Genetic variation and population structure of Italian native sheep breeds undergoing in situ conservation

C. Dalvit, M. De Marchi, E. Zanetti, and M. Cassandro

Dipartimento di Scienze Animali, Università degli Studi di Padova, Viale dell’Università 16, 35020 Legnaro (PD), Italy

ABSTRACT: The genetic variability and presence of population substructures in 4 native Northern Italian sheep breeds, Alpagota, Brogna, Foza, and Lamon, undergoing in situ conservation, and 1 widespread Italian breed, Bergamasca, were studied by investigating 19 microsatellite markers. The breeds showed considerable genetic variability in terms of number of alleles and heterozygosity, with the exception of Alpagota, which was the least variable (0.607). Nevertheless, a significant deficit of heterozygotes was observed in each breed due to rather increased levels of inbreeding or to the presence of population substructures, probably caused by increased genetic variation in the founder populations. The analyses evidenced clear genetic differentiation ($F_{ST} = 0.085$), reduced levels of admixture, and presence of private alleles among the breeds, confirming their genetic uniqueness. In particular, according to Reynolds genetic distances, Alpagota was the most differentiated, perhaps because it had been bred mostly in a rather isolated area. Loss of any of the investigated breeds would result in a loss of genetic diversity ranging from 0.5 to 1.6% of the total observed gene diversity. Results supported the decision to safeguard these breeds as important reservoirs of genetic diversity and suggested breeding and mating practices to maintain variability and to overcome within-breed substructures.

Key words: genetic variability, in situ conservation, microsatellite, population structure, sheep

INTRODUCTION

In Italy, sheep breeding has always played an important role, mainly in less developed, rural, and mountain areas, where the production systems are interrelated with local traditions and native breeds (Rancourt et al., 2006). In recent decades, the focus on production has remarkably changed sheep management, emphasizing the breeding of a small number of very productive and selected cosmopolitan breeds. This trend has also been observed in cattle breeding and is the consequence of socio-economic changes. In fact, the continuous improvement of crop productivity and the subsequent low price of cereals led farmers to feed animals with grain instead of exploiting pastures and to change the production systems in favor of more productive and highly selected breeds (Mendelsohn, 2003). Therefore, the intensification of production systems and the success of industrial breeding have led farmers to abandon several autochthonous genotypes (Taberlet et al., 2008). In the northeastern part of Italy, the Veneto region, the sheep population decreased drastically from 101,170 animals in 1953 (Montanari, 1954) to 34,734 animals in 1991, and several native breeds disappeared (Pastore and Fabbris, 1999). Nowadays, only 4 native breeds are still reared in Veneto: Alpagota (ALP), Brogna (BRO), Foza (FOZ), and Lamon (LAM). They are mainly found on farms located in mountainous and hilly areas, and due to their good adaptive traits and to their environmental role in exploiting marginal pastures, these breeds represent an important reservoir of genetic resources. The aim of this paper was to study the genetic variability and presence of population substructures in these 4 native breeds using microsatellite markers. In fact, several studies have already indicated the presence of significant inbreeding in many native sheep breeds (Pariset et al., 2003; Rendo et al., 2004; Santos-Silva et al., 2008), suggesting that even domestic breeds belonging to species such cattle, sheep, and goat could be considered endangered (Taberlet et al., 2008).
MATERIALS AND METHODS

Research protocols followed the guidelines stated in the Guide for the Care and Use of Agricultural Animals in Agricultural Research and Teaching (FASS, 1999).

Breed Description

The 4 Venetian native breeds (ALP, BRO, FOZ, and LAM) analyzed in the present paper are historically tied to mountain communities of the Veneto Italian region and have always been used for the production of traditional dairy or meat products. The following is a brief description of their main features.

The ALP is a small-sized sheep that originated in the Alpago mountains, and nowadays about 1,500 animals are enrolled in the herd book; it is listed in the FAO global databank for farm animal genetic resources where it is classified as endangered. It is a multipurpose breed (meat, milk, and wool), but now it is bred only for meat production due to its purported high quality. For these reasons, ALP lamb is 1 of the 270 presidia in the world recognized by the Slow Food Foundation (2008), and its meat is sold to the best restaurants of the Veneto region. According to Di Stasio (1980), the good adaptive traits to the mountain environment exhibited by ALP are due to the greater frequency of blood group A with respect to other sheep breeds in which the B group is more frequent.

