Trop Anim Health Prod
DOI 10.1007/s11250-011-9873-6

SI FAT TAILED SHEEP

Molecular and pedigree analysis applied to conservation of animal genetic resources: the case of Brazilian Somali hair sheep

Samuel R. Paiva · Olivardo Facó · Danielle A. Faria · Thaísa Lacerda · Gabriel B. Barretto · Paulo L. S. Carneiro · Raimundo N. B. Lobo · Concepta McManus

Accepted: 12 April 2011

Abstract The first registers of Somali sheep in Brazil are from the beginning of the 1900s. This breed, adapted to the dry climate and scarce food supply, is restricted in the northeast region of the country. Molecular marker technologies, especially those based on genotyping microsatellite and mtDNA loci, can be used in conjunction with breeding (pedigree analysis) and consequently the maintenance of genetic variation in herds. Animals from the Brazilian Somali Conservation Nuclei from Embrapa Sheep and Goats in Ceará State were used to validate genetic monitoring by traditional pedigree methods and molecular markers. Nineteen microsatellite markers and 404 base pairs from the control region of mtDNA were used. For total herd diversity, an average 5.32 alleles were found, with expected heterozygosity of 0.5896, observed heterozygosity of 0.6451, 0.4126 for molecular coancestrality, and coefficient of inbreeding (F_{IS}) was −0.095. Comparing molecular coancestrality means over the years, there was a consistent increase in this parameter within the herd, increasing from 0.4157 to 0.4769 in 2 years (approx. 12% variation). Sixteen mtDNA haplotypes were identified. Inbreeding and other estimates from genealogical analyses confirm the results from molecular markers. From these results, it is possible to state that microsatellites are useful tools in genetic management of herds, especially when routine herd recording is not carried out, or there were gaps in recent generations. As well as pedigree control, genetic diversity can be optimized. Based on the results, and despite herd recording in the herd of Brazilian Somali of Embrapa Sheep and Goats, additional management measures need to be carried out in this herd to reduce inbreeding and optimize genetic variation.

Keywords *Ovis aries* · Genetic management · STRs · D-loop · Molecular markers · Genealogical analysis

Electronic supplementary material The online version of this article (doi:10.1007/s11250-011-9873-6) contains supplementary material, which is available to authorized users.

S. R. Paiva · D. A. Faria · T. Lacerda
LGV/PCC, Embrapa Recursos Genéticos e Biotecnologia,
02372 Brasília, DF, Brazil

O. Facó · R. N. B. Lobo
Embrapa Caprinos e Ovinos,
Sobral, CE, Brazil

G. B. Barretto
Faculdades Integradas, UPIS,
Brasilia, DF, Brazil

P. L. S. Carneiro
Universidade Estadual do Sudoeste da Bahia,
Jequié, BA, Brazil

C. McManus (✉)
Universidade Federal do Rio Grande do Sul,
Porto Alegre, RS, Brazil
e-mail: concepta.mcmanus@ufrgs.br

Introduction

The Brazilian Somali fat-tailed sheep has its origins in the horn of Africa and is thought to have the Urial as its ancestor. This animal first arrived in Brazil in 1939, brought by farmers from Rio de Janeiro State, but it did not adapt well to the climatic conditions. The drier and hotter climates found in the northeast of the country were more adequate. At present, the population of this breed is small even though it was used extensively in crosses to create the Dorper, which has recently been imported from South Africa. This means that this breed is threatened with

Published online: 02 May 2011

extinction (Mariante et al. 2009). It is used preferentially for meat production and is well adapted to extensive and semi-extensive farming.

Little information is available on Somali sheep in Brazil, most studies with this breed being based on crossbreeding to produce heavier animals for meat (McManus et al. 2010). Using RAPD markers to study Brazilian breeds, the Somali showed some of the lowest expected heterozygosis under Hardy–Weinberg equilibrium, and the proportion of polymorphic loci (Paiva et al. 2005). These same authors concluded that due to processes of genetic drift and endogamy, Somali has been affected by stochastic events. Among all breeds, Somali showed highest level of differentiation from the other breeds. Carneiro et al. (2010) differentiated between naturalized and commercial breeds in Brazil, Uruguay, and Colombia using morphological descriptors. In this study with 24 breeds, the Somali was one of the smallest breeds being grouped with the Morada Nova sheep, also from the Brazilian Northeast (Ianella et al. 2010) using 13 breeds of sheep in Brazil also saw that the Somali had a low level of susceptibility to Scrapie based on *PRNP* allele frequencies.

These same studies along with Castanheira et al. (2010) and McManus et al. (2009) showed how selection and crossbreeding in the main Brazilian hair breed (Santa Inês) has de-characterized the breed, making it more susceptible to heat and gastrointestinal parasites. This had led more recently to some farmers in arid areas looking to other naturalized breeds to substitute the Santa Inês. Knowledge on genetic issues in a population is important not only as a guide for genetic interventions but also to evaluate the results of selection of individuals for reproduction such as available genetic variability, effective population size, and generation intervals. Knowledge of population structure combined with information on genetic changes in the population can influence future management actions. This can be used to develop strategies for use of a breed in a particular ecosystem and guide genetic improvement programs.

