Analysis of the genetic diversity between Gentile di Puglia, Sopravissana and Sarda sheep breeds using microsatellite markers

Bianca Moioli, Francesco Napolitano, Luigi Orrù, Gennaro Catillo

Istituto Sperimentale per la Zootecnia. Monterotondo (RM), Italy

Corresponding author: Dr. Bianca Moioli. Istituto Sperimentale per la Zootecnia. Via Salaria 31, 00016 Monterotondo (RM), Italy – Tel. +39 06 90090213 – Fax: +39 06 9061541 – Email: bianca.moioli@isz.it

Paper received March 23, 2005; accepted October 11, 2005

ABSTRACT

Two Italian sheep breeds – Gentile di Puglia and Sopravissana – have undergone a consistent decline in numbers and have been widely substituted by the Sarda breed, which gives higher milk yield. With the aim to provide a basis for conservation decisions, the genetic variability of the three breeds was investigated in this study. DNA of 60 animals was PCR amplified at the following microsatellite loci: CSSM43, CSSM47, CSSM60, TGLA110, TGLA122, TGLA126, TGLA377, ETH3, ETH10, ETH225, NRAMP1, OARCP20 and SPS115. Allele frequencies, gene diversity and genetic distances were calculated. The highest differences in allele frequencies were found at the following loci: TGLA377, TGLA122, NRAMP1, OARCP20 and ETH3, while at the other loci the most frequent alleles were the same. The average inbreeding rate was 0.156 in the Gentile di Puglia, 0.158 in the Sopravissana and 0.137 in the Sarda. Average gene diversity of the analysed microsatellites was 0.59. Genetic distance between Gentile di Puglia and Sopravissana (0.081) indicates moderate differentiation; distances between the Sarda and the endangered breeds – 0.111 from the Gentile di Puglia and 0.107 from the Sopravissana – indicate a medium-high differentiation rate. The disappearance of the two less productive breeds would entail a consistent loss of genetic diversity. The inbreeding values are low enough to allow the implementation of sound conservation programmes.

Key Words: Sheep, Gentile di Puglia, Sopravissana, Sarda, Genetic diversity.

RIASSUNTO

ANALISI DELLA DIVERSITÀ GENETICA TRA LE RAZZE OVINE GENTILE DI PUGLIA, SOPRAVISSANA E SARDA, TRAMITE MARCATORI MICROSATELLITI

Le razze ovine Gentile di Puglia e Sopravissana hanno subìto negli ultimi quarant’anni un impressionante declino numerico, fino all’orlo dell’estinzione. Si è assistito, infatti, a un processo di sostituzione delle due razze con la Sarda, che, grazie alla elevata produzione di latte, è risultata più competitiva nella maggior parte delle regioni italiane. Su 60 soggetti appartenenti alle tre razze è stata misurata la variabilità genetica sulla base delle frequenze alleliche a 13 loci microsatelliti: CSSM43, CSSM47, CSSM60, GLA110, TGLA122, TGLA126, TGLA377, ETH3, ETH10, ETH225, NRAMP1, OARCP20 e SPS115. Il numero di alleli individuati a ogni locus va da 2 (CSSM60 e ETH10) a 14 (TGLA126). Le differenze maggiori nelle frequenze alleliche tra le tre razze si sono riscontrate ai loci TGLA377, TGLA122, NRAMP1, OARCP20, ETH3, mentre negli altri loci gli alleli più frequenti sono gli stessi. Il deficit di eterozigosità medio è risultato di 0,156 per la Gentile di Puglia, 0,158 per la Sopravissana e 0,137 per la Sarda, valori che indicano che quest’ultima è la meno consanguinea. La diversità genetica media considerando tutti i loci è risultata di 0,59. Il grado di differenziazione genetica tra la Gentile di Puglia e la Sopravissana è risultato di 0,081 mentre tra la Sarda e le razze Gentile di Puglia e Sopravissana i valori sono risultati di 0,111 e 0,107 rispettivamente. Pertanto, le due razze ritenute a rischio sono da considerarsi moderatamente differenziate tra loro, mentre la differenziazione di queste con la Sarda è risultata medio-alta. In conclusione, la
Introduction