The BRO breed has been known since the 12th century and is native to the Lessini mountains, although its origin are uncertain; at present the herd book accounts for about 1,300 animals. It is not present in the FAO databank, but according to the population size, it can be classified as endangered. The BRO is a multipurpose breed (meat, milk, and wool), and its milk has been traditionally transformed to produce typical cheeses. The animals are medium to small in size and are very prolific, with a twinning rate of 58% as reported by Pastore and Fabbri (1999).

The first evidence of the existence of the LAM breed traces to the 19th century; it was described as a rustic breed with low mortality rates and was the primary economic resource of local farmers (Pastore, 2002). Today, about 300 animals are enrolled in the herd book, and according to the FAO global databank for farm animal genetic resources, it is classified as endangered. The LAM sheep are well adapted to adverse conditions and are maintained in a semi-wild state without the use of any housing throughout the year. Its traditional management provided for flock migration from hill to plain in winter, from plain to hill in spring, and from hill to mountains in summer (Pastore, 2002). It is a multipurpose breed (meat, milk, and wool), but nowadays is kept only for meat, in particular for the production of a typical smoked meat.

The FOZ breed origins are uncertain, but it is native to the Asiago Mountains; nowadays the situation of FOZ is very critical because its population size is estimated at only about 100 animals. The FOZ breed was traditionally reared for the good quality of its wool and was typically reared in small herds either transhumance or farm flock (Pastore, 2002).

Bergamasca (BER) is a native breed from the Lombardy region, which is widespread mainly in northern Italy. This breed specializes in meat production and is widely used also in crossing with other breeds. In the present work, it was used as a reference population.

Animal Sampling

A total of 170 individual blood samples were collected from the 5 sheep breeds involved in the study: ALP (n = 38), BRO (n = 38), FOZ (n = 33), LAM (n = 16), and BER (n = 45). Samples from ALP, BRO, FOZ, and LAM were collected on the experimental farm where they are maintained according to an in situ conservation program. All of the animals present in the conservation nucleus were sampled to estimate the degree of variability and the population structure of the conserved animals at the beginning of the conservation activities. The conservation scheme provided for the choice of founder animals according to the breed standards; founders belonged to different farms and were moved to the conservation nucleus on the experimental farm.

Blood samples of BER were collected in several flocks and were used as a reference population. Blood samples were collected from each animal in 5-mL vacutainer tubes containing sodium citrate as anticoagulant and stored at −20°C until analyses were performed. The DNA extraction was carried out employing the Gentra System PUREGENE DNA purification kit (Gentra System, Minneapolis, MN) starting from 300 µL of whole blood.

Amplification and Genotyping of Microsatellite Markers

A panel of 19 microsatellite markers was established according to ISAG/FAO Standing Committee (2004) recommendations and to previous studies (Baumung et al., 2006) to investigate highly polymorphic markers spread throughout the genome. The loci studied were OarAEB45, OarFCB20, URB58, McM527, INRA23, TGLA53, MAF65, OarCP49, MAF214, HSC, INRA63, McM42, OarAE119, OarAE129, ILSTS087, OarFCB304, OarCP34, OarCP20, and CSD247 (Table 1). Details on the amplification protocol and primer annealing temperatures can be found in Dalvit et al. (2008). Allele size was determined with a CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA).

