M., Korol, A.B., David, L., Kirzhner, V.M., Burke, T., Barre-Dirie, A., Crooijmans, R.P.M.A., Elo, K., Feldman, M.W., Freidlin, P.J., Mäki-Tanila, A., Oortwijn, M., Thomson, P., Vignal, A., Wimmers, K., Weigend, S., 2003. Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genet. Sel. Evol.* 35, 533–557.
- Hudson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267.
- Kalinowski, S.T., 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv. Genet.* 5, 539–543.
- Kohn, M.H., Murphy, W.J., Ostrander, E.A., Wayne, R.K., 2006. Genomics and conservation genetics. *TREE* 21, 629–637.
- Kumar, P., Freeman, A.R., Loftus, R.T., Gaillard, C., Fuller, D.Q., Bradley, D.G., 2003. Admixture analysis of South Asian cattle. *Heredity* 91, 43–50.
- Laval, G., SanCristobal, M., Chevalet, C., 2002. Measuring genetic distances between breeds: use of some distances in various short term evolution models. *Genet. Sel. Evol.* 34, 481–507.
- Laval, G., Iannuccelli, N., Legault, C., Milan, D., Groenen, M.A.M., Giuffra, E., Andersson, L., Nissen, P.H., Jørgensen, C.B., Beeckmann, P., Geldermann, H., Foulley, J.-L., Chevalet, C., Ollivier, L., 2000. Genetic diversity of eleven European pig breed. *Genet. Sel. Evol.* 32, 187–203.
- Le Corre, V., Kremer, A., 2003. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* 164, 1205–1219.
- Leinonen, T., O'Hara, B., Cano, J.M., Merilä, J., 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* 21, 1–17.
- Lewontin, R.C., Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selected neutrality of polymorphism. *Genetics* 74, 175–195.
- Li, M.H., Adamowicz, T., Switonski, M., Ammosov, I., Ivanova, Z., Kiselyova, T., Popov, R., Kantanen, J., 2006. Analysis of population differentiation in North Eurasian cattle (*Bos taurus*) using single nucleotide polymorphism in three genes associated with production traits. *Anim. Genet.* 37, 390–392.
- Li, M.H., Nogovitsina, E., Ivanova, Z., Erhardt, G., Vilkkilä, J., Popov, R., Ammosov, I., Kiselyova, T., Kantanen, J., 2005. Genetic contribution of indigenous Yakutian cattle to two hybrid populations, revealed by microsatellite variation. *Asian-Australas. J. Anim. Sci.* 18, 613–619.
- Lindgren, G., Backström, N., Swinburne, J., Hellborg, L., Einarsson, A., Sandberg, K., Cothran, G., Vilà, C., Binns, M., Ellegren, H., 2004. Limited number of patrines in horse domestication. *Nat. Genet.* 36, 335–336.
- López-Fanjul, C., Fernández, A., Toro, M.A., 2003. The effect of non-additive gene action on the neutral quantitative index of population divergence. *Genetics* 164, 1627–1633.
- Lynch, M., 1996. A quantitative-genetic perspective on conservation issues. In: Avise, J., Hamrick, J. (Eds.), *Conservation Genetics: Case Histories From Nature*. Chapman & Hall, New York, pp. 471–501.
- Meirmans, P.G., 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60, 2399–2402.
- McKay, J.K., Latta, R.G., 2002. Adaptive population divergence: markers, QTL and traits. *TREE* 17, 285–291.
- Malécot, G., 1948. *Les mathématiques de l'hérédité*. Masson et Cie, Paris, France.
- Marletta, D., Tupac-Yupanqui, I., Bordonaro, S., García, D., Guastella, A.M., Criscione, A., Cañón, J., Dunner, S., 2006. Analysis of genetic diversity and the determination of relationships among western Mediterranean horse breeds using microsatellite markers. *J. Anim. Breed. Genet.* 123, 315–325.
- Merilä, J., Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* 14, 892–903.
- Meuwissen, T.H.E., 2007. Operation of conservation schemes. In: Oldenbroek, K. (Ed.), *Utilisation and Conservation of Farm Animal Genetic Resources*. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 167–193.
- Moser, R.J., Reverter, A., Kerr, C.A., Beh, K.J., Lehnert, S.A., 2004. A mixed-model approach for the analysis of cDNA microarray gene expression data from extreme-performing pigs after infection with *Actinobacillus pleuropneumoniae*. *J. Anim. Sci.* 82, 1261–1271.