The Brazilian Agricultural Research Company (Embrapa) is responsible for a network of Conservation Nuclei for locally adapted sheep and goats in Brazil. The objectives of this program include: (1) maintain herds with a minimum level of genetic variability in accordance with FAO recommendations; (2) maintain a germplasm bank with semen and embryos; (3) develop and adapt methods for optimizing the conservation of Animal Genetic resources in Brazil. Even with this network, more information is needed on the Somali breed as well as other available genetic resources for future use in dissemination programs. The Embrapa Goat and Sheep in Ceará State maintains a herd of Somali sheep for research purposes. The aim of this paper is to analyze the genetic structure of this specific population using molecular markers and pedigree analysis to subsidize and improve actual conservation efforts as well as future breeding programs.

Material and methods

Pedigree data and biological samples for genetics analysis were obtained from the Brazilian Somali herd form Embrapa Goats and Sheep Research Station in Ceará, State, Northeastern Brazil between 1997 and 2010. Pedigree analysis and parameter estimates based on gene origin probability, inbreeding, and generation interval were performed using the ENDOG program (Gutiérrez and Goyache 2005) in a dataset of 1,089 animals. A brief resumption from the principal analysis are described below.

1. The pedigree completeness level was computed as the proportion of ancestors known per parental generation (MacCluer et al. 1983).
2. The number of equivalent to discrete generations (t) for each individual in a pedigree was computed as the sum of $(1/2)^n$, where n is the number of generations separating the individual to each known ancestor (Boichard et al. 1997).
3. The inbreeding coefficient (F), defined as the probability that two alleles at a randomly chosen locus are identical by descent (Malécot 1948), and the average relatedness coefficient (AR), defined as the probability that an allele randomly chosen from the whole population belongs to a given animal (Goyache et al. 2003a; Gutiérrez et al. 2003) was computed for each individual. The AR coefficient of a founder means its genetic contribution to the population.
4. The probability of gene origin was characterized by computing the following parameters:
   (a) Effective number of founders (f_e), which is the reciprocal of the probability that two alleles drawn at random in the studied population originate from the same founder (James 1972), and this is computed from the genetic contribution of founders to the descendant gene pool of the population (Lacy 1989)
   (b) Effective number of ancestors (f_a), defined as the minimum number of ancestors, not necessarily founders, explaining the complete genetic diversity of a population (Boichard et al. 1997). Parameter f_a does not fully account for gene loss by drift from the ancestors to a reference population but complements the information offered by f_e accounting for the losses of genetic variability produced by the unbalanced use of reproductive individuals producing bottlenecks (Gutiérrez et al. 2005a)

(c) Founder genome equivalents (Ballou and Lacy 1995), defined as the theoretically expected number of founders that would be required to provide the genetic diversity in the actual population if the founders were equally represented and had lost no alleles, was obtained by the inverse of twice the average coancestry of the individuals within the population (Caballero and Toro 2000).

5. Effective population size (N_e) was computed following Gutiérrez et al. (2003) approaching the increase in inbreeding (ΔF) as the regression coefficient (b) of the individual inbreeding coefficient over the equivalent discrete generations (t), and considering the corresponding regression coefficient as the increase in inbreeding between two generations, $b = F_t - F_{t-1} = \frac{F_t - F_{t-1}}{1 - F_{t-1}}$ and consequently $N_e = 1/(2b)$.
6. The migration rate (m) was computed for the dataset as the ratio between the number of animals from the parental populations acting as parents over the whole number of parents (from the parental population or not) identified in the reference population.
7. The Wright's F statistics—F_{IT}, F_{ST}, and F_{IS}—were calculated as: F_{IT} is the average inbreeding coefficient of population; F_{ST} represents the expected average inbreeding coefficient, estimated in a hypothetic populations produced by random mating of parents within each period; and F_{IS} expresses the deviation generated from actual mating.
8. The coefficient of coancestry of two individuals is defined as the probability that two gametes taken at random, one from each, carry alleles that are identical by descent (Falconer and Mackay 1996). The within-breed coancestry (f_{ii}) and the between-breeds coancestry matrix (f_{ij}) were computed averaging all pairwise coancestry coefficients of the individuals belonging, respectively, to a given breed i or to two different breeds i and j. Following Caballero and Toro (2002), the between-breeds Nei's minimum distance (D_m) matrix was also computed as $D_m = ((f_{ii} + f_{jj})/2) - f_{ij}$ where f_{ii} and f_{jj} are the average coancestry within two breeds i and j and f_{ij} the coancestry between two breeds i and j. In the present, sires were considered separate populations.
9. The genetic conservation index (GCI) was estimated as stated by Alderson (1992).

$$GCI = 1/\Sigma P_i^2$$

Where P_i is the proportion of genes of founder animal i in the pedigree.

For molecular analysis, 19 microsatellite loci (Table 1) were used on 48 animals of the Brazilian Somali Conservation Nucleus over the last 4 years of data collection. The microsatellites were genotyped in an automatic sequencer with capillary electrophoresis with detection of specific fluorescent dyes contained in the oligonucleotides present in the PCRs. After analysis of allelic frequencies for all loci genetic population, parameters were estimated for the herd using FSTAT (Goudet 2002), GenAlex (Peakall and Smouse 2006), and Molkin (Gutiérrez et al. 2005b) programs. An analysis of principal components from the allele shared index matrix for all individuals was obtained using Ntsys v.2.02g.

