Proceedings, 10th World Congress of Genetics Applied to Livestock Production

New Parentage Testing SNP Panel for Commercial Breeds will be a Useful Tool for Conservation of Creole Sheep

F. Macedo[1], *E. A. Navajas*[2], I. Aguilar[2], A. N. Grasso[2], F. Pieruccioni[2], and G. Ciappesoni[2]
[1]Universidad de la República, Uruguay [2]Instituto Nacional de Investigación Agropecuaria, Uruguay

ABSTRACT: Accurate and cost-effective DNA-based parentage assignment tools are relevant for genetic improvement and conservation programs. Parentage-testing SNPs selected in commercial sheep breeds were evaluated in Creole sheep in Uruguay using information provided by the 600k SNP chip. Genotyping results indicated a high proportion of fixed and rare SNPs (21.98%; MAF<=0.001), and a relatively low proportion of highly polymorphic SNPs (25.86%; MAF>=0.3). Nevertheless, subsets of 72 and 93 SNPs of the most informative markers for parentage testing, reached probability exclusions of 0.9999 and 0.99999, with average MAF values of 0.43. Comparison of the genotypes of 34 ewe-lamb pairs showed that both subsets of SNPs successfully found correct/incorrect pairs. Although these are preliminary results they show the feasibility of using these SNPs for parentage verification in Creole sheep. This will help to correctly record the genealogy of the population in the future conservation plan.
Keywords: Creole sheep; Parentage testing; Validation

Introduction

DNA-based parentage assignment is usually emphasized from a genetic improvement point of view to minimize genetic gain losses due to pedigree errors. Although both microsatellite and single-nucleotide polymorphism (SNP) parentage assignment in cattle are commercially available, SNP panels for sheep are still being investigated (Heaton et al. (2014)). The National Agricultural Research Institute of Uruguay is developing a parentage-testing SNP panel for sheep based on sequencing and 50k SNP genotyping data of Corriedale and Merino breeds (Macedo et al. (2014)). A subset of 319 SNPs was preselected including those described by Kijas et al. (2012), and it is now on validation in breeding flocks of commercial breeds.

Genetic diversity decline of the world's livestock species was addressed internationally in the Global Plan of Action for Animal Genetic Resources, which was adopted at the International Technical Conference on Animal Genetic Resources for Food and Agriculture in 2007. The implementation of conservation programs is one of four strategic priority areas (FAO (2013)). DNA-based parentage tests are also useful tools for in situ animal genetic conservation programs by helping pedigree reconstruction and sire identification (Allendorf et al. (2010)).

Creole sheep in Uruguay are being phenotypically and genetically characterized using genomic information. Because of the risk of extinction of this population, a conservation program is being designed simultaneously. The availability of a cost-effective test for parentage assignment would be an important contribution given the difficulties to have accurate parentage recording in the harsh environment and extensive production system where the Creole sheep are situated.

The objective of this paper was to evaluate the feasibility of the parentage testing SNPs in Creole sheep, using SNPs preselected in commercial breeds.

Materials and Methods

Data. Genotyping data from 606,006 SNPs (HD chip, Anderson et al., 2014) of 158 Creole sheep were used in this study, including 21 rams, 60 ewes and 77 females and male lambs. The DNA samples were of 75% of the largest Creole flock in Uruguay and main reservoir of the population, which is located in a National Protected Area.

A total of 278 SNPs, out of the 319 preselected SNPs for parentage testing (Macedo et al. (2014)), were within the HD chip. They were identified based on chromosome, position and name using the Sheep Genome Assembly OARv2.0

Although historical pedigree information is not available in the Creole sheep, at the time of starting the comprehensive characterization that is being carried out, all animals were tagged with unique identification, and mother-progeny pairs are regularly recorded in the field by visual matching since then. This proxy method is one of the cost-effective tools that may help keeping dam pedigree records. The mother-progeny relationship was evaluated with the HD chip information, and the results used for a first preliminary validation of the parentage testing panel.

Data analysis. The minor allele frequency (MAF) for each SNP was calculated using the genotypic data available. In addition, deviation from Hardy-Weinberg equilibrium was also tested using Genepop v.4.2.2 (Rousset (2008)).

Based on the allele frequencies of the SNP subset for parentage testing, exclusion probabilities by marker (Pk) and combined exclusion probability (P) were calculated as described by Jamieson and Taylor (1997):

$$P=1-(1-P1)(1-P2)...(1-Pk)$$

where k is the total number of markers, and

$$P_K = 1 - 4\sum_{i=1}^{n} p_i^{\,2} + 2\left(\sum_{i=1}^{n} p_i^{\,2}\right)^2 + 4\sum_{i=1}^{n} p_i^{\,3} - 3\sum_{i=1}^{n} p_i^{\,4}$$

where p_i is frequency of allele i and n the number of alleles for each marker.

To compute the combined exclusion probability the markers were arranged by MAF in decreasing order, so the most informative SNPs were combined first. Two panels of SNPs were selected to be tested based on a combined probability exclusion of 0.9999 and 0.99999.

Parentage verification was performed with the HD chip and selected SNP for parentage testing. Rules for assignment were based on the count of Mendelian conflicts, number of opposites homozygous genotypes, between pair of animals (e.g. Hayes (2011)). Genotypes of parents and progeny were compared. A pair parent-progeny with 2% or lower number of conflicts was assigned.