- Nei, M., 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U. S. A.* 70, 3321–3323.
- Nei, M., 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- Notter, D.R., 1999. The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci.* 77, 61–69.
- Oliehoek, P.A., Windig, J.J., van Arendonk, J.A.M., Bijma, P., 2006. Estimating relatedness between individuals in general populations with a focus on their use in conservation programs. *Genetics* 173, 483–496.
- Ollivier, L., Foulley, J.L., 2005. Aggregate diversity: new approach combining within- and between-breed genetic diversity. *Liv. Prod. Sci.* 95, 247–254.
- Ollivier, L., Alderson, L., Gandini, G.C., Foulley, J.-L., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A.M., Amigues, Y., Boscher, M.-Y., Russell, G., Law, A., Davoli, R., Russo, V., Matassino, D., Désautés, C., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, A.M., Glodek, P., Meyer, J.-N., Plastow, G.S., Siggers, K.W., Archibald, A.L., Milan, D., San Cristobal, M., Laval, M.G., Hammond, K., Cardellino, R., Chevalet, C., 2005. An assessment of European pig diversity using molecular markers: Partitioning of diversity among breeds. *Conserv. Genet.* 6, 729–741.
- Pariset, L., Cappuccio, I., Joost, S., D'Andrea, M., Marletta, D., Ajmone Marsan, P., Valentini, A., The ECONOGENE Consortium, 2006. Characterization of single nucleotide polymorphisms in sheep and their variation as evidence of selection. *Anim. Genet.* 37, 290–292.
- Pearse, D.E., Crandall, K.A., 2004. Beyond FST: analysis of population genetic data for conservation. *Conserv. Genet.* 5, 585–602.
- Pemberton, J., 2004. Measuring inbreeding depression in the wild: the old ways are the best. *TREE* 19, 613–615.
- Peter, C., Bruford, M., Perez, T., Dalamitra, S., Hewitt, G., Erhardt, G., the ECONOGEN Consortium, 2007. Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Anim. Genet.* 38, 37–44.
- Petit, R.J., El Mousadik, A., Pons, O., 1998. Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* 12, 844–855.
- Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., Estoup, A., 2004. GENECLASS 2: a software for genetic assignment and first-generation migrant detection. *J. Heredity* 95, 536–539.
- Piyasatian, N., Kinghorn, B.P., 2003. Balancing genetic diversity, genetic gain and population viability in conservation programmes. *J. Anim. Breed. Genet.* 120, 137–149.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Reed, D.H., Frankham, R., 2001. How closely related are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55, 1095–1103.

- Rege, J.E.O., Gibson, J.P., 2003. Animal genetic resources and economic development: issues in relation to economic valuation. *Ecol. Econ.* 45, 319–330.
- Reist-Marti, S.B., Simianer, H., Gibson, J., Hanotte, O., Rege, J.E.O., 2003. Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. *Conserv. Biol.* 17, 1299–1311.
- Ritland, K., 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. *Mol. Ecol.* 9, 1195–1204.
- Rodríguez-Ramilo, S.T., Toro, M.A., Caballero, A., Fernández, J., 2007. The accuracy of a heritability estimator using molecular information. *Conserv. Genet.* 8, 1189–1198.
- Rosenberg, N.A., Burke, T., Elo, K., Feldman, M.W., Freidlin, P.J., Groenen, M.A.M., Hillel, J., Mäki-Tanila, A., Tixier-Boichard, M., Vignal, A., Wimmers, K., Weigend, S., 2001. Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* 159, 699–713.
- SanCristobal, M., Chevalet, C., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groen, M.A.M., Amigues, Y., Boscher, M.-Y., Russell, G., Law, A., Davoli, R., Russo, V., Désautès, Alderson, L., Finland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martínez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G.C., Matassino, D., Plastow, G.S., Siggens, K.W., Laval, G., Archibald, A.L., Milan, D., Hammond, K., Cardellino, R., 2006. Genetic diversity within and between European pig breeds using microsatellite markers. *Anim. Genet.* 37, 189–198.
- Saura, M., Pérez-Figueroa, A., Fernández, J., Toro, M.A., Caballero, A., in press. Preserving population allele frequencies in ex situ conservation programs. *Conserv. Biol.*, accepted.
- Simianer, H., 2005a. Using expected allele number as objective function to design between and within breed conservation of farm animal biodiversity. *J. Anim. Breed. Genet.* 122, 177–187.
