Genetic relationships between Spanish Assaf (Assaf.E) and Spanish native dairy sheep breeds


a Castellana de Ganaderos Sociedad Cooperativa, Campo Real, E-28510 Madrid, Spain
b SERIDA-Somió, C/Camino de los Claveles 604, E-33203 Gijón (Asturias), Spain
c Dpto. de Producción Animal, Facultad de Veterinaria, Universidad Complutense, E-28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 9 April 2008
Received in revised form 1 September 2008
Accepted 4 September 2008
Available online 19 October 2008

Keywords:
Assaf
Genetic distance
Microsatellite
Molecular kinship
Sheep breeds

ABSTRACT

At present, the Assaf is the main dairy sheep in Spain. The Spanish Assaf (Assaf.E) was formed by male-mediated absorption of native Spanish sheep. Here we assess the genetic relationships among the Assaf.E and major native Spanish dairy breeds using microsatellites to contribute to the knowledge of the formation and within-population genetic variability of the breed. Blood samples from 44 unrelated Assaf.E individuals from 23 different Assaf.E flocks spread throughout 6 different Spanish provinces were obtained and genotyped using 14 microsatellites. Up to 312 additional samples belonging to the Awassi and Milchschaf sheep breeds and to six native Spanish dairy sheep breeds (Castellana, Churra, Latxa, Manchega, and Rubia de El Molar) as well as samples from Merino individuals to be used as the outgroup were also analysed (H₀) and expected (Hₑ) heterozygosity, rarefacted number of alleles per locus and distances based on molecular coancestry information were computed. Probabilities of assignment of the Assaf.E individuals to native Spanish dairy sheep breeds and cryptic genetic structure in the whole dataset were also assessed. It can be concluded that the Assaf.E breed has low genetic variability and high genetic distance with respect native Spanish dairy sheep breeds. From our results, the formation of the Assaf.E breed basically occurred via the absorption of individuals belonging to the Entrefino type, particularly to the Castellana and Manchega populations. Furthermore, Churra individuals may have participated in the formation of the Assaf.E breed at an early moment of the introduction of the breed into Spain.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Recent reports (Ugarte et al., 2001) estimated that 45% of the roughly 5 million head of sheep exploited for dairy in Spain are foreign breeds (mainly Assaf, Awassi and Lacaune) or crosses. The Assaf sheep breed was formed in Israel by crossbreeding the East Friesian (Milchschaf) with the Awassi breed (Goot, 1986). The first Assaf individuals were introduced into Spain in 1977, and, at present, the breed is well developed with a number of dairy recording schemes having recently been established (Jiménez and Jurado, 2005; Gutiérrez et al., 2007).

The expansion of the Assaf sheep breed has occurred basically by the male-mediated absorption of Spanish native dairy sheep breeds such as Castellana, Churra, Manchega or, to a lesser extent, Latxa (Ugarte et al., 2002), i.e., continuous crossing of the Spanish breeds and their crosses to Assaf rams. Recent studies (Pedrosa et al., 2007) did not find differences at the mitochondrial DNA level among the Spanish Assaf and the Spanish native dairy sheep breeds, thus supporting the hypothesis of a male-mediated absorption of native sheep during the formation of the Spanish Assaf breed.

0921-4488/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.smallrumres.2008.09.001
Admixture, intercrossing among subpopulations or breeds, is a common occurrence within livestock species and affects breeding for production because of heterosis and also elicits growing scientific interest, for example, with regards to the potential efficiency of admixture mapping of important traits (McKeigue, 2005). The aim of this paper is to assess the genetic relationships among the Spanish Assaf (Assaf.E) and major native Spanish dairy breeds using microsatellites to contribute to the knowledge of the formation and within-population genetic variability of the breed.