In Italy the Sarda sheep breed has played the same role as the Holstein Friesian, although on a smaller scale. In fact, being a high milk yielding indigenous breed of the Sardinian island, it spread out to the Italian peninsula half century ago and was preferred by many shepherds who abandoned their low-producing sheep breeds: Gentile di Puglia and Sopravissana. Until 1963 these two multi-purpose indigenous sheep breeds represented the most important sheep resource in Southern and Central Italy, respectively numbering to about one million head of each of the two (ASSONAPA, 1972), and successfully fulfilled the market demand for sheep cheese. In 1992 there were about 300,000 of each of these breeds (Sarti, 1992), while estimates for 2004 indicate that no more than 20 thousand Gentile di Puglia and 4 thousand Sopravissana are left (ASSONAPA, 2004). Conversely, the population of Sarda sheep, which numbered about 2.5 million heads in 1963 (ASSONAPA, 1972), increased to over 4 million in 1992 (Sanna, 1992). We can assume that the economic competition of the Sarda breed will lead to a dangerous loss of genetic diversity that nobody will be able to reemploy in the future. In fact, although no official milk recording exists for the two indigenous breeds, shepherds state the following: 1. They produce 0.3-0.5 litres milk/day, 2. The length of the lactation is extremely variable (60 to 150 days), 3. The lamb has a very good meat conformation so that it can be sold to the butcher at a higher price than the Sarda lamb, 4. They are docile and have an excellent and “sparing” grazing system. This information should be considered as a starting point for a detailed scientific investigation of their productive and morphological traits in order to propose them in the context of sustainable agriculture to compete with the Sarda breed, which produces 198 litres in one lactation (AIA, 2003). The purpose of the present work was simply to estimate the genetic diversity between these breeds and their inbreeding rate using microsatellites as markers in the first step toward the knowledge of the extent of the genetic diversity (Hammond, 1998).

Material and methods

Sheep for genotyping were chosen in order to assure that they were a representative sample of each breed, therefore non-related animals, originating from different flocks. The sample consisted of 25 Gentile di Puglia (GP), 20 Sopravissana (SV) and 15 Sarda (SA). Extracted DNA from frozen blood (GENOMIX extraction kit, Talent, Trieste) was PCR amplified at the following microsatellite loci: CSSM43, CSSM47, CSSM60 (Moore et al., 1994); TGLA110, TGLA122, TGLA126, TGLA377 (Georges and Massey, 1992); ETH3, ETH10 (Solinas Toldo et al., 1993); ETH225 (Steffen and Eggen, 1993); NRAMP1 (Matthews and Crawford, 1998); OARCP20 (Ede et al., 1995) and SPS115 (Moore and Byrne, 1993). Primer sequences and amplification conditions were those referred in literature. The selected microsatellites are located on 7 different chromosomes (Solinas Toldo et al., 1993; Moore et al., 1994; McGraw et al., 1997).

Allele size was measured by Genescan software on the detected DNA fragments by the Perkin Elmer ABI Prism 310 DNA sequencer.

Two approaches were used in analysing the genotyping results. The first was to calculate all common estimators of genetic variability (Wright, 1943,1951; Nei, 1973) between and within breeds. The second approach was to pretend to ignore the breed of each sample and to make groups of the animals on the basis of the most common allele recurrence in each group, in order to maximise allelic similarities within each group, thereby maximising allelic differences between groups.

The analyses of the genetic variation of each breed and locus were performed using the FSTAT computer programme (Goudet, 1995). This programme calculates average gene diversity at each
locus according to Nei's formula (Nei, 1973); observed and expected heterozygosity at each locus and overall (Nei, 1978); and Weir and Cockerman (1984) estimators of Wright's fixation indices ($F_{st}$, $F_{is}$, and $F_{it}$) of genetic differentiation.