Statistical Analysis

Number of alleles per locus, allelic frequencies, and observed and expected heterozygosity were calculated
using Genetix version 4.05.2 (Belkhir et al., 1996–2004). Exact tests for deviation from Hardy-Weinberg equilibrium (Guo and Thompson, 1992) were applied using a Markov Chain Monte Carlo simulation (100 batches, 5,000 iterations per batch, and a dememorization number of 10,000) as implemented in GENEPOP version 3.4 (Raymond and Rousset, 1995). A test for population differentiation was performed, as implemented in GENEPOP 3.4; for each locus, an unbiased estimate of Fisher’s exact test was computed to verify if the allelic distribution was different among breeds. The Fstat 2.9.3 software (Goudet, 1995) was employed in calculations of allelic richness (an estimation of mean number of alleles per locus corrected by sample size), gene diversity (Nei, 1987), and estimation of Wright’s fixation index (Weir and Cockerham, 1984). Molecular coancestry coefficients within and between breeds were measured according to Caballero and Toro (2002) using MolKin 3.0 (Gutiérrez et al., 2005); to avoid bias, because of unequal sample sizes, 100 samples with exactly 50 individuals per breed were generated with a bootstrap procedure. To help set conservation priorities, Molkin 3.0 (Gutiérrez et al., 2005) was used to quantify the contribution of each analyzed population to the diversity of the whole data set using the method proposed by Caballero and Toro (2002). Because BER is not a Veneto native breed involved in the conservation scheme but has been used as reference population, its data were not included in the approach for setting conservation priorities. The genetic distances of Reynolds (\(D_R\)) among breeds were estimated, a neighbor-joining consensus tree was reconstructed, and tree robustness was evaluated by bootstrapping over loci (1,000 replicates) using the PHYLIP package (Felsenstein, 1993–2002); the dendrogram was depicted using the software package TreeView version 1.6.6 (Page, 2001). Reynolds distance is the most suited for relatively closely related populations such as breeds in Europe that diverged during short time periods; in fact, in this case, the amount of mutation is negligible and the main factor to describe genetic variability is random drift (Eding and Laval, 1999). Moreover, genetic distances among individuals were estimated as the proportion of shared alleles (\(D_{PS}\)) using MICROSAT 1.5 (Minch et al., 1998); individual distances were represented by a neighbor-joining tree and displayed using MEGA 4 (Tamura et al., 2007). To study population structure and to detect the most likely number of clusters (\(K\)) in the data set, the software STRUCTURE version 2.2 (Pritchard et al., 2000) was used. The analysis involved an admixture model with correlated allele frequencies as suggested by several authors (Pritchard et al., 2000; Vicente et al., 2008; Zuccaro et al., 2008). To choose the appropriate number of inferred clusters to model the data, 2 to 10 inferred clusters were performed with 50 independent runs each. All analyses used a burn-in period of 30,000 iterations and then 150,000 iterations for data collection. The best number of clusters fitting the data was established by plotting the Ln Pr(X|K) over the 50 independent runs for each K, as suggested by Pritchard et al. (2000). To find optimal alignments of independent runs, the computer program CLUMPP 1.1 (Jakobsson and Rosenberg, 2007) was used; the output obtained was used directly as input by the cluster visualization program DISTRUCT (Rosenberg, 2004).

### Table 1. Microsatellite markers with corresponding fragment size, chromosomal location, total number of detected alleles (TNA), allelic richness (AR), and gene diversity (GD)

| Locus   | Fragment size | Chromosome | TNA | AR  | GD  |
|---------|---------------|------------|-----|-----|-----|
| CSRD274 | 213–259       | 14         | 22  | 8.0 | 0.779 |
| ILSTS87 | 138–178       | 6          | 19  | 8.8 | 0.814 |
| OarCP20 | 68–102        | 21         | 15  | 5.5 | 0.745 |
| OarCP34 | 100–128       | 3          | 12  | 6.5 | 0.788 |
| OarFCB304 | 150–198     | 19         | 20  | 7.3 | 0.702 |
| McM527  | 171–189       | 5          | 10  | 7.0 | 0.738 |
| OarAE54 | 122–156       | 25         | 14  | 8.2 | 0.797 |
| OarFCB20| 92–122        | 2          | 16  | 8.9 | 0.836 |
| URB58   | 161–209       | 13         | 18  | 8.0 | 0.808 |
| OarAE129| 137–157       | 5          | 8   | 4.7 | 0.651 |
| OarAE119| 147–185       | 19         | 13  | 7.7 | 0.770 |
| INRA63  | 162–210       | 14         | 21  | 7.4 | 0.785 |
| HSC     | 263–299       | 20         | 19  | 9.3 | 0.835 |
| McM42   | 82–106        | 9          | 9   | 5.1 | 0.634 |
| INRA23  | 196–224       | 1          | 15  | 9.0 | 0.834 |
| MAF214  | 185–261       | 16         | 13  | 4.4 | 0.538 |
| MAF65   | 121–143       | 15         | 12  | 7.5 | 0.821 |
| OarCP49 | 69–119        | 17         | 22  | 8.1 | 0.793 |
| TGLA53  | 140–168       | 12         | 15  | 7.9 | 0.846 |
| Mean    |               |            | 15.4| 7.3 | 0.764 |
Table 2. Number of analyzed samples, total number of alleles (TNA), allelic richness (AR), expected (H. exp.) and observed (H. obs.) heterozygosity, within-breed heterozygote deficiency (FIS), and within-breed molecular coancestry (fij) for each breed analyzed