Table 1 Genetic diversity using 19 microsatellite loci in a conservation nucleus of Somali sheep

Loci	*k*	*Ho*	*He*	*PIC*	*PExcl*(1)	*PExcl*(2)
Inra23	4	0.63	0.57	0.49	0.164	0.297
Oar304	4	0.70	0.64	0.56	0.208	0.353
MAF214	4	0.87	0.61	0.54	0.184	0.331
INRA63	8	0.86	0.74	0.69	0.335	0.512
OARHH35	9	0.85	0.86	0.83	0.534	0.699
INRA35	5	0.66	0.62	0.54	0.198	0.34
OMHC1	8	0.79	0.79	0.75	0.396	0.575
ILSTS87	8	0.83	0.81	0.77	0.427	0.605
ILSTS05	4	0.27	0.24	0.22	0.029	0.12
ILSTS11	4	0.49	0.64	0.56	0.205	0.353
MAF65	6	1.00	0.74	0.68	0.312	0.485
BM827	6	0.15	0.30	0.28	0.046	0.167
OARFCB20	6	0.92	0.74	0.68	0.31	0.481
OARCP20	5	0.59	0.62	0.58	0.219	0.398
OAR129	3	0.55	0.41	0.33	0.082	0.171
INRA172	6	0.69	0.72	0.67	0.306	0.484
HUJ616	4	0.11	0.10	0.10	0.005	0.052
SRCRSP05	3	0.67	0.50	0.43	0.122	0.247
BM6526	4	0.65	0.58	0.52	0.171	0.327
Mean	5.32	0.65	0.59	0.54	0.9941[a]	0.9999[a]

k number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *PIC* polymorphic information content, *PExcl*(*1*) probability of exclusion using information from one parent and offspring, *PExcl*(*2*) probability of exclusion using information from both parents and offspring

[a] Group probability

A total of 460 base pairs were sequenced from the first half of the control region (D-loop) from 34 of the 48 Somali sheep sampled from the Conservation Nuclei of Embrapa Sheep and Goats. Embrapa is the Brazilian Research Company for Agriculture and Livestock linked to the Brazilian Ministry of Agriculture, Livestock, and Supply. Sequencing was carried out directly on the PCR product using primers suggested by (Hiendleder et al. 1998a) and Big Dye Terminator Cycle Sequencing kit in an Automatic ABI 3700 sequencer following factory recommendations (Applied Biosystems). The sequences were analyzed using Sequence Analysis v.5.1 (Applied Biosystems) software

and edited using SeqScape v2.5 (Applied Biosystems) and the reference sequence AF010406 deposited in NCBI (Hiendleder et al. 1998b). The statistical analyses were carried out using MEGA (Kumar et al. 2008), DNAsp (Librado and Rozas 2009), and Network (Bandelt et al. 1999) softwares. An additional analysis was carried out to verify how much of the haplotype diversity is specific to Brzilian Somali breed. Multiple alignments were made with the following accesses number form GenBank: EF490451–EF490456, EF490493, DQ320065–DQ320075, DQ903190–DQ903207, DQ903228–DQ903304, DQ852128–DQ852142, DQ852148, DQ852152, DQ852153, DQ852156, DQ852158–DQ852159, DQ852161–DQ852246, DQ852280–DQ852289, DQ790890–DQ791249, DQ491576–DQ491736, DQ460429–DQ460467, DQ097431–DQ097468, AY879348–AY879355, AY879361–AY879371, AY879375–AY879385, AY879392–AY879399, AY879404, AY879407, AY879409–AY879412, AY879415–AY879417, AY879420, AY879421, AY879424, AY879426–AY879428, AY879432–AY879439, DQ087255, DQ073049–DQ073053, AY858379, AY829376–AY829430, AY582800–AY582820, AY091487, AY091488, AY091495–AY091500, AF089809, AF010406, AF010407, AF039577–AF039579, Z35293, Z35228–Z35268, L29055.

Results

Pedigree data

Forty-eight percent of the animals born were male and 52% female. Births were distributed throughout the year with no concentration in any season, indicating that breed is not a seasonal breeder in this ecosystem. There was an increased number of lambings between 2001 and 2005, stabilizing thereafter in about 80 lambings a year. This was observed because the flock had been moved from an experimental farm to another and forced the reduction of the herd. This also explains the absence of lambs born in 2006.

The reference population was considered as the animals with both parents known. This population will be smaller than that one used to analyze founders. The effective founder population size was 22.69 with 902 animals in the reference population and 151 ancestors contributing to the reference population. The effective number of founders/ancestors for the reference population was 16/16 with five animals supplying 50% of the ancestors. There were 33 matings between sibs and four parent offspring matings. Mean inbreeding was 0.54% and mean average relatedness 5.80%.

Mean effective herd size has decreased over the years (Table 2). This was caused by a reduction in the number of sires used but principally by the reduction in the number of females. Males had on average 15 offspring while females had approximately 1.5. There were just over one generation in the time period studied (Table 3) with low inbreeding rate per generation (<1%). Generation intervals (Table 4) averaged 3.91 years and were higher for dams (4.74 years) compared with sires (2.8 years). When all animals were considered, the generation interval was slightly higher for males but similar for females. This is because a higher number of females are used for breeding purposes.