2. Materials and methods

2.1. Samples

Blood samples were obtained from 44 unrelated Assaf.E individuals from 23 different Assaf.E flocks spread throughout six different Spanish provinces (Zamora, Madrid, Segovia, Toledo, Salamanca, and Valladolid). A maximum of 2 individuals per flock were sampled. Samples from 32 individuals belonging to the Assaf breed and 14 from the Milchschaf breed were also obtained in order to account for the original genetic background of the Assaf sheep. Additionally, a total of 234 unrelated individuals corresponding to six native Spanish dairy sheep breeds were obtained: breeds and number of individuals sampled (in brackets) were: Castellana (59; including 29 White-Castellana and 30 Black-Castellana individuals), Churra (37), Latxa (67; including 33 Blonde-faced and 34 Black-faced Latxa individuals), Manchega (33) and Rubia de El Molar (39). The Castellana and Churra breeds are located in Northern Castile; the Manchega breed being the major native dairy sheep in Southern Castile; Rubia de El Molar sheep located in Madrid; and Latxa the main dairy sheep of the Basque country and Cantabrian range. Black Castellana and Rubia de El Molar individuals are expected to characterise the genetic background of endangered populations that could have been, at least partially, absorbed by foreign sheep breeds. Finally, samples from 32 Merino individuals were obtained to be used as the outgroup. A total of 356 samples were analysed.

Total DNA was isolated from blood samples following standard procedures (Sambrook et al., 1989). Up to 14 microsatellites (BM8125, BM6526, CP34, BM7575, INRA006, BM6606, BM1818, FCB128, CSSM31, CSMM66, ILTST011, McM53, RM006, ILTST005) previously used to characterise both between- (Álvarez et al., 2004, 2005) and within-breed (Álvarez et al., 2007, 2008) genetic relationships in sheep were analysed in all the individuals. The PCR products were labelled with a fluorescent method (Cy5 labelled primer) and genotyping was performed on an ALFExpress II automated sequencer (Amersham Biosciences, Barcelona).

2.2. Statistical analysis

Molecular information was analysed using the program MolKIN v2.0 (Gutiérrez et al., 2005). The following parameters were computed at the breed level: observed (Hg) and expected (He) heterozygosity and number of alleles per locus (A) corrected using Hurlbert’s rarefaction method (1971) as

\[
A[g] = \sum_{k=0}^{g-1} \left[ 1 - \frac{N!}{(N-k)!N^k} \right]^{-1},
\]

where g is the specified sampled size, N the number of gene copies examined in a given locus (N×g), and N the number of occurrences of the i-th allele among the N sampled gene copies to account for sample size. Here, g was fitted to 24, which is twice the minimum number of individuals within a cohort with genotype known for all the microsatellites.

Additionally, the within- and between-individuals molecular coancestry matrixes were computed. The molecular coancestry (fij) between two individuals i and j is the probability that two randomly sampled alleles from the same locus in two individuals are identical by state (Caballero and Toro, 2002). Molecular coancestry between two individuals i and j(fij) is

\[
f_{ij} = \frac{1}{4[L_i + L_j - 1] + f_{ij}},
\]

where L is 1 when allele x on locus i in individual i and allele y in the same locus in individual j are identical, and zero otherwise. Notice that this value can only have four values: 0, 1/4, 1/2 and 1. The molecular coancestry between two individuals i and j(fij[mj]) can be obtained by simply averaging over L analysed loci as

\[
f_{ij[mj]} = \frac{1}{L} \sum_{l=1}^{L} f_{ij[l]/L}.
\]

Moreover, the between-individual kinship distance (Dk) matrix was computed as

\[
D_k = \frac{1}{2} - \frac{f_{ij}}{k},
\]

where j is the molecular self-coancestry for individual i, computed as

\[
f_{ij} = \frac{1}{2} \left[ 1 - \frac{\sum_{x=1}^{2} f_{ix} f_{jx}}{2} \right],
\]

where fij is the homozygosity for individual i and fij the molecular coancestry between individuals i and j. Within- and between-population molecular coancestry and Dk were computed by simply averaging the corresponding values for all the within- or between-population pairs of individuals.

To avoid bias because of unequal sample sizes, particularly for the Milchschaf breed, the bootstrapping method recommended by Simianer (2002) and Baumung et al. (2006) was applied using 100 samples with exactly 30 individuals per breed.