| Breed      | Sample size | TNA | AR | H. exp. ± SD | H. obs. ± SD | FIS    | fij ± SD |
|------------|-------------|-----|----|-------------|-------------|--------|---------|
| Bergamasca | 45          | 9.6 | 6.3| 0.782 ± 0.090 | 0.722 ± 0.149 | 0.078*** | 0.236 ± 0.008 |
| Alpagota   | 38          | 7.5 | 5.3| 0.699 ± 0.187 | 0.607 ± 0.265 | 0.134*** | 0.318 ± 0.011 |
| Broga      | 38          | 9.3 | 6.1| 0.764 ± 0.087 | 0.675 ± 0.164 | 0.118*** | 0.246 ± 0.012 |
| Foza       | 33          | 9.1 | 6.2| 0.770 ± 0.845 | 0.698 ± 0.186 | 0.095*** | 0.250 ± 0.011 |
| Lamon      | 16          | 7.1 | 6.2| 0.796 ± 0.063 | 0.736 ± 0.230 | 0.078*** | 0.239 ± 0.008 |

***P < 0.001.

RESULTS

Genetic Variability at Microsatellite Loci

The total number of alleles detected in the entire data set was 293, and all markers were polymorphic in each breed. Information about the variability of the investigated loci is shown in Table 1. The largest number of alleles was found at loci CSRD274 and OarCP49 (22) and the smallest at locus OarAE129 (8). Global mean number of alleles per locus was 15.4, and allelic richness, an estimate of number of alleles per locus corrected by sample size, was 7.3. The gene diversity across loci was 0.764, ranging from 0.846 (TGLA53) to 0.651 (OarAE129). Private alleles were detected in each breed and for each marker; they represented 30% of the total number of observed alleles. However, allele frequencies were small and exceeded 0.10 in few cases. Bergamasca showed the largest number of private alleles with increased frequencies, displaying 4 alleles with frequencies greater than 0.10; ALP and BRO exhibited 2 private alleles each with frequencies greater than 0.15, whereas FOZ and LAM displayed private alleles with rather small frequencies.

Breed Variability and Differentiation

The genetic variability of each breed was studied, first, in terms of number of observed alleles and allelic richness, as shown in Table 2. Bergamasca showed the largest number of alleles per locus (9.6), followed by BRO (9.3), and LAM the least (7.1). On the other hand, allelic richness was very similar among breeds, with the exception of ALP, which exhibited the least value (5.3). Bergamasca also exhibited considerable heterozygosity (0.722), whereas in the native breeds, it was generally less, except for LAM (0.736). In general, all breeds showed heterozygosity estimates less than expected; a highly significant (P < 0.001) departure from Hardy-Weinberg equilibrium was observed in each of the breeds. This divergence was reflected in the FIS index within breed, which was always positive and rather large, especially in ALP and BRO, which are the least variable breeds according to observed heterozygosity.

Another way to measure within-breed diversity is the estimation of molecular coancestry, a measure of relatedness among individuals. Molecular coancestry estimates varied from 0.236 (BER) to 0.318 (ALP), as shown in Table 2.