The probabilistic assignment of the animals to different populations on the basis of allele frequencies was performed through the program Structure (Pritchard et al., 2000) that implements a model-based clustering method for inferring population structure using genotype data consisting of unlinked markers. The percentage of animals of each breed admitted to the appropriate cluster at different probability levels was determined by direct counting.

**Results**

In the three breeds together, number of detected alleles ranges between 2 (CSSM60 and ETH10) and 16 (TGLA122), with the average equal to 7.9. Major differences between breeds in allele frequencies were found at: TGLA377 (allele 92 is in 36.7% of the Sarda animals and in less than 12% of the others; allele 104 in 40% of the Sopravissana and absent in the Sarda); TGLA122 (allele 129 is in 46.7% of the Sarda animals and absent in the others; allele 131 in 54% of the Gentile di Puglia and 3.3% in the Sarda); OARCP20 (allele 66 is in 40% of the Sarda animals and in 5% of the Sopravissana); and TGLA126 (alleles 112 and 128 are present only in the Gentile di Puglia; allele 118 only in the Sopravissana and allele 122 only in the Sarda). At the remaining loci the most frequent alleles are the same in the three breeds.

Average gene diversity (Nei, 1973) over all loci and for the three breeds together was 0.59 (Table 1). The microsatellite loci showing the highest gene diversity were TGLA126 (0.88); TGLA122 (0.83); TGLA377 (0.80). Of the three breeds, GP showed the highest gene diversity over all considered loci (0.59), followed by SA (0.53) and SV (0.51).

The observed heterozygosity (Table 2) was also highest in the GP (0.127) compared to SA (0.122) and SV (0.121), although the values were not significantly different. Moreover, the observed heterozygosity was slightly lower than the expected one according to Hardy-Weinberg equilibrium,

| Locus      | N. alleles | Gene diversity % | N. alleles | Gene diversity % | N. alleles | Gene diversity % | N. alleles | Gene diversity % |
|------------|------------|------------------|------------|------------------|------------|------------------|------------|------------------|
| CSSM43     | 13         | 0.79             | 8          | 0.85             | 5          | 0.56             | 9          | 0.84             |
| CSSM47     | 6          | 0.35             | 2          | 0.54             | 3          | 0.19             | 3          | 0.07             |
| CSSM60     | 2          | 0.12             | 1          | 0.21             | 2          | 0.18             | 2          | 0.25             |
| TGLA122    | 16         | 0.83             | 7          | 0.68             | 4          | 0.81             | 6          | 0.69             |
| TGLA126    | 13         | 0.88             | 8          | 0.89             | 5          | 0.79             | 7          | 0.89             |
| TGLA377    | 9          | 0.80             | 9          | 0.82             | 6          | 0.73             | 7          | 0.78             |
| ETH3       | 6          | 0.75             | 5          | 0.66             | 5          | 0.74             | 4          | 0.55             |
| ETH10      | 2          | 0.05             | 1          | 0.13             | 2          | 0.16             | 1          | 0.1             |
| NRAMP1     | 8          | 0.73             | 7          | 0.75             | 5          | 0.72             | 6          | 0.70             |
| OARCP20    | 9          | 0.77             | 8          | 0.82             | 4          | 0.64             | 4          | 0.71             |
| SPS115     | 7          | 0.67             | 4          | 0.74             | 3          | 0.53             | 8          | 0.63             |
| ETH225     | 6          | 0.68             | 6          | 0.69             | 4          | 0.59             | 4          | 0.67             |
| TGLA110    | 6          | 0.11             | 6          | 0.20             | 1          | 0.24             | 3          | 0.13             |

Mean/locus 7.9 0.59 5.3 0.59 3.7 0.51 4.4 0.53
Table 2. Sample size, number of alleles per locus and heterozygosity averaged over 13 microsatellite loci.