Mean inbreeding levels and average relatedness was low (Table 5). No inbreeding was seen in the most recent generation, but this has few animals and is not yet complete. As the generations increased, there was little change in inbreeding and average relatedness (Table 5). This was due to careful control of reproduction in the nucleus. Although very low (<0.01%), there is an increase over the generations, which should be carefully monitored. There were 5.5% matings between sibs and 0.44% between parent and offspring.

The amount of pedigree information increased per generation (Supplemental Material 1) and was relatively similar for males and females. The lack of sire information in the earlier generations was due to multiple ram mating groups, errors in recording, and frequent exchange of farm helpers. In order to help management, conservation genetic index was estimated—CGI (Fig. 1) for males and females increased along the period studied (0.31/year in males and 0.28/year in females).

Table 2 Average number of offspring for sires and dams along the period studied in a Somali sheep conservation nuclei in Brazil

Period	Sires	Av offspring males	Dams	Av offspring females	*Ne*
1997–2000	6	10.67	63	1.56	12.07
1998–2001	5	12.60	31	2.03	5.48
1999–2002	8	14.13	43	1.91	7.81
2000–2003	5	13.40	37	1.62	8.11
2001–2004	4	16.00	24	1.33	5.21
2002–2005	4	16.00	26	1.27	5.83
2003–2006	1	14.00	14	1.00	5.21
2004–2007	2	5.50	8	1.00	3.28

Ne is Effective population size

Table 3 Average number of generations, inbreeding per generation and effective population size in a Somali sheep conservation nuclei in Brazil

	Maximum	Complete	Equivalent
Mean number of generations	1.64	1.11	1.34
Inbreeding per generation %	0.32	0.84	0.58
Effective population size	15.94	59.22	85.48

Using sires to analyze differentiation, F_{IS} was found to be −0.1262; F_{ST}=0.0904 and F_{IT}=−0.0240, all indices being close to zero. Fixation indices *(F* statistics) describe the level of heterozygosity in a population; more specifically, the degree of (usually) a reduction in heterozygosity when compared to Hardy–Weinberg expectation. In the present study, when sires were considered subpopulations, there is low genetic differentiation (11.68% of total genetic variation). F_{IS} can range from −1.0 (all individuals heterozygous or totally outbred) to +1.0 (no observed heterozygotes or totally inbred), while F_{IT} is the mean reduction in heterozygosity of an individual relative to the total population. Self-coancestry was 0.50, inbreeding 0.005, and Nei distance between sires 0.088, showing that no immediate problems in maintenance of genetic diversity within the herd was seen.

Molecular data

It was observed an average of 5.32 alleles, with expected heterozygosity of 0.5896, observed heterozygosity of 0.6451 (Table 6). The inbreeding index ($F_{IS\text{-}M}$) and overall molecular coancestry of the herd was −0.095 and 0.3935, respectively.

Table 4 Generation intervals in Somali sheep conservation nuclei in Brazil

Type	Number	*L*	*s*
Generation intervals[a]			
Father–son	8	2.47	1.12
Father–daughter	131	3.12	1.74
Mother–son	8	4.69	2.51
Mother–daughter	129	4.76	1.78
Total	276	3.91	1.95
Mean age of parents when offspring born[b]			
Father–son	451	3.36	2.22
Father–daughter	458	3.29	2.12
Mother–son	452	4.71	2.11
Mother–daughter	457	4.64	2.03
Total	1818	4.00	2.22

L generation interval, *s* standard deviation

[a] Breeding animals only

[b] All animals

The comparison of various population parameters from two classes of markers over the years (Table 6) showed a decrease in genetic variability along with an increase in molecular coancestrality from 0.4157 to 0.4769 in a 2-year period (years 2 to 4 had a variation of approximately 12%). In order to see the relationship among the animals analyzed, a principal component analysis (PCA) was done (Fig. 2). The PCA shows the differentiation of sampled individuals for each year and the three first components explained 63% (p<0.01) of all genetic variation observed.

A total of 460 base pairs were sequenced from mtDNA D-loop region per animal. After editing, 404 sites were maintained for analysis. Of the 34 individuals that showed the minimum sequence size, 16 haplotypes (GenBank accession numbers HQ593514-HQ593529) and 17 polymorphic sites (SNP) between transitions and tranversions as well as one insertion/deletion were seen (Supplementary material 2). The haplotype H3 was the most frequent, and it was observed in almost all years except year 1. None of the 16 haplotypes were seen a single time. The nucleotide diversity (*Pi*) was 0.0177 and haplotype diversity (*Hd*) was 0.891. In the early years, an increase was observed in the frequencies of H3, H8, and H9 haplotypes. From 2000, the frequency of H3 and H9 declined. No significant alterations were observed in the frequencies of the other haplotypes.

The network analysis (Fig. 3) shows the relationship between all identified haplotypes. Mitochondrial genetic variability of this nucleus shows at least two large groups of haplotypes, one formed by H6, H13, H15, and H16 versus the other 12. From a genetic variability viewpoint, the maintenance of many similar haplotypes in the conservation herd is not desirable. It is preferable to maintain individuals with distinct haplotypes.