When necessary for descriptive purposes multidimensional scaling analysis was carried out on the genetic distance matrices using the Proc MDS of SAS/STAT™ (1999). Multidimensional scaling analysis is an exploratory technique that allows the visualization of proximities in a low dimensional space. The interpretation of the dimensions obtained from the analysis can lead to an understanding of the processes underlying the perceived nearness of entities.

With the aim of ascertaining the recent history of the Assaf.E breed, probabilities of assignment of the Assaf.E individuals to native Spanish dairy sheep breeds were computed using the Bayesian method inspired from Rannala and Mountain (1997) as implemented in the program GENECLASS 2.0 (Cornuet et al., 1999; Piry et al., 2004). This Bayesian procedure computes the likelihood of a genotype in a given population assuming an equal prior probability density to the allelic frequencies of each locus in each population. This method shows better assignment performance than frequentist methods or distance-based methods both in simulated or real populations (Cornuet et al., 1999; Arranz et al., 2001). Moreover, using the program GENECLASS 2.0 (Piry et al., 2004), detection of first generation migrants from the native Spanish dairy sheep breeds into the Assaf.E breeds was carried out employing the methodology proposed by Paetkau et al. (2004) which seeks the maximization of the ratio of the likelihood computed from the population where the individual was sampled over the highest likelihood value among all population samples including the population where the individual was sampled.

The program STRUCTURE (Pritchard et al., 2000) was used to ascertain a possible cryptic genetic structure in the analysed dataset. The program estimates, using the Markov Chain Monte Carlo method, the natural log of the probability of a given genotype X is part of a given population K (In Pr(X|K)). This ensures that the groups are, as representative as possible, samples from a simple population. As the implemented algorithm uncovers ‘hidden structure’ without using a priori knowledge about the number of clusters (populations or breeds) present in a dataset, we carried out 10 different runs from K = 1 to K = 15 to identify the most likely number of clusters present in the dataset (the most likely parameter K). All runs used a burn-in period of 100,000 iterations and a period of data collection of 100,000 iterations.

Results from the programs GENECLASS 2.0 (Piry et al., 2004) and STRUCTURE (Pritchard et al., 2000) were summarised computing the similarity between the Assaf.E breed and the other breeds in the analysed dataset. Similarity was computed on: (a) the probabilities of assignment of each Assaf.E individual in the dataset to each of the other breeds in the dataset, including the Assaf.E breed itself, obtained from GENECLASS 2.0; and (b) the probabilities of assignment of each Assaf.E individual in the dataset to each of the clusters inferred by STRUCTURE for the most likely value of K. Similarity was computed as

\[
\theta = \sum_{i=1}^{N} \sqrt{\frac{q_{ik}}{N}},
\]

(see Nei, 1987 and Söelkner et al., 1998, as examples) where θ is the cosine of the angle between the vectors of the square roots of the assignment probabilities of a given individual i into the breeds or clusters A(qik) or B(qik). Since the assignment probabilities obtained from GENECLASS 2.0 are computed at the individual level, similarity was computed as

\[
\theta = \sum_{i=1}^{N} \sqrt{\frac{q_{ik}}{N}},
\]

