Dipartimento di Produzioni Animali, Università della Tuscia, via S. Camillo de Lellis, Viterbo, Italy

Use of microsatellites for genetic variation and inbreeding analysis in Sarda sheep flocks of central Italy

By L. Pariset, M. C. Savarese, I. Cappuccio and A. Valentini

Summary
A set of 11 polymorphic microsatellites has been used to assess the distribution of genetic variability in 17 flocks of the Sarda sheep breed in the Viterbo province, located in central Italy. The suitability of samples size and number of loci analysed were tested using a bootstrap procedure. The data obtained were used to estimate the genetic diversity within and among flocks, and to test the presence of inbreeding within flocks. To study the genetic relationship among flocks, a principal component analysis based on Nei standard distances was performed. The results of genetic analysis show a significant excess of homozygotes in some of the flocks. The use of genetic data for maximum likelihood tests of assignment is investigated in order to establish suitable breeding policies. This research represents a case study for the development of mating strategies suitable for incrementing genetic variability within flocks that can be applied in other domestic species showing similar breeding problems (goat, buffalo and small populations of cattle).

Zusammenfassung

Introduction
In central Italy, sheep breeding is an important livestock sector, as a source of milk and meat. In the last 40 years breeding techniques and genotypes chosen for breeding have sensibly changed in this area, because of the extensive use of highly specialized breeds (de Haan et al. 1997). When a limited number of parents are employed, an increase of inbreeding in the populations is likely to occur (Meszaros et al. 1998) because a founder effect in most of the flocks living in the area may take place, reducing the number of possible genotypes. In this case, the wrong choice of mating animals can determine the fixation of unfavourable alleles in the populations, such as the Welsh allele of αS1-casein (Crisà et al. 2001). Moreover, a high level of inbreeding together with a reduced gene flow may lead to genetic
fragility of populations that are unfit to environmental changes (Bijlsma et al. 1997). Compared with other livestock species, information on ovine genetics are still relatively scarce, although recently more attention is being paid to this taxon (Hedrick 1994; Maddox et al. 2001; Reed and Beattie 2001; EU funded projects Biotech2 and Econogene).

Different molecular markers have been used for genetic analysis in sheep: DNA fingerprints (Laneluc et al. 1992), random amplified polymorphic DNA (RAPDs) (Kantanen et al. 1995; Gutierrez-Adan et al. 1997; Matthews and Crawford 1998), restriction fragment length polymorphisms (RFLPs) (Parsons et al. 1996), microsatellites (Moore et al. 1991; Paterson 1998; Diez-Tascón et al. 2000; Shay et al. 2001). In this study we have used microsatellites because they are a powerful tool for the evaluation of paternity, to maintain pedigree records (Marklund et al. 1994; Glowatzki-Mullis et al. 1995; Talbot et al. 1996), for tracking alleles through a population (Moazami-Goudarzi et al. 1997; Arranz et al. 1998; Wimmers et al. 2000) and to estimate genetic variability and inbreeding (Zajc et al. 1997; Rooney et al. 1999; Cañon et al. 2000; Hedrick et al. 2001; Goudet and Keller 2002).

The aim of this study is the use of molecular data for evaluating genetic variability, gene flow and inbreeding in sheep flocks, and for designing a breeding strategy aimed at incrementing genetic diversity within and across flocks.

## Materials and methods

### Samples and DNA extraction

Blood samples of 376 individuals, all females, were randomly collected in 17 farms (Table 1) located in the Viterbo area (northern Latium, central Italy) with each farm having at least 300 individuals. DNA was extracted from 0.1 ml of fresh or frozen blood using the Wizard Genomic DNA Extraction Kit (Promega, Madison, WI) according to manufacturer’s instructions.

### Table 1. Location of the farms analysed, number of animals sampled (n), microsatellite expected (He) and observed (H0) heterozygosity, mean number of alleles (m.n.a.), FIS values per flock with relative standard deviation (SD) estimated on 1000 bootstrap replicates, and percentage of correct assignment of individual genotypes to the flock of origin by maximum likelihood (m.l.c.a)