| Breed | N. animals | Average heterozygosity | observed | expected |
|-------|------------|------------------------|----------|----------|
| GP    | 25         | 0.127 ± 0.080          | 0.146 ± 0.082 |
| SV    | 20         | 0.122 ± 0.064          | 0.141 ± 0.062 |
| SA    | 15         | 0.121 ± 0.078          | 0.138 ± 0.076 |
| Total | 60         |                        |           |          |

Average inbreeding rate (Table 3) was only slightly higher in the two indigenous breeds (0.156 in GP; 0.159 in SV) than in the SA (0.137), and differences between breeds, performed by Students t, were not significant.

Genetic differentiation between GP and SV (0.081; Table 3) is moderate (Hartl, 1980); while the differentiation between the SA and the indigenous ones is of medium-high rate. Therefore, the disappearance of the two less productive breeds would entail a loss of genetic diversity. Although many people are convinced that GP and SV are very similar - because their morphology and performances are similar - the genetic differentiation that was found in the present work shows that the belief is not true. This is confirmed by the fact that the two breeds have evolved separately in two different Italian areas, the Southern Apennines for the former, the Central Apennines and the Roman lowland for the latter, with different socio-economic constraints and no exchange of breeding animals between them. This difference should be preserved through opportune programmes, which aim for the sustainable use of such breeds.

Further comments are necessary to discuss the results (Table 4) of the probabilistic assignment of plankton species to each of the 3 clusters (Structure software) indicate that 80% SA, 85% SV and 68% GP can be assigned to the appropriate breed with a probability higher than 80% (Table 4).

Discussion

Among the obtained results, two items indicate that the three breeds maintain a random mating structure. The first is that the observed heterozygosity is not significantly different from the expected one, according to Hardy-Weinberg equilibrium (Table 2); the second is that the average inbreeding rate (Table 3), although slightly higher in the two indigenous breeds, is to be considered low. In fact, no organised selection scheme for the improvement of productivity has been applied in the two indigenous ones; while in the SA, the selection scheme that has been applied during the last two decades for the improvement of milk yield is too recent to have made consistent genomic changes at breed level. This is a good starting point for a conservation and improvement programme for both the GP and the SV breeds, while the inbreeding rate is still more than acceptable.

According to Hartl (1980) the differentiation values that were found between GP and SV (Table 3) should be considered moderate, while the differentiation between the SA and the indigenous ones is of medium-high rate. Therefore, the disappearance of the two less productive breeds would entail a loss of genetic diversity. Although many people are convinced that GP and SV are very similar - because their morphology and performances are similar - the genetic differentiation that was found in the present work shows that the belief is not true. This is confirmed by the fact that the two breeds have evolved separately in two different Italian areas, the Southern Apennines for the former, the Central Apennines and the Roman lowland for the latter, with different socio-economic constraints and no exchange of breeding animals between them. This difference should be preserved through opportune programmes, which aim for the sustainable use of such breeds.

Further comments are necessary to discuss the results (Table 4) of the probabilistic assignment of plankton species to each of the 3 clusters (Structure software) indicate that 80% SA, 85% SV and 68% GP can be assigned to the appropriate breed with a probability higher than 80% (Table 4).

Table 3. Fixation indexes: total (on diagonal Fis = inbreeding within breed) and between breeds (Fst = genetic distances).