The alignment results show that 10 haplotypes (H1, H2, H4, H6, H7, H10, H12, H13, H15, and H16) from the total of 16 observed in the Brazilian Somali were not identified in other breeds. The remaining haplotype identified, when compared with another GenBank sequences, had the following number of similarities: H8 (220 hits), H11 (62 hits), H9 (60 hits), H3 (39 hits), and H12 (eight hits).

Discussion

This the first study in Brazil to merge molecular and pedigree data for conservation purposes. As well as increasing herd size, improvements are underway to minimize conservation issues found here. Other conservation flocks have been seen to have a low effective number of ancestors. For example, the effective number of ancestors for the reference population of conservation nucleus (CN) of Galego sheep in Spain was 13 animals with five explaining 50% of the ancestors (Adán et al.

Table 5 Inbreeding (*F*), Average Relatedness (*AR*) and Effective Population Size (*Ne*) per generation in a Somali sheep conservation nuclei in Brazil

	Number	Mean *F*	% *F*	Mean *F* for Inbred	% Mean *AR*	*Ne*
Complete generation[a]						
0	185	0.000			0.006	
1	605	0.002	0.010	0.198	0.057	254.70
2	296	0.016	0.159	0.100	0.091	35.90
3	3	0.000			0.084	
Maximum generation[b]						
0	175	0.000			0.006	
1	310	0.000			0.047	
2	377	0.012	0.088	0.136	0.076	41.80
3	182	0.005	0.049	0.104	0.086	
4	45	0.010	0.244	0.040	0.076	

[a] Number of full generations traced

[b] Maximum number of generations traced (Gutiérrez et al. 2005a)

2007). The Somali CN has a slightly higher genetic base, but this number is still low. The low number of ancestors which explain half the genetic variation indicates an over usage of a low number of sires. This trend tends to be overcome with the introduction of new sires to form other families inside this CN.

The mean number of animals was reduced over the years. This was due to problems maintaining large effective herd sizes with public resources. The strategy used at present is to optimize the effective herd size to maintain lower herd numbers and intensify the germplasm of the families in this conservation nucleus. Alternatively, this work will be expanded to local farmers who have the breed or are interested in rearing it, so that growth of the breed will be stimulated and that future studies can be regional instead of local.

Generation intervals show a relatively fast turnover of breeding animals. In terms of mean age of parents when offspring were born, these were higher for dams than sires. This is due to the fact that sire turnover is faster than that for dams. These numbers are similar to those found for a CN of Galega (Adán et al. 2007) and Mallorquina sheep (Goyache et al. 2010) in Spain. These are on the lower limit for those calculated by Huby et al. (2003) in six French meat sheep breeds in France but higher than those found by Goyache et al. (2003b) for the Xalda de Asturias breed which is also threatened with extinction.

The values computed for mean inbreeding (0.54%) and mean average relatedness (5.80%) are comparable to those for the Galega (Adán et al. 2007) and Mallorquina (Goyache et al. 2010) sheep in Spain and well below the dangerous level (10%). Although the population here was small, the inbreeding is less than found in the international breeds. This may be due to two factors: careful choice of sires in the conservation herd or lack of pedigree information (Alberto et al. 2007). Low *AR* means that an animal shares a low percentage of genes with the rest of the population; these animals could, therefore, serve to disseminate the breed to other farms or regions in the country.

Regarding the trend of pedigree information, it can be noted that these numbers are significantly higher than those found by Goyache et al. (2010) and Álvarez et al. (2007)

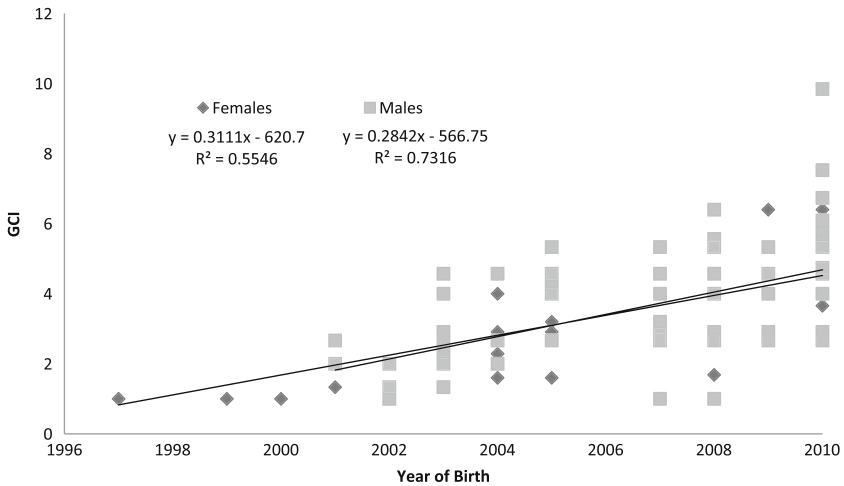

Fig. 1 Genetic Conservation Index in Male and female in Somali sheep conservation nuclei in Brazil over a 13-year period

Table 6 Genetic Diversity indexes in a Somali sheep conservation nuclei in Brazil along four sampling events