<table>
<thead>
<tr>
<th>Flock</th>
<th>Locality</th>
<th>n</th>
<th>He ± SD</th>
<th>Ho ± SD</th>
<th>m.n.a. ± SD</th>
<th>FIS ± SD</th>
<th>p &lt; 0.05</th>
<th>m.l.c.a</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Acquapendente</td>
<td>22</td>
<td>0.81 ± 0.03</td>
<td>0.70 ± 0.04</td>
<td>7.63 ± 2.62</td>
<td>0.14 ± 0.06</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>CM</td>
<td>Valentano</td>
<td>18</td>
<td>0.68 ± 0.00</td>
<td>0.67 ± 0.08</td>
<td>6.00 ± 1.41</td>
<td>0.01 ± 0.12</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>CPL</td>
<td>Ischia di Castro</td>
<td>25</td>
<td>0.82 ± 0.08</td>
<td>0.48 ± 0.08</td>
<td>10.50 ± 0.71</td>
<td>0.41 ± 0.09</td>
<td>**56%</td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td>Collevalle (Bomarzo)</td>
<td>29</td>
<td>0.75 ± 0.05</td>
<td>0.64 ± 0.04</td>
<td>7.89 ± 3.41</td>
<td>0.14 ± 0.08</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td>ET</td>
<td>Ischia di Castro</td>
<td>21</td>
<td>0.66 ± 0.10</td>
<td>0.51 ± 0.05</td>
<td>6.25 ± 3.28</td>
<td>0.22 ± 0.09</td>
<td>**91%</td>
<td></td>
</tr>
<tr>
<td>FT</td>
<td>Ischia di Castro</td>
<td>24</td>
<td>0.80 ± 0.02</td>
<td>0.67 ± 0.04</td>
<td>7.50 ± 1.77</td>
<td>0.16 ± 0.06</td>
<td>75%</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Gallesè</td>
<td>20</td>
<td>0.71 ± 0.03</td>
<td>0.50 ± 0.05</td>
<td>5.86 ± 1.57</td>
<td>0.30 ± 0.09</td>
<td>**90%</td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>Ischia di Castro</td>
<td>25</td>
<td>0.80 ± 0.03</td>
<td>0.60 ± 0.05</td>
<td>7.50 ± 2.81</td>
<td>0.25 ± 0.05</td>
<td>***68%</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>Ischia di Castro</td>
<td>20</td>
<td>0.80 ± 0.02</td>
<td>0.65 ± 0.04</td>
<td>7.83 ± 2.64</td>
<td>0.19 ± 0.09</td>
<td>73%</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>Ischia di Castro</td>
<td>21</td>
<td>0.76 ± 0.05</td>
<td>0.44 ± 0.08</td>
<td>6.50 ± 0.71</td>
<td>0.42 ± 0.12</td>
<td>**84%</td>
<td></td>
</tr>
<tr>
<td>PB</td>
<td>Grotte S. Stefano</td>
<td>35</td>
<td>0.75 ± 0.04</td>
<td>0.60 ± 0.07</td>
<td>6.50 ± 3.42</td>
<td>0.20 ± 0.11</td>
<td>76%</td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Castiglione in teverina</td>
<td>17</td>
<td>0.75 ± 0.07</td>
<td>0.51 ± 0.05</td>
<td>8.67 ± 2.16</td>
<td>0.32 ± 0.07</td>
<td>***94%</td>
<td></td>
</tr>
<tr>
<td>PO</td>
<td>Valentano</td>
<td>20</td>
<td>0.74 ± 0.02</td>
<td>0.56 ± 0.09</td>
<td>6.00 ± 0.00</td>
<td>0.24 ± 0.13</td>
<td>88%</td>
<td></td>
</tr>
<tr>
<td>RA</td>
<td>Valentano</td>
<td>24</td>
<td>0.71 ± 0.10</td>
<td>0.56 ± 0.06</td>
<td>7.50 ± 4.36</td>
<td>0.21 ± 0.06</td>
<td>**64%</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>Acquapendente</td>
<td>23</td>
<td>0.67 ± 0.05</td>
<td>0.76 ± 0.05</td>
<td>5.50 ± 1.73</td>
<td>0.12 ± 0.10</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Valentano</td>
<td>22</td>
<td>0.80 ± 0.06</td>
<td>0.65 ± 0.04</td>
<td>8.57 ± 3.21</td>
<td>0.18 ± 0.05</td>
<td>**83%</td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>Acquapendente</td>
<td>22</td>
<td>0.77 ± 0.10</td>
<td>0.76 ± 0.07</td>
<td>8.50 ± 2.12</td>
<td>0.02 ± 0.09</td>
<td>73%</td>
<td></td>
</tr>
</tbody>
</table>

