Genetic database development for the characterisation of Sicilian sheep population

Viviana Giangreco, Tiziana Lupo, Ignazio Sammarco, Maurizio Bivona, Gabriele Ciaccio, Luca Sineo & Stefano Reale*

Istituto Zooprofilattico Sperimentale della Sicilia “A. Mirri” Via Gino Marinuzzi 3, 90129 Palermo, Italy
*Corresponding author, e-mail: stefano.reale@izssicilia.it

ABSTRACT

The most representative sheep in Sicily are Belicina, Comisana, Pinzirita, Barbaresca and the crossbred derived sheep from all this species. In this study, the allelic frequencies of the Sicilian sheep population were investigated. It currently represents the best way to determine the genetic identity and/or family even with limited amounts of sample or when the DNA is degraded. The aim of the study was to provide a reference data bank and to evaluate a microsatellite panel for pedigree analysis as suggested by the International Society for Animal Genetics (ISAG). There are various studies on European sheep, but few datasets were developed on the population of Sicilian sheep. The reference database will include allele frequencies at each locus and will determine genetic parameters for Sicilian ovine species selection. Our results indicated that Hardy Weinberg equilibrium was not always maintained. These results could be explained by a non-random mating. The database is useful to investigate the relationship, the parentage the meat traceability and in disease control programs. The standardized panels of allele frequencies represent a molecular fingerprinting characterizing the subjects with very high definition level and can be useful to control all the livestock. The parentage identification could be important for the veterinary police to investigate the theft or the animal substitutions in the Sicilian farms.

KEY WORDS

Ovine; microsatellite; locus; allele.

INTRODUCTION

The most representative species of sheep in Sicily are Belicina, Comisana, Pinzirita and Barbaresca. The purpose of this study was to investigate the allelic frequencies in the Sicilian sheep population to provide a reference data bank and to evaluate the use of ISAG (International Society for Animal Genetics) through a microsatellite panel for pedigree analysis. Tandem repeats are used as an effective method to track DNA markers in genotyping field. The deriving database can be useful for the traceability of meat, risk assessment and consumer warranty. The microsatellites are employed as genetic markers for their random distribution, the codominant inheritance (Barbarà et al., 2007), discriminative power and possibility of simultaneous analysis. The microsatellite markers can also be used to trace the meat through the production processes (Vázquez et al., 2004), to study the ge-
MATERIAL AND METHODS

We tested 10 microsatellite markers recommended by ISAG on 452 Sicilian sheep. The microsatellite loci were employed in two homogeneous multiplex group of loci, we use ten ISAG loci for present study: OarFCB011, INRA0063, HSC, OarCP0049, OarFCB0304, CSRD0247, OarFCB020, D5S2, SPS0113, INRA005. Whole blood samples were taken from the Sicilian typical half-breed, species representative of the Sicilian population. Genomic DNA was extracted and purified using a commercial kit (Ezna WVR). Two different PCR test were employed to investigate a panel of ten microsatellites. DNA targets were amplified in a 6-plex and 4-plex PCRs system respectively as follows: 12.5 µl Type-it 2X master mix (Qiagen), 2.5 µl Primer mix (2 µM for each primer), 20 ng DNA. PCRs were carried out using a thermocycler (9700 Applied Biosystems, San Diego, CA, USA). Multiplex-PCR products were analyzed using an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems). Genotypic profiles

Figure 1. Allele frequency observed at each locus for every examined samples. (GenAlEx v6.5 software).
were read and analyzed by GeneMapper software v4.0 (Applied Biosystems). Statistical analysis of data generated from the 10 markers that were analyzed was performed using GenAlEx (Peakall & Smouse, 2012), PowerMarker (Liu & Muse, 2005), and Micro-Checker (Van Oosterhout et al., 2004) softwares. All these are a useful set of bioinformatic tools specific for genetic populations analysis implementing various data management algorithms. The number of effective alleles (Ne), allele number (Na), allele frequency, observed and expected heterozygosities (Ho and He, respectively), probability of Identity (Pi), defined as the probability that two individuals drawn at random from a population will have the same genotype at multiple loci, probability of exclusion of a locus (Pe), and the devia-