| Breed | GP  | SV  | SA  |
|-------|-----|-----|-----|
| GP    | 0.156 | 0.081 | 0.111 |
| SV    | 0.158 | 0.107 | 0.137 |
| SA    | 0.111 | 0.107 | 0.137 |
the animals to their breed on the basis of allele frequencies, performed through the program \textit{Structure}. The GP again shows the highest genetic variability: only 15 of them (60\%) can be assigned to the correct breed with a probability higher than 90\%. Even more surprising is the fact that 28\% of the sheep can be assigned to their own breed with a probability lower than 60\%. The SV breed demonstrates opposite features: 85\% of the sheep fall in the correct breed cluster with a probability higher than 80\% and all of them fall in it with probability higher than 60\%. We tried to explain this unexpected difference between the GP and the SV, and tried to answer the question as to why 28\% of the GP sheep are assigned to the GP cluster with probability lower than 60\%? Apparently, GP and SV have suffered the same destiny: the triple-purpose attitude became less and less important due the loss of value of Italian wool; milk production was not competitive; lamb production could become the only way to make the breeds survive. In fact, during the Sixties, extensive crossbreeding with foreign ram breeds used for meat production (Ile-de-France, Berrichon, Rambouillet) was performed by many shepherds to such an extent that in 1997 the Herdbook Society, established new regulations for admission to the Herdbook of a new breed, named “Merinizzata”, which originated through the above-mentioned crossbreeding practices (ASSONAPA, 1997).

The creation of the new officially recognised breed also aimed to protect those shepherds who wanted to maintain the original triple-purpose type, i.e. the pure GP and SV breeds. In fact, each of the three breeds is now considered a peculiar genetic entity, and admission to the respective Book occurs when the animal fulfils specific standard requirements. Someone might infer that the higher genetic variability of the GP is the consequence of uncontrolled crossbreeding practices. However, because the analysed sheep were chosen among pure-bred animals, according to the judgement of official experts, we believe that such variability is merely the mirror of the intrinsic genetic variability of the breed, maintained thanks to the random-mating structure and the sufficient number of animals. Therefore, we can also conclude that the move from the pure breed to the Merinizzata affected the SV more than the GP, very likely because the geographical areas where the GP is reared include marginal lands where higher milk yielding breeds cannot survive. On the contrary, most milk producing shepherds of the Lazio region have moved to more productive breeds. If we compare the present total numbers of the three breeds: 20,000 head the GP, 4000 the SV and 600,000 the Merinizzata (ASSONAPA, 2004), with the numbers of the Herdbook registered animals: 4587; 3472 and 15,040 for the three breeds, respectively (ASSONAPA, 2005), it is evident that what is left of the SV is almost all registered in the Herdbook, which explains the little genetic variability that we have found, although animals were sampled in different flocks. Of the considered breeds, the SV is, therefore, the most endangered, but the substrate for a conservation as well as improvement programme exists. The inbreeding rate is still at an acceptable level; the heterozygosity level is also good (similar to that of the SA breed); the genetic difference from the other

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{P \%} & \textbf{GP} & & \textbf{SV} & & \textbf{SA} \\
\hline
& \textbf{N.} & \textbf{\%} & \textbf{N.} & \textbf{\%} & \textbf{N.} & \textbf{\%} \\
\hline
\textbf{> 90} & 15 & 60 & 16 & 80 & 10 & 66 \\
\textbf{> 80} & 17 & 68 & 17 & 85 & 12 & 80 \\
\textbf{> 70} & 18 & 72 & 19 & 95 & 13 & 86 \\
\textbf{> 60} & 18 & 72 & 20 & 100 & 14 & 93 \\
\hline
\end{tabular}
\caption{Number and percentage of animals that can be assigned to their breed at different probability levels (P).}
\end{table}
breeds made evident in this study with the use of anonymous non-coding markers should be studied within the candidate genes of potential economic traits that might make SV different, and that could lead to its sustainable use.

Conclusions

Of the three considered breeds, the GP shows the highest genetic variability, while the SV is the most inbred, although inbreeding is maintained at an acceptable level, thus making conservation and improvement programmes feasible. Differentiation rate between GP and SV is moderate; therefore, the disappearance of either of the two would entail a consistent loss of genetic diversity. The creation of the “Merinizzata” new breed by the Herdbook Society made the majority of the SV shepherds abandon the old breed and move toward a more specialised meat breed. This move affected the GP breed to a lesser extent, very likely because the triple-purpose attitude is still necessary in the local conditions of the Southern Apennines. The genetic differentiation between the SA and the two indigenous breeds is of medium-high rate, confirming that the breeds have evolved separately in different geographical areas. In order not to lose such genetic variability, programmes aimed at their sustainable use should be accompanied by the study of polymorphisms within the candidate genes for potential economic traits that might make the indigenous breeds different.