Mean values over all flocks 0.75 ± 0.05 0.60 ± 0.06 7.34 ± 2.23 0.19 ± 0.01

*p < 0.05 applying the Bonferroni criterion; **p < 0.05 of FIS ≥ 0 without correction for multiple comparisons.
Amplification and screening of microsatellites

The 11 microsatellite markers used in this study were OarFCB0011, OarFCB0304, OarCP0049, OarFCB0048, OLADRB2abcde ILSTS0011, ILSTS0028, ILSTS0005, MAF0214, McM0527 and CSR0247. Microsatellites were amplified in a reaction mix containing 1.25 units of Taq DNA polymerase, 2.5 mm MgCl₂, 200 μM of each dNTP (all from Amersham Pharmacia, Little Chalfont, UK), and 25–100 ng of genomic DNA. PCR reactions were performed using near-infrared (NIR) fluorescent dye (iridine 700 or 800) end-labelled primers (MWG Biotech, Ebersberg, Germany) in a MJ Research thermocycler (MS Research, Waltham, MA). Amplification products were electrophoresed on a 6.5% KBPlus gel (Licor Inc., Lincoln, NE) with 0.25 mm spacers under the following conditions: 1500 V, 40 W, 40 mA and 45°C. Alleles were visualized and digitalized on a Licor 4200 DNA sequencer using the e-Scan Global Controller software, and true image data were scored and analysed using Gene ImagIR software Fragment Analysis (both software from LICOR Inc., Lincoln, NE, USA). The length of the amplification products was calculated by the 50–350 IRD-700 and 800 size standards (MWG Biotech) and by three reference samples included in each gel.

Statistical analysis

Number of alleles, expected heterozygosity (He; Nei 1987), observed heterozygosity (Ho; Hedrick 1983) were estimated. The coefficient of variation of the Nei’s estimator of $F_{ST}$ was computed on 1000 bootstrap replicates over individuals and loci (Almudevar 2001), and used to evaluate the suitability of sample size and number of loci. The principal coordinate analysis (PCA) was performed on Nei standard distances (Takezaki and Nei 1996) by the Jacobi method (Press 1992). In order to assess the presence of inbreeding within flock, the Nei’s estimator of $F_{IS}$ (Nei 1987) was calculated for each flock according to original formulas, a bootstrap analysis on 1000 replicates was performed to estimate the standard deviation of $F_{IS}$ and finally the probability of $F_{IS} \neq 0$ was tested by Student’s $t$-test with and without applying a standard Bonferroni correction.

To individualize the animals more suitable to be included in breeding schemes, an approach based on the computation of the log-likelihood of each individual genotype to belong to each flock (Paetkau et al. 1995, 1997; Waser and Strobeck 1998) was performed using ARLEQUIN version 2000 (Schneider et al. 2000).

Results

All loci were found to be polymorphic in the 17 flocks, and generated a total of 165 alleles across the 11 loci from the 376 individuals analysed, with locus ILSTS0028 showing the highest number of alleles. The number of alleles detected for each marker ranged from 7 to 28. Figure 1 shows the values of the coefficient of variation (CV) of the fixation index ($F_{ST}$) versus a number of individuals ranging from 3 to 21 and a number of loci ranging from 2 to 11. The flex of the CV curves corresponds to 12 individuals and eight microsatellite loci, and represents the point after which very little precision is gained by increasing the size of the experiment.

Table 1 summarizes the population statistics generated by the 11 microsatellites for different flocks. The observed and expected heterozygosity per flock ranged from 0.44 to 0.76 and from 0.66 to 0.82, respectively. The mean expected heterozygosity (He) over all flock was 0.75, ranging from below 0.66 to 0.82 and the mean observed heterozygosity (Ho) over all flock was 0.60, ranging from below 0.44 to 0.76. $F_{IS}$ measures the reduction of heterozygosity in an individual because of non-random mating within population and hence $F_{IS}$ values significantly higher or lower than 0 reveal inbreeding or outbreeding, respectively.
The $F_{IS}$ significance tested by Student’s $t$-test (Table 1) reveals that flocks A, CM, EP, FT, P, PB, PO, SG and VR do not vary significantly from the condition of equilibrium. Five flocks (ET, G, PA, RA and V) show a significantly high $F_{IS}$ value ($p < 0.05$) only if a standard Bonferroni correction for multiple comparisons is not applied, while the other three (CPL, MA and PI) show a significant inbreeding ($p < 0.05$) also after its application.