<table>
<thead>
<tr>
<th>Locus</th>
<th>N</th>
<th>NA</th>
<th>NE</th>
<th>NE/NA</th>
<th>MAF</th>
<th>Ho</th>
<th>He</th>
<th>Fis</th>
<th>PIC</th>
<th>Pe</th>
<th>Pi</th>
</tr>
</thead>
<tbody>
<tr>
<td>OarF8011</td>
<td>452</td>
<td>22</td>
<td>8.42</td>
<td>0.383</td>
<td>0.187</td>
<td>0.771</td>
<td>0.881</td>
<td>0.007</td>
<td>0.853</td>
<td>0.80</td>
<td>0.033</td>
</tr>
<tr>
<td>INRA0063</td>
<td>452</td>
<td>31</td>
<td>13.40</td>
<td>0.433</td>
<td>0.138</td>
<td>0.772</td>
<td>0.925</td>
<td>0.090</td>
<td>0.914</td>
<td>0.85</td>
<td>0.023</td>
</tr>
<tr>
<td>HSC</td>
<td>452</td>
<td>20</td>
<td>9.75</td>
<td>0.488</td>
<td>0.188</td>
<td>0.783</td>
<td>0.897</td>
<td>0.024</td>
<td>0.876</td>
<td>0.81</td>
<td>0.022</td>
</tr>
<tr>
<td>OarCP049</td>
<td>452</td>
<td>36</td>
<td>12.51</td>
<td>0.347</td>
<td>0.161</td>
<td>0.763</td>
<td>0.922</td>
<td>0.091</td>
<td>0.907</td>
<td>0.89</td>
<td>0.014</td>
</tr>
<tr>
<td>OarF8034</td>
<td>452</td>
<td>28</td>
<td>5.54</td>
<td>0.191</td>
<td>0.345</td>
<td>0.651</td>
<td>0.819</td>
<td>0.024</td>
<td>0.754</td>
<td>0.83</td>
<td>0.063</td>
</tr>
<tr>
<td>CSR0247</td>
<td>452</td>
<td>37</td>
<td>9.68</td>
<td>0.262</td>
<td>0.197</td>
<td>0.804</td>
<td>0.897</td>
<td>0.001</td>
<td>0.875</td>
<td>0.87</td>
<td>0.024</td>
</tr>
<tr>
<td>OarF8020</td>
<td>452</td>
<td>24</td>
<td>8.79</td>
<td>0.366</td>
<td>0.187</td>
<td>0.758</td>
<td>0.886</td>
<td>0.030</td>
<td>0.860</td>
<td>0.82</td>
<td>0.028</td>
</tr>
<tr>
<td>DSS2</td>
<td>452</td>
<td>13</td>
<td>4.14</td>
<td>0.319</td>
<td>0.329</td>
<td>0.614</td>
<td>0.759</td>
<td>-0.050</td>
<td>0.642</td>
<td>0.68</td>
<td>0.095</td>
</tr>
<tr>
<td>SP50113</td>
<td>452</td>
<td>18</td>
<td>7.11</td>
<td>0.395</td>
<td>0.235</td>
<td>0.642</td>
<td>0.859</td>
<td>0.112</td>
<td>0.820</td>
<td>0.77</td>
<td>0.034</td>
</tr>
<tr>
<td>INRA005</td>
<td>452</td>
<td>30</td>
<td>12.22</td>
<td>0.407</td>
<td>0.130</td>
<td>0.681</td>
<td>0.918</td>
<td>0.176</td>
<td>0.905</td>
<td>0.82</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Table 2. Genetic parameters in Sicilian sheep population. NA = number of alleles; NE = number of effective alleles; MAF = minor allele frequency; Ho = observed heterozygosity; He = expected heterozygosity; Fis = inbreeding coefficient; PIC = polymorphic information content; Pe = probability of exclusion; Pi = probability of identity.
tion probability from the Hardy-Weinberg equilibrium (HWE) were obtained using the GenAlEx v6.5 software. The expected homozygosity (HomE), observed homozygosity (HomO), homozygosity excess, evidence for null allele, evidence for large allele dropout, and evidence for scoring error due to stuttering were obtained using the MicroChecker v2.2 software. The polymorphism information content (PIC), inbreeding coefficient, genotype number (NG), major allele frequency (MAF), and major genotype frequency (MGF) were investigated using the PowerMarker v3.25 software.

RESULTS

In total, 452 related and unrelated sheep were genotyped, the average observed heterozygosity was lower than the expected value (0.701 vs 0.852). The exact test for Hardy-Weinberg proportion, allele number and inbreeding coefficient were calculated. These results could be explained by a non-random mating studies on a larger number of samples. The polymorphic information content (PIC) calculated according Botstein et al. (1980) ranged from 0.642 for locus D5S2 to 0.914 for locus INRA0063. PIC is a parameter that indicates the degree of marker informativeness describing genotypic variation in single base pair or in larger sequence repeats. The PIC value will be almost zero if there is no allelic variation. All locus were informative (PIC > 0.5) (Table 1), with a mean PIC of 0.841. The heterozygosity expected HE and observed HO, as measures of genetic diversity at a single locus, are shown in Table 1. In all cases, HO was lower than HE.

DISCUSSION AND CONCLUSIONS

To establish a livestock conservation program it is fundamental the genetic characterization of the entire population under study. For different breeds must have taken into account also phenotypic differences (morphology, milk production, disease resistance etc) and information about provenience of samples. A database of frequencies for the different alleles of known microsatellite markers it’s important to help researcher in studying the phylogeny of one or more populations, to discover patterns of relationships among different groups or associate the genetic markers with important productive characteristics. Analysis of different samples of sheep resulted in a set of genotype profiles of the most representative ovine populations from Sicily. The principal statistics parameter of sheep population were obtained elaborating microsatellites alleles frequencies through a set of statistical analysis tools in particular Genalex and PowerMarker software and Micro-Checker. The final data show a significant deficiency in the HO value compared to the HE value. This deviation from HWE can be caused by inbreeding, assortative mating or Wahlund effect, due to a fragmentation of the original population into subpopulations. But loss of heterozygosity can also include genotyping errors due to nonamplified alleles (null alleles) caused by mutations in primer binding site (Pemberton et al., 1995), short allele dominance (large allele dropout) and the scoring of stutter peaks dropout. Finally, we observed an homozygote excess in all loci compared to the expected value. Five loci (OarFCB0304, CSRD0247, D5S2, SPS0113, INRA005) showed a strong evidence for scoring error due to stuttering, and none of these loci had large allele dropout. It is not clear whether the homozygote excess is due to null alleles or if it really reflects the genotypes of the Sicilian sheep population.

REFERENCES

Barendse W., Armitage S.M., Kossarek L.M., Shalom A., Kirkpatrick B.W., Ryan A.M., Clayton D., Li L.,