- Simianer, H., 2005b. Decision making in livestock conservation. *Ecol. Econ.* 53, 559–572.
- Simianer, H., Marti, S.B., Gibson, J., Hanotte, O., Rege, J.E.O., 2003. An approach to optimal allocation of conservation funds to minimize loss of genetic diversity between livestock breeds. *Ecol. Econ.* 45, 377–302.
- Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E., McEwan, J.C., 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* 93, 255–265.
- Sonesson, A.K., Meuwissen, T.H.E., 2000. Mating schemes for optimum contribution selection with constrained rates of inbreeding. *Genet. Sel. Evol.* 32, 231–248.
- Spitze, K., 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135, 367–374.
- Tapio, M., Tapio, I., Grislis, Z., Holm, L.-E., Jeppsson, S., Kantanen, J., Miceikiene, I., Olsaker, I., Viinalass, H., Eythorsdottir, E., 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. *Mol. Ecol.* 14, 3951–3963.
- Tapio, I., Varv, S., Bennewitz, J., Maleviciute, J., Finland, E., Grislis, Z., Meuwissen, T.H.E., Miceikiene, I., Olsaker, I., Viinalass, H., Vilkki, J., Kantanen, J., 2006. Prioritization for conservation of northern European cattle breeds based on analysis of microsatellite data. *Conserv. Biol.* 20, 1768–1769.
- Thaon d'Arnoldi, C., Foulley, J.L., Ollivier, L., 1998. An overview of the Weitzman approach to diversity. *Genet. Sel. Evol.* 30, 149–161.
- Toro, M.A., Caballero, A., 2005. Characterisation and conservation of genetic diversity in subdivided populations. *Philos. Trans. R. Soc. Ser. B* 360, 1367–1378.
- Toro, M.A., Pérez-Enciso, M., 1990. Optimization of selection response under restricted inbreeding. *Genet. Sel. Evol.* 22, 93–107.
- Toro, M.A., Fernández, J., Caballero, A., 2006. Scientific basis for policies in conservation of farm animal genetic resources. *Proc. of the 8th World Congress on Genetics Applied to Livestock Production*, CD-Rom Communication No. 33-05.
- Toro, M.A., Silió, L., Rodríguez, M.C., Rodríguez, J., Fernández, J., 1999. Optimal use of genetic markers in conservation programmes. *Genet. Sel. Evol.* 31, 255–261.
- Toro, M.A., Barragán, C., Óvilo, C., Rodríguez, J., Rodríguez, C., Silió, L., 2002. Estimation of coancestry in Iberian pigs using molecular markers. *Conserv. Genet.* 3, 309–320.
- Van Tassell, C., Wiggans, G., 2007. Large scale bovine SNP genotyping for genomic selection and Hapmap development. *USDA Research project number 1265-31000-081-06*.
- Vasemägi, A., Nilson, J., Primmer, C.R., 2005. Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol. Biol. Evol.* 22, 1067–1076.
- Vekemans, X., Beauwens, T., Lemaire, M., Roldán-Ruiz, I., 1990. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol.* 11, 139–151.
- Villanueva, B., Woolliams, J.A., Simm, G., 1994. Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genet. Sel. Evol.* 26, 517–535.
- Wang, J., 2001. Optimal marker-assisted selection to increase the effective size of small populations. *Genetics* 157, 867–874.
- Weitzman, M.L., 1992. On Diversity. *Q. J. Econ.* 107, 363–405.
- Weitzman, M.L., 1993. What to preserve? An application of diversity theory to crane conservation. *Q. J. Econ.* 108, 157–183.
- Wiener, P., Burton, D., Ajmone-Marsan, P., Dunner, S., Mommens, G., Nijman, I.J., Rodellar, C., Valentini, A., Williams, J.L., 2003. Signatures of selection? Patterns of microsatellite diversity on a chromosome containing a selected locus. *Heredity* 90, 350–358.
- Woolliams, J.A., Toro, M.A., 2007. What is genetic diversity? In: Oldenbroek, K. (Ed.), *Utilisation and Conservation of Farm Animal Genetic Resources*. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 55–77.
- Wong, G.K., Liu, B., Wang, J., Zhang, Y., Yang, X., et al., International Chicken Polymorphism Map Consortium, 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432, 717–722.
- Wright, S., 1969. *Evolution and the genetics of populations. The Theory of Gene Frequencies*, vol. 2. University of Chicago Press, Chicago.