where N is the 44 Assaf.E individuals sin the dataset.
Table 1

<table>
<thead>
<tr>
<th>Breed</th>
<th>N</th>
<th>He</th>
<th>fii</th>
<th>FIS</th>
<th>A</th>
<th>A(24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assaf.E</td>
<td>44</td>
<td>0.670</td>
<td>0.332</td>
<td>0.041</td>
<td>7.0</td>
<td>5.2</td>
</tr>
<tr>
<td>Awassi</td>
<td>32</td>
<td>0.741</td>
<td>0.307</td>
<td>-0.025</td>
<td>7.1</td>
<td>5.8</td>
</tr>
<tr>
<td>Castellana</td>
<td>59</td>
<td>0.796</td>
<td>0.265</td>
<td>0.047</td>
<td>9.5</td>
<td>6.8</td>
</tr>
<tr>
<td>Churra</td>
<td>37</td>
<td>0.761</td>
<td>0.297</td>
<td>0.060</td>
<td>8.3</td>
<td>6.5</td>
</tr>
<tr>
<td>Latxa</td>
<td>67</td>
<td>0.764</td>
<td>0.290</td>
<td>0.055</td>
<td>8.6</td>
<td>6.1</td>
</tr>
<tr>
<td>Manchega</td>
<td>33</td>
<td>0.774</td>
<td>0.275</td>
<td>0.101</td>
<td>8.7</td>
<td>6.7</td>
</tr>
<tr>
<td>Milchschaf</td>
<td>32</td>
<td>0.789</td>
<td>0.268</td>
<td>0.081</td>
<td>8.3</td>
<td>6.7</td>
</tr>
<tr>
<td>Rubia de El Molar</td>
<td>39</td>
<td>0.706</td>
<td>0.340</td>
<td>0.051</td>
<td>6.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Total</td>
<td>357</td>
<td>0.762</td>
<td>0.240</td>
<td>0.129</td>
<td>12.2</td>
<td>6.8</td>
</tr>
</tbody>
</table>

3. Results

Parameters characterising genetic variability of the analysed sheep breeds are given in Table 1. Closer to the Milchschaf individuals, the Assaf.E breed showed the lowest values for expected heterozygosity (0.670) and rarefied average number of alleles per locus (5.2) and the highest values of within-breed molecular coancestry (0.332). In general, the other analysed sheep breeds showed higher diversity (He higher than 0.7), even though breeds such as the endangered Rubia de El Molar (Álvarez et al., 2004, 2005) showed high values of fii (0.340) and A(24) (5.0).

The between-breeds molecular coancestry (fij) and kinship distance (Dk) matrices are given in Table 2. As expected, a lower molecular coancestry value was found not only between the Assaf.E breed and the Merino breed (0.204), which was used as the outgroup, but also for the pair Assaf.E–Latxa (0.207) showing that there exists little genetic identity between these two breeds. The Assaf.E breed had the highest fij values with the Churra breed (0.250), those computed with the Milchschaf and Rubia de El Molar breeds (0.234 and 0.236, respectively), also being high. The higher genetic differentiation of the Assaf.E breed, assessed using Dk, were found in the Merino and Milchschaf breeds (above 0.46) while the lower were found in the Churra and Awassi breeds (below 0.44). Consistent with Álvarez et al. (2004, 2005), the lowest differentiation assessed in the dataset was for the pair Castellana–Latxa (0.408). This scenario is illustrated in Fig. 1 which shows the bidimensional scaling plot constructed using the between-breeds Dk matrix. In dimension 1 (X-axis) the Milchschaf breeds is separated from the other analysed sheep breeds while the Assaf.E and the foreign breeds (Awassi and Milchschaf) are differentiated with respect to most of the Spanish native sheep (except for Churra).

Assignment of the individuals to the corresponding breeds was assessed using the program GENECLASS 2.0 under the algorithm by Rannala and Mountain (1997). Overall, there were large assignment errors (18.5%) indirectly showing low between-breeds differentiation in the available dataset. Forty out of 44 Assaf.E individuals were correctly assigned to their breed. Two of the individuals incorrectly assigned were considered most likely to be Manchega individuals, and the other 2 were assigned to the Castellana and Awassi breeds, respectively. Additionally, the same 4 incorrectly assigned individuals were detected as first generation migrants from the same sheep breeds into the Assaf.E breed using the methodology described by Paetkau et al. (2004).

Population structure and degree of admixture were assessed using STRUCTURE. The most likely number of clus-