The log-likelihood-based assignment of genotypes correctly attributed a great proportion of individuals to their flock of origin (Table 1).

The first two axes from the PCA are plotted in Fig. 2. While V, PO, PI, CPL, VR, CM and G are well separated, little differentiation is revealed among flocks A, ET, SG, MA, FT, PB, PE, EP, RA and PA, and virtually no difference is shown between MA and FT, SG and ET, PB and PE.

**Discussion**

Although most of the microsatellites used were derived from bovine studies, they were suitable to detect polymorphisms in sheep (Slate et al. 1998). In fact, all the four microsatellite loci showed a high number of alleles within flock, providing a useful tool to assess the genetic diversity at flock level (Goudet and Keller 2002).

Positive $F_{IS}$ values could be derived from inbreeding or from the presence of a substructure within the population. In this case, a substructure would be present if rams are bred only with a particular group of ewes within the flock, a practice not very much in use in the livestock system considered, where rams breed with all the ewes in the flock.

---

![Fig. 1. Coefficient of variation of $F_{ST}$ computed by bootstrapping (1000 samples) on number of individuals (a) and of loci (b).](image)
Therefore, the heterozygote deficiency is more likely to arise from the relationship of the individuals used for reproduction.

The different $F_{IS}$ values of the flocks reflect different levels of inbreeding. Usually, rams are bought from a few highly reputed breeders and used within the flock for several breeding seasons (Provincial Association of Breeders, pers. comm.); therefore males may mate with relatives as close as daughters. This is particularly likely where the assessment of parentage is not carried out. The three flocks showing the highest level of inbreeding (CPL, MA and PI) should be considered to be at risk as this condition can lead to genetic diseases carried by rare recessive alleles and moreover their fitness may be strongly reduced (Meszaros et al. 1998). Flocks ET, G, PA, RA and V, showing a significant but lower excess of homozygotes, should be also carefully considered, in that they are likely to be in a borderline condition, with a suspect of genetic fragility.

The policy of exchanging rams between farms may sound appealing to reduce inbreeding, and many breeders in the area adopt it. However, when ram exchanges are made between flocks having a similar genetic pool, as revealed by the PCA, very few genetic variants are gained by this practice.

The log-likelihood of individual genotypes belonging to a flock can be a method to select rams suitable for exchanging between two or more flocks. In fact, if individuals can be attributed to more flocks, then any exchange between them is not likely to improve the genetic variability, as they are likely to carry the same alleles. Rules for including/excluding rams for exchange are easily build-up by likelihood ratio at any desired threshold. Maximum likelihood is a relatively robust method for assessing similarities in comparison with phylogenetic approaches (like genetic distances), and as it is independent from the model assumed, it can also be used in the presence of very recent separations (Davies et al. 1999) and it can be used sequentially as genetic information is gathered.

Taking into account the very limited radius of the area considered (approximately 50 km), it is interesting to note that the spatial distribution of the flocks explained by PCA partially fits with the geographical locations of the farms in the study area (Fig. 3), indicating that the gene flow among the flocks is rather limited.

This study confirms microsatellites as a useful tool for the assessment of genetic variability within and among flocks and for the selection of breeding animals from divergent groups in order to maximize the genetic variation and consequently the fitness (Smith et al. 1998; Goudet and Keller 2002) and can be a model for future studies in

![Fig. 2. Principal coordinate analysis of allele frequencies from 11 microsatellite loci typed in 17 populations from the Viterbo area, defined by the first and second dimensions. Their respective contributions to the grouping are shown as percentages](image-url)
different conditions of rearing, and in other species showing the same breeding problems (goats, buffalo and small populations of cattle).

In particular, the assignment of individual genotypes to flocks performed by the log-likelihood analysis represents a quick and easy method for identifying rams more likely to increase genetic diversity.

Acknowledgements

We wish to thank Monica Agostini for collecting most of the samples, the breeders that kindly provided them, and Loredana Bedini for DNA extractions. This work has been supported by grants of Regione Lazio (OB5b cod. 11207003 and 1128003).

References

Davies, N.; Villalobos, F. X.; Roderick, G. K., 1999: Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. TREE 14: 17–21.


*Author's address:* L. Pariset, Dipartimento di Produzioni Animali, Università della Tuscia, via S. Camillo de Lellis, 01100 Viterbo, Italy. Tel.: +39 761 357447; Fax: +39 761 357434; E-mail: pariset@unitus.it