Genetic index	Year of sampling			
	Y1	Y2	Y3	Y4
N (STR/mtDNA)	(10/8)	(14/15)	(16/5)	(8/6)
Na	3.789	4.368	4.000	3.368
Na Freq. ≥ 5%	3.789	3.632	3.211	3.368
Ne_M	2.705	2.974	2.517	2.582
I	1.030	1.116	0.996	0.935
No. LComm alleles (≤50%)	0.421	0.474	0.368	0.211
Mean molecular coancestry	0.434	0.416	0.465	0.477
Number of mtDNA haplotypes	6	6	6	5

N number o individuals analyzed for microsatellites (STRs) and mtDNA, *Na* number of different alleles, *Na* (*Freq ≥5%*) number of different alleles with a frequency ≥5%, Ne_M number of effective alleles=1/(Sum π^2); *I* Shannon's information index=−1×Sum (π×Ln (π)), *number of LComm Alleles* (≤*50%*) number of locally common alleles (Freq. ≥5%) found in 50% or fewer populations

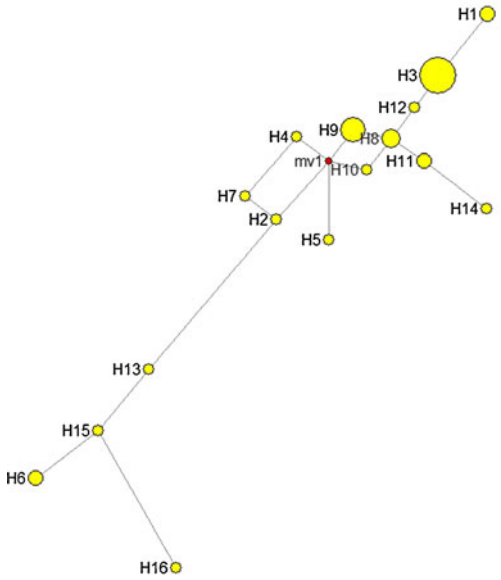

Fig. 3 Network analysis of 16 haplotypes (*yellow*) identified in Somali sheep conservation nuclei in Brazil after sequencing the second half of control region of mtDNA. The *red circle* highlights median vectors generated by Network software

for Mallorquina and by Alfonso et al. (2006) for Laxta Cara Negra Navarra sheep in Spain. These latter authors note that the lack of genealogies leads to computation of high generation intervals, unlike those noted in the present study.

Alderson (1990) used the *GCI* to calculate an effective number of founders in the pedigree of an animal. The higher the *GCI* value the higher the values of an animal for conservation. The *GCI* can be used either by individual breeders as an aid to the selection of herd sire or within an overall breed policy to formulate a group breeding program. However, the index has limitations such as not accounting for any concentration of breeding to non-founder animals in subsequent generations in a pedigree and is inapplicable without pedigree records (Alderson 1992).

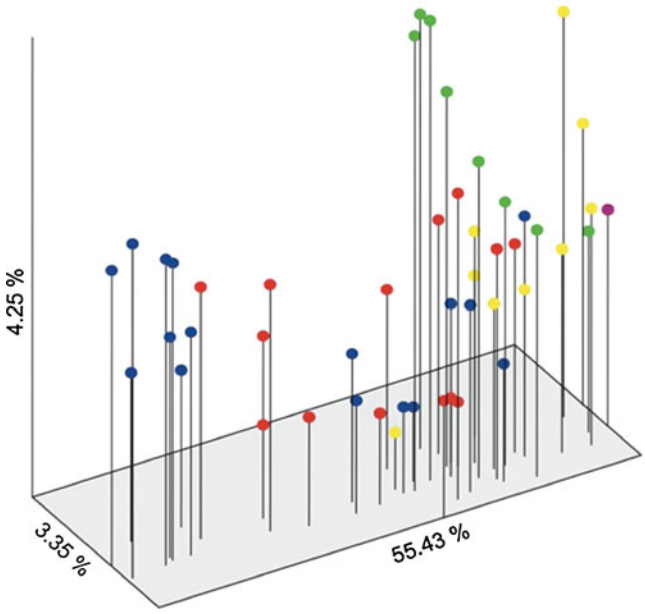

Fig. 2 Principal component analysis from matrix of proportion of alleles shared between 48 individuals from a Somali sheep conservation nuclei in Brazil: *Yellow* animals sampled year 1, *red* animals sampled year 2, *blue* animals sampled year 3, and *green* animals sampled year 4, *purple* ram sampled year 2

Molecular results agreed with pedigree data as they: (1) showed a reduction in genetic diversity over the years (Table 5) and (2) showed that despite this reduction, there is still important genetic diversity (Figs. 2 and 3), which, if managed correctly, may contribute to recovery variability within the population. With the exception of the $F_{IS\text{-}M}$ value, the Brazilian Somali presented lower variability when compared with other Brazilian hair sheep breeds studied previously (McManus et al. 2010; Paiva et al. 2005), which may be explained by the sampling was within a single herd.