Table 2

<table>
<thead>
<tr>
<th>Breed</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assaf.E</td>
<td>0.437</td>
<td>0.442</td>
<td>0.424</td>
<td>0.464</td>
<td>0.456</td>
<td>0.457</td>
<td>0.476</td>
<td>0.447</td>
<td>0.447</td>
</tr>
<tr>
<td>Awassi</td>
<td>0.225</td>
<td>0.426</td>
<td>0.450</td>
<td>0.430</td>
<td>0.431</td>
<td>0.440</td>
<td>0.464</td>
<td>0.449</td>
<td>0.449</td>
</tr>
<tr>
<td>Castellana</td>
<td>0.223</td>
<td>0.221</td>
<td>0.438</td>
<td>0.408</td>
<td>0.418</td>
<td>0.410</td>
<td>0.449</td>
<td>0.416</td>
<td>0.416</td>
</tr>
<tr>
<td>Churra</td>
<td>0.250</td>
<td>0.207</td>
<td>0.221</td>
<td>0.453</td>
<td>0.445</td>
<td>0.451</td>
<td>0.487</td>
<td>0.431</td>
<td>0.431</td>
</tr>
<tr>
<td>Latxa</td>
<td>0.208</td>
<td>0.224</td>
<td>0.250</td>
<td>0.214</td>
<td>0.431</td>
<td>0.417</td>
<td>0.451</td>
<td>0.431</td>
<td>0.431</td>
</tr>
<tr>
<td>Manchega</td>
<td>0.221</td>
<td>0.228</td>
<td>0.243</td>
<td>0.227</td>
<td>0.238</td>
<td>0.438</td>
<td>0.486</td>
<td>0.431</td>
<td>0.431</td>
</tr>
<tr>
<td>Merina</td>
<td>0.204</td>
<td>0.214</td>
<td>0.246</td>
<td>0.216</td>
<td>0.247</td>
<td>0.231</td>
<td>0.477</td>
<td>0.438</td>
<td>0.438</td>
</tr>
<tr>
<td>Milchschaf</td>
<td>0.234</td>
<td>0.229</td>
<td>0.247</td>
<td>0.218</td>
<td>0.251</td>
<td>0.222</td>
<td>0.226</td>
<td>0.489</td>
<td>0.489</td>
</tr>
<tr>
<td>Rubia de El Molar</td>
<td>0.236</td>
<td>0.217</td>
<td>0.252</td>
<td>0.247</td>
<td>0.244</td>
<td>0.249</td>
<td>0.237</td>
<td>0.225</td>
<td>0.225</td>
</tr>
</tbody>
</table>
Table 3
Number of individuals per breed (N) and proportion of membership of each of the analysed sheep breeds in each of the 11 most likely clusters inferred using the program STRUCTURE

<table>
<thead>
<tr>
<th>Breed</th>
<th>Inferred clusters</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Assaf.E</td>
<td>0.019</td>
<td>0.055</td>
</tr>
<tr>
<td>Awassi</td>
<td>0.014</td>
<td>0.016</td>
</tr>
<tr>
<td>Castellana</td>
<td>0.271</td>
<td>0.168</td>
</tr>
<tr>
<td>Churra</td>
<td>0.032</td>
<td>0.048</td>
</tr>
<tr>
<td>Latxa</td>
<td>0.056</td>
<td>0.047</td>
</tr>
<tr>
<td>Manchega</td>
<td>0.193</td>
<td>0.373</td>
</tr>
<tr>
<td>Merina</td>
<td>0.094</td>
<td>0.118</td>
</tr>
<tr>
<td>Milchschaf</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>Rubia de El Molar</td>
<td>0.016</td>
<td>0.024</td>
</tr>
</tbody>
</table>

The program STRUCTURE was used to infer the most likely number of clusters (K) and the membership of each individual to each cluster. The most likely K was 11, and this was confirmed using 10 different runs of the program. The highest similarity values were assessed, regardless of the source of information, between the Assaf.E breed and both the Castellana and the Manchega breeds.

4. Discussion

Overall, the Assaf.E breed has a well-differentiated genetic background. Pedrosa et al. (2007) were unable to assess differences, at the maternal level, between the mitochondrial DNA polymorphism found in the Spanish Assaf sheep and that of the Spanish native dairy sheep breeds. The information by Pedrosa et al. (2007) would support the well-known fact that the formation of the Assaf.E breed was carried out via a male-mediated absorption of native sheep (Gutiérrez et al., 2007). In any case, the absorption of the Spanish native sheep may have happened using a limited number of (probably related) Assaf rams thus leading to a significant differentiation of the Assaf.E individuals with respect to the other breeds in the dataset but also to a low within-breed genetic variability. In a context of low between-breed differentiation (see results from the programs GENECLAS 2.0 and STRUCTURE), the Assaf.E breed is in the low range of the values obtained for expected heterozygosity and rarefacted average number of alleles per locus while the between-Assaf.E individuals genetic identity assessed using molecular coancestry \( f_{ii} \) is high (Table 1) and that roughly 80% of the Assaf.E individuals genotyped were assigned to a single cluster after the analyses carried out with the program STRUCTURE (Table 3).