The breeds showed considerable genetic differentiation; in fact, the FST index was equal to 0.085 (0.066 to 0.107; 99% confidence interval). Moreover, the test for population differentiation revealed a highly significant (P < 0.001) difference in the distribution of allelic frequencies among the breeds. To test if the large FST estimate was due to the presence of BER, which is not a native Venetian breed, it was computed again after the removal of this breed; the estimate was even greater as FST was equal to 0.097 (0.070 to 0.128, 99% confidence interval). The neighbor-joining tree constructed on D_R (Figure 1) showed a clear distinction between ALP and BER and the other 3 local breeds, which appeared to be more similar.

Results obtained with the Caballero and Toro (2002) approach to set up conservation priorities are illustrated in Table 3. The removal of 1 breed from the data set resulted in a loss of genetic diversity in the population, which ranged from −0.51 to −1.64% when ALP and FOZ, respectively, were removed. The greatest loss of between-breed diversity was found by removing ALP (−3.25%); on the other hand, its removal resulted in a contribution to the within-breed diversity (0.27%). On the contrary, removal of BRO gave a modest contribution to the between-breed diversity (0.08%) and a loss in the within-breed diversity (−1.13%).

Population Structure

The structure of the breeds was analyzed using a Bayesian approach that inferred the number of clusters (K) present in the population, permitting detection of differences among breeds and hidden substructures within breeds. The Ln Pr(X|K) increased sharply from K = 2 to K = 5 and reached a plateau without showing a significant decrease from K = 6 to K = 10, as illustrated in Figure 2. The greatest mean Ln Pr(X|K) over the 50 independent runs performed for each value of K was found at K = 7; at K = 5 and K = 6, estimated Ln Pr(X|K) were more consistent across runs leading to some difficulties in identifying the most probable number of clusters in the population. Results of STRUCTURE analyses are shown in Figure 3 for K ranging from 5 to 8. It can be noticed that, apart from the number of clusters, a clear dis-
tinction among the 5 breeds was possible and only a small level of admixture was detectable. In particular, ALP and BER formed 2 distinct clusters with high proportion of membership (0.876 and 0.890, respectively, considering K = 7), and admixture was nearly absent. On the other hand, BRO, FOZ, and LAM exhibited a more complex structure with proportion of membership split into 2 or more clusters and an underlying presence of within breed substructures; nevertheless, admixture seemed to be negligible even among these breeds. It is worth mentioning that, even at K > 8, ALP and BER still evidenced increased membership to only 1 cluster, whereas BRO, FOZ, and LAM presented a more and more complex structure.

The neighbor-joining tree of individuals, based on the proportion of shared alleles (Figure 4), showed that individuals grouped together very well by breed. Some confusion was present only between FOZ and LAM animals, which, in some cases, grouped together. It is interesting to see that the tree could be separated into 2 groups, one composed of BRO, FOZ, and LAM groups and one composed of ALP and BER.

**DISCUSSION**

The 4 local breeds analyzed in this study are the only native Venetian breeds that survived extinction and are still reared on small farms. Nevertheless, to maintain them and to increase their population size, it has been necessary to set up a conservation flock on an experimental farm with financing support of the Veneto regional government. Reasons for conserving animal genetic resources have already been underlined by several authors (Mendelsohn, 2003; Rege and Gibson, 2003), and in our case, the conservation of ALP, BRO, FOZ, and LAM is carried out for 3 main reasons. 1) Their important environmental role in exploiting mountain and marginal pastures. In fact, the abandonment of pastures results in a landscape change in favor of the expansion of woodlands, as observed by Giupponi et al. (2006) in a study of the land use changes in the Belluno Venetian province. This loss of landscape diversity is negative for the aesthetic-recreational appeal of the area because it has been shown that tourism activities favor maintaining a more pleasant landscape and avoiding excessive afforestation (Giupponi et al., 2006). This question is of outstanding importance as tourism plays an important role in the economy of the Veneto region. In addition, landscape changes also have negative effects on the habitats of species of naturalistic interests (Giupponi et al., 2006). 2) The cultural value in that these breeds are used for the production of typi-