REFERENCES

AIA, 2003. Milk Recording Activity: Official Statistics. Italian Breeders Association (AIA), Roma, Italy.
ASSONAPA, 1972. Razze: consistenza e distribuzione. Ed. R.E.D.A., Roma, Italy.
ASSONAPA, 1997. Caratteri tipici e indirizzi di miglioramento delle razze Derivate Merine. Regolamenti. National Association of Sheep and Goats Breeders (ASSONAPA), Roma, Italy.
ASSONAPA, 2004. Internal statistics. National Association of Sheep and Goats Breeders (ASSONAPA), Roma, Italy.
ASSONAPA, 2005. Home page address: http://www.assonapa.it
EDK, A.I., PIETRASCA, C.A., CRAWFORD, A.M., 1995. Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26 loci. Anim. Genet. 26:129-130.
GEOGHEGAN, M., MARSHALL, J., 1992. Polymorphic DNA markers in Bovidae. Patent WO 92/13102.
GOUDEN, J., 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. J. Hered. 86:485-486.
HAMMOND, K., 1998. Development of the global strategy for the management of farm animal genetic resources. Proc. 6th World Congr. on Genetics Applied to Livestock Production, Armidale, Australia, 28: 43-50.
HARTL, D., 1980. Principles of population genetics. Sinauer Associates Inc., Sunderland, MA, USA.
MATTHEWS, G.D., CRAWFORD, A.M., 1998. Cloning, sequencing and linkage mapping of the NRAMP1 gene of sheep and deer. Anim. Genet. 29:1-6.
MCGRAW, R.A., GROSE, WM, KAPPES, S.M., BEATTIE, C.W., STONE, R.T., 1997. Thirty-four bovine microsatellite markers. Anim. Genet. 28:66-68.
MOORE, S., BYRNE, K., 1993. Dinucleotide polymorphism at the bovine calmodulin independent adenocyclase locus. Anim. Genet. 24:150-158.
MOORE, S., BYRNE, K., BERGER, K.T., BARENDSE, W., McCARTHY, F., WOMACK, J.E., HETZEL, D.J.S., 1994. Characterization of 65 bovine microsatellites. Mamm. Genome 5:84-90.
NEI, M., 1973. Analysis of gene diversity in subdivided populations. P. Natl. Acad. Sci. USA 70: 3321-3323.
NEI, M., 1978. Estimation of average heterozygosity and genetic distance from small number of individuals. Genetics 89:583-590.
PRITCHARD, J.K., STEPHENS, M., DONNELLY, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
SANNA, S., 1992. Le razze ovine italiane da latte ed il loro miglioramento. In: Ovinicoltura, Ed. UNAPOC, Roma, Italy, pp 115-139.
SARTI, D.M., 1992. Razze ovine italiane da carne. In: Ovinicoltura, Ed. UNAPOC, Roma, Italy, pp. 241-249.
SOLINAS TOLDO, S., FRIES, R., STEFFEN, P., NEIBERGS, H.L., BARENDSE, W., 1993. Physically mapped cosmid-derived microsatellite markers as anchor loci on bovine chromosomes. Mamm. Genome 4:720-727.
STEFFEN, P., EGGERT, A., DIETZ, A.B., WOMACK, J.E., STRANZINGER, G., 1993. Isolation and mapping of polymorphic microsatellites in cattle. Anim. Genet. 24:121-124.
WEIR, B.S., COCKERHAM, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358-1370.
WRIGHT, S., 1943. Isolation by distance. Genetics. 28:114-138.
WRIGHT, S., 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.