The PCA agrees with molecular coancestrality between individuals such that the individuals that were closed in the PCA had highest molecular coancestrality between each other. These tools may aid in culling decisions when choosing parents of the next generation. Animals born on year 3 (lowest mean molecular coancestrality) were more dispersed in Fig. 2 than those born in year 5 (highest molecular coancestrality). This pattern agrees with the diversity observed for year 3. However, unfortunately, there was a variation in sampling for each year that might have influenced the estimations, but new analysis should be done to confirm the observer pattern.

Álvarez et al. (2008), using a Spanish sheep breed as a model, observed that the use of microsatellite markers could be an additional tool for herd management when pedigree is unknown or incomplete. Other authors found similar results (Fernández et al. 2005). Genetic management of a population may be carried out efficiently and at low cost with pedigree data. Molecular markers (such as microsatellites) may be used to monitor pedigree errors. Nevertheless, the use of pedigree and markers together may be a reality with the development of new methods of wide-scale genotyping (Kijas et al. 2009; Meuwissen 2009) and new generation sequencing (Archibald et al. 2010).

The most frequent haplotypes were H3 and H9, being considered haplotypes which were representative of this nucleus (Fig. 2); however, it is important to notice that they represent just one of the two main haplogroups evidenced by network analysis (Fig. 3). The second haplogroup, formed by haplotypes H6, H13, H15, and H16 are specific, until now, for the Brazilian Somali breed. These results will aid in the definition and monitoring of families within the nucleus so that maximum variability can be maintained. Pedigree data should also be evaluated together with molecular results to understand the distribution of haplotypes over the years and help the definition of the families that should be maintained in the conservation nuclei.

The mtDNA haplotypes and microsatellite markers not only confirm the reduction in variability over the years but also show a wide haplotypic diversity compared with other Brazilian breeds (Silvério et al. 2006). It was not an expected result for a highly inbred herd such as this. In this way, these results reinforce two major points in conservations studies: (1) the utilization of just one class of molecular marker (in general, microsatellites) to help conservation polices and managements is not desirable, as they do not show all information. (2) studies with mtDNA in livestock species should address the importance of higher sampling per breed.

The results of the present study show that different methods can be used to accelerate the conservation process so that breed studies can be inserted more quickly into market niches. In this way, for example, microsatellite loci are useful for herd management especially when pedigree control is absent. However, additional management measures need to be taken (e.g., raise the number of rams used per generation and introduce new animals) to reduce inbreeding in the studied herd to optimize genetic variability. Those results also will be used to design families within the conservation nucleus to optimize genetic variability and increase effective size of the herd. These analyses will also aid in the ex situ conservation program for the choice of animals that will have their germoplasm collected for inclusion in the National Animal Germplasm bank of Embrapa's Animal Genetic Conservation Network.

References

Adán, S.A., Fernández, M.F., Justo, J.R.J., Rivero, C.J.R., Rois, D.R., Lama, J.L., 2007. Análisis de la información genealógica en la raza ovina ovella galega. Archivos de zootecnia 56, 587–592.

Alberto, S., Valentino, B., Francesca, T.M., Paola, S., 2007. Genetic variability and population structure in the italian haflinGer horse from pediGree analysis1. Ann. Fac. Medic. Vet. di Parma XXVII, 199–210.

Alderson, L.A., 1990. Genetic Conservation of Domestic Livestock. CAB International, Wallingford, UK.

Alderson, G.L.H., 1992. A system to maximize the maintenance of genetic variability in small populations. Genetic Conservation of Domestic Livestock 2, 18–29.

Alfonso, L., Parada, A., Legarra, A., Ugarte, E., Arana, A., 2006. Effects on genetic variability of selection against scrapie sensitivity in the Latxa Black-Faced sheep. Genet. Sel. Evol 38, 495–511.

Álvarez, I.A., Fernández, I.F., Espinhosa, M.A.E., Gutiérrez, J.P.G., Royo, L.J.R., Goyache, F.G., 2007. Anális del libro genealógico de La raza ovina mallorquina. SEOC.

Álvarez, I., Royo, L.J., Gutiérrez, J.P., Fernández, I., Arranz, J.J., Goyache, F., 2008. Relationship between genealogical and microsatellite information characterizing losses of genetic variability: Empirical evidence from the rare Xalda sheep breed. Livestock Science 115, 80–88.

Archibald, A.L., Cockett, N.E., Dalrymple, B.P., Faraut, T., Kijas, J. W., Maddox, J.F., McEwan, J.C., Hutton Oddy, V., Raadsma, H. W., et al., 2010. The sheep genome reference sequence: a work in progress. Anim Genet 41, 449–453.

Ballou, J.D., Lacy, R.C., 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations. Population management for survival and recovery: analytical methods and strategies in small population conservation, 76–111.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16, 37.

Boichard, D., Maignel, L., Verrier, E., 1997. The value of using probabilities of gene origin to measure genetic variability in a population. Genetics, Selection, Evolution: GSE 29, 5.

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. Conservation Genetics 3, 289–299.

Carneiro, H., Louvandini, H., Paiva, S.R., Macedo, F., Mernies, B., McManus, C., 2010. Morphological characterization of sheep breeds in Brazil, Uruguay and Colombia. Small Ruminant Research.