| Breed     | GD | Within breeds | Between breeds | Loss (−)/Gain (+) |
|-----------|----|---------------|----------------|------------------|
| All breeds| 0.813 |        |                |                  |
| Alpagota  | 0.806 | +2.740 | −3.245 | −0.505 |
| Brogna    | 0.801 | −1.126 | +0.081 | −1.045 |
| Foza      | 0.796 | −0.889 | −0.752 | −1.641 |
| Lamon     | 0.801 | −0.538 | −0.531 | −1.069 |
cal foods linked to old traditions; authors have usually referred to this aspect as existence value (Rege and Gibson, 2003). 3) The interesting adaptive traits representing genetic stocks that could transmit favorable traits to new breed crosses. Thanks to these features, these breeds have been enrolled in an in situ conservation program aiming to increase the population size, to maintain genetic variability, to organize matings, and to avoid crossbreeding, which was rather common in small flocks. In fact, in situ preservation of endangered breeds should be preferred to ex situ or cryo-conservation because in situ preservation promotes sustainable management of genetic resources and permits conservation of all features carried by a specific breed (Taberlet et al., 2008). Moreover, in situ conservation allows animals to be kept in the original area of production. This aspect is important considering the link between Venetian sheep breeds and their typical production. Product valorization is also a step toward the sustainability of these breeds in case Italian or European government subsidies are not available in the future. An example of the profitability of rearing local and less productive breeds in marginal areas, when there is a strong link among breed, territory, and product, has been examined by Pretto et al. (2009) in a study of the native Italian Burlina cattle breed. The first step toward safeguarding of ALP, BRO, LAM, and FOZ was the choice of the founder animals to rear in the conservation nucleus, according to their phenotypic traits. It is worth mentioning that, for the FOZ and LAM breeds, the situation was extremely critical because their population size is extremely small. Nevertheless, these were the 2 breeds showing the greatest variability and the least, but significant, deficit of heterozygotes. This lack of heterozygotes could be caused by the presence of popu-

Figure 3. Graphical representation of the estimated membership fractions of individuals of the breeds analyzed in each of the K inferred clusters, for K = 6 to K = 8. Color version available in the online pdf.

Figure 4. Neighbor-joining tree based on the proportion of shared alleles of Alpagota (ALP = ●), Brogna (BRO = ■), Foza (FOZ = ▲), Lamon (LAM = ∆), and Bergamasca (BER = ○).
Genetic variation in native Italian sheep breeds

In conclusion, the present study highlighted the importance of the in situ conservation program set up to safeguard native sheep breeds. In fact, these breeds are an important reservoir of genetic variation and of genetic uniqueness, and their extinction would cause a loss of diversity. Our study will be useful in outlining further development of the in situ conservation scheme; particular attention should be paid to organizing matings to minimize inbreeding, especially in ALP, which is already suffering from a lack of heterozygosity. The presence of substructures observed in BRO, FOZ, and LAM was probably due to fragmentation in isolated small flocks before their inclusion in the conservation flock. This population structure should be overcome in the next generations by rotating the rams and favoring mating between genetically distinct animals. In addition, to achieve sustainable management of the described genetic resources, the production of these breeds should be valorized by taking into account their value as reservoirs of unique diversity, as suggested by Taberlet et al. (2008).

**LITERATURE CITED**

Álvarez, I., J. P. Gutiérrez, L. J. Royo, I. Fernández, E. Gómez, J. J. Arranz, and F. Goyache. 2005. Testing the usefulness of the molecular coancestry information to assess genetic relationships in livestock using a set of Spanish sheep breeds. J. Anim. Sci. 83:737–744.

Baumung, R., V. Cubric-Curik, K. Schwend, R. Achmann, and J. Sölkner. 2006. Genetic characterization and breed assignment in Austrian sheep breeds using microsatellite markers information. J. Anim. Breed. Genet. 123:265–271.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 1996–2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions. CNRS UMR 5000, Université de Montpellier II, Montpellier, France. http://www.genetix.univ-montp2.fr/genetix/genetix.htm Accessed Nov. 5, 2008.

Caballero, A., and M. A. Toro. 2002. Analysis of genetic diversity for the management of conserved subdivided populations. Conserv. Genet. 3:289–299.