With respect to the genetic relationship between the Assaf.E breed and the others in the dataset, we obtained complimentary information. The \( D_0 \) matrix allowed us to assess the relationship between the Assaf.E breed and their 'original' source populations: Awassi and Milchschaf (see Fig. 1). The \( D_0 \) distance is appropriate for assessing relationships between co-extant breeds with different expected genetic origins (Álvarez et al., 2005). The formula

![Fig. 2. Ln Pr(G|K) values of 10 different runs of the program STRUCTURE for each K value varying from 1 to 15.](image)
used to compute the kinship distance \( D_{ki} = \frac{(s_i + s_j)}{2} - f_{ij}; \) Caballero and Toro, 2002) has two terms that summarise the recent between-population differentiation (term \( s_i + s_j/2 \)) corrected by the genetic identity of the populations at the moment of the original (ancestral) metapopulation fission \( f_{ij}; \) Eding and Mieuwissen, 2001). Related to that, the between coancestry matrix informed us that the highest gene flow from a Spanish native dairy sheep breed into the Assaf.E breed was from the Churra breed. In view of the assignment analyses carried out using the programs GENELIAS 2.0 and STRUCTURE, this result is not easy to interpret. The high coancestry value between the pair Assaf.E-Churra can be interpreted as a reflection of the gene flow from the Churra to the Assaf populations which occurred during the formation of the Assaf.E breed. If so, the Churra breed would be the most important source of native Spanish sheep genes used in the formation of the Assaf.E breed. However, no clear introgression of Churra individuals into the Assaf.E breed could be assessed using the assignment procedures, thus telling us that, if it did happen, this introgression did not occur recently, i.e., in at least the last two generations. Moreover, Álvarez et al. (2004, 2005) state that the Churra breed had a particular genetic background with respect to the major sheep breeds analysed here. In this respect, the ‘association’ between the Assaf.E and the Churra breeds assessed using molecular coancestry-based parameters can also be explained by the grouping of ‘rare’ breeds into the Spanish sheep context rather than a ‘real’ genetic relationship.

At least during the most recent period, the Assaf.E breed seems to have been formed via the absorption of Manchega and Castellana individuals. The Manchega breed is the main native dairy sheep in Southern Castile which is also going through a large introgression of the Lacaune breed. Furthermore, until recently, the Castellana breed has not had a well-established breeders association and performance-recording organization thus making it an obvious target for the formation of the Assaf.E flocks in Northern Castile. Note that both the Manchega and Castellana breeds are classified as belonging to the same type of Spanish native sheep, the so-called Entrefino type. As expected, the relationship between the Latxa and the Assaf.E breeds is low due to the low number of Assaf.E flocks existing in the Basque Country and Navarre which are the main areas of spreading of the Latxa sheep breed.

Summarising the information obtained here, we can suggest that the formation of the Assaf.E breed was basically done via the absorption of individuals belonging to the Entrefino type, particularly to the Castellana and Manchega populations. A significant participation of Churra individuals in the formation of the Assaf.E breed is less clear and, if it did happen, it probably occurred at an early point in the formation of the Assaf.E breed. Moreover, formation of the Assaf.E breed was carried out using a limited number of Assaf rams thus providing a relatively low genetic variability, at the microsatellite level, in the breed.

Acknowledgements

This work was partially funded by a grant from MECINIA, No. RZ2004-00007-C02. Authors are indebted to Henner Simianer and Roswitha Baumung for providing the FORTRAN routines for bootstrapping equalising for sampling size that have been included in the current version of the program MolKin.

References


Prichard, J., Stephens, M., Donnelly, P., 2000. Inference of population struc-