Castanheira, M., Paiva, S.R., Louvandini, H., Landim, A., Fiorvanti, M.C., Dallago, B.S., Correa, P.S., McManus, C., 2010. Use of heat tolerance traits in discriminating between groups of sheep in central Brazil. Trop Anim Health Prod 42, 1821–1828.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to quantitative genetics (1996) Essex (UK). Longman Group Limited, Edinburgh.

Fernández, J., Villanueva, B., Pong-Wong, R., Toro, M.A., 2005. Efficiency of the use of pedigree and molecular marker information in conservation programs. Genetics 170, 1313–1321.

Goudet, J., 2002. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3.2). Lausanne: University of Lausanne, Department of Ecology & Evolution.

Goyache, F., Gutiérrez, J.P., Alvarez, I., Fernández, I., Royo, L.J., Gómez, E., 2003a. Genetic analysis of calf survival at different preweaning ages in beef cattle. Livestock Production Science 83, 13–20.

Goyache, F., Gutiérrez, J.P., Fernández, I., Gómez, E., Alvarez, I., Diez, J., Royo, L.J., 2003b. Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. J Anim Breed Genet 120, 95–105.

Goyache, F., Fernández, I., Espinosa, M.A., Payeras, L., Pérez-Pardal, L., Gutiérrez, J.P., Royo, L.J., Álvarez, I., 2010. Análisis demográfico y genético de la raza ovina Mallorquina. ITEA, información técnica económica agraria, 3.

Gutiérrez, J.P., Goyache, F., 2005. A note on ENDOG: a computer program for analysing pedigree information. J Anim Breed Genet 122, 172–176.

Gutiérrez, J.P., Altarriba, J., Díaz, C., Quintanilla, R., Cañón, J., Piedrafita, J., 2003. Pedigree analysis of eight Spanish beef cattle breeds. Genet Sel Evol 35, 43–63.

Gutiérrez, J.P., Marmi, J., Goyache, F., Jordana, J., 2005a. Pedigree information reveals moderate to high levels of inbreeding and a weak population structure in the endangered Catalonian donkey breed. J Anim Breed Genet 122, 378–386.

Gutiérrez, J.P., Royo, L.J., Alvarez, I., Goyache, F., 2005b. MolKin v2.0: a computer program for genetic analysis of populations using molecular coancestry information. J Hered 96, 718–721.

Hiendleder, S., Lewalski, H., Wassmuth, R., Janke, A., 1998a. The complete mitochondrial DNA sequence of the domestic sheep (Ovis aries) and comparison with the other major ovine haplotype. J Mol Evol 47, 441–448.

Hiendleder, S., Mainz, K., Plante, Y., Lewalski, H., 1998b. Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. J Hered 89, 113–120.

Huby, M., Griffon, L., Moureaux, S., De Rochambeau, H., Danchin-Burge, C., Verrier, 2003. Genetic variability of six French meat sheep breeds in relation to their genetic management. Genet Sel Evol 35, 637–655.

Ianella, P.I., Caetano, A.R.C., McManus, C.M., Paiva, S.R.P., 2010. In: Evaluation of PRNP Polymorphism in Brazilian Local Sheep Breeds. In: 9th World Congress on Genetics Applied to Livestock Production., pp. 1–4.

James, J.W., 1972. Computation of genetic contributions from pedigrees. TAG Theoretical and Applied Genetics 42, 272–273.

Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J. F., McGrath, A., Wilson, P., Ingersoll, R.G., McCulloch, R., et al., 2009. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS One 4, e4668.

Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9, 299–306.

Lacy, R.C., 1989. Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biology 8, 111–123.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451.

MacCluer, J.W., Boyce, A.J., Dyke, B., Weitkamp, L.R., Pfenning, D. W., Parsons, C.J., 1983. Inbreeding and pedigree structure in Standardbred horses. J Hered 74, 394.

Malécot, G., 1948. Les mathématiques de l'hérédité. Masson.

Mariante, A.S., Albuquerque, M., Egito, A.A., McManus, C., Lopes, M.A., Paiva, S.R., 2009. Present status of the conservation of livestock genetic resources in Brazil. Livestock Science 120, 204–212.

McManus, C., Louvandini, H., Paiva, S.R., de Oliveira, A.A., Azevedo, H.C., de Melo, C.B., 2009. Genetic factors of sheep affecting gastrointestinal parasite infections in the Distrito Federal, Brazil. Vet Parasitol 166, 308–313.

McManus, C., Paiva, S.R., Araújo, R.O., 2010. Genetics and breeding of sheep in Brazil. Revista Brasileira de Zootecnia 39, 236–246.

Meuwissen, T., 2009. Towards consensus on how to measure neutral genetic diversity? J Anim Breed Genet 126, 333–334.

Paiva, S.R.P., Silvério, V.C.S., Egito, A.A.E., McManus, C.M.M., Faria, D.A.F., Mariante, A.S.M., Castro, S.T.R.C., Albuquerque, M.S.M.A., Dergam, J.A.D., 2005. Genetic variability of Brazilian Hair Sheep Breeds using RAPD-PCR markers. Pesquisa Agropecuária Brasileira 40, 887–893.

Peakall, R.O.D., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288–295.

Silvério, V.C., Paiva, S.R., FARIA, D.A., 2006. Phylogenetic study on the main brazilian naturalized sheep breeds. In: 8th World Congress on Genetics Applied to Livestock Production. pp.1–4.