Results and Discussion

Polymorphism and homozygosis. Distribution of MAF values presented in Figure 1 shows that 13 SNPs were fixed (MAF=0) and 5 SNPs were rare (MAF<=0.01). Nevertheless, a significant proportion of SNPs in this subset for parentage testing were polymorphic, with a mean MAF of 0.25. More than 43.88% of the SNPs were highly polymorphic (MAF>=0.3). Only 6 SNPs were not in Hardy-Weinberg equilibrium, suggesting that most of the preselected SNPs are not under selection.

The significant proportion of highly informative SNP in the subset is particularly interesting given that HD genotyping results showed a very high proportion of fixed and rare SNP (21.98%; MAF<=0.001), and a relatively low proportion of highly polymorphic SNPs (25.86%; MAF>=0.3) in the Creole sheep (Figure 2). A previous study by Grasso et al. (2014), based on the 50k SNP panel data of 10 Creole rams, reported 36% of highly polymorphic SNPs compared to approximately 50% in Corriedale and Merino. A high proportion of homozygosis was expected in Creole sheep because of the small effective population size (Allendorf et al. (2010)). In addition, other unknown events in the history of this Creole population, such as any severe population bottleneck, could have also contributed to a low percentage of polymorphic SNPs. Nevertheless, the low MAF values may also be due to the fact that the Creole sheep was not used in the ascertainment of SNPs on the HD chip.

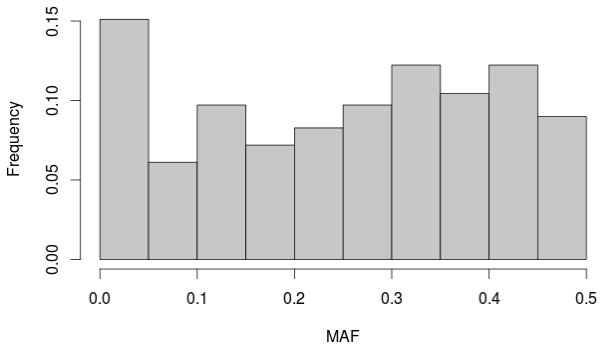

Figure 1: MAF distribution of 278 SNPs of parentage testing panel in Creole sheep.

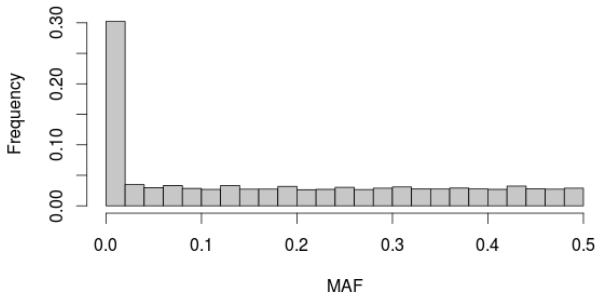

Figure 2: MAF distribution of 600k SNPs of the Creole sheep largest flock (158 animals).

Exclusion probabilities. Besides the low genetic variability in the Creole population, very high cumulative exclusion probabilities can be achieved with the parentage-testing SNPs using a relatively low number of markers. As indicated in Figure 3, probability values of 0.999 and 0.9999 were reached with 72 and 93 SNPs, respectively. MAF averages were 0.44 and 0.42 with 72 and 93 SNPs.

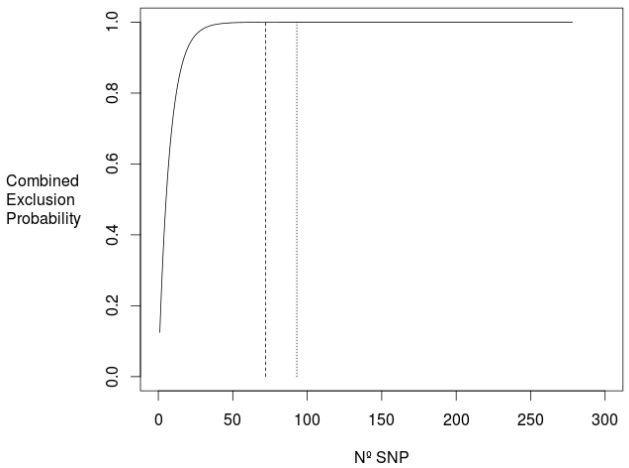

Figure 3: Combined Exclusion Probability. Probabilities of .9999 and .99999 are achieved with 72 (vertical dotted line) and 93 SNPs (vertical solid line), respectively.

Exclusion probabilities and MAF values are comparable with other panels reported in outbred populations that have higher levels of heterozygosis, as well as mean MAF and proportions of informative SNP. Kijas et al. (2012) selected 89 SNPs with MAF values above 0.20. Validation studies with 108 Corriedale and 110 Merino in Uruguay showed MAF averages of 0.32 and maximum cumulative exclusion probabilities of 0.9998 and 0.99999 with Corriedale and Merino, respectively (Macedo et al. (2014)).

A validation of the 72 and 93 SNPs for parentage testing was carried out with 34 pair ewe-lamb. Results are presented in Table 1. First, the genotypes comparison using the HD chip indicated that two of the pairs did not match and 32 were correct. Both parentage testing SNP panels successfully identified the correct pairs (0 conflicts) and clearly showed those that did not match.

Table 1. Parentage confirmation of 34 pairs of ewes-lamb identified in the field using HD chip (600k SNPs) and parentage testing SNP panels with 72 and 93 SNPs.

Test result	HD chip	72 SNP	93 SNP
Confirmed Lamb-ewe	32	32 (0 conflict)	32 (0 conflict)
Rejected Lamb-ewe	2	2 (7 conflicts)	2 (10 conflicts)

In general terms, genomic approaches provide useful information for the characterization of animal genetic resources and their conservation. The HD SNP panel provides very useful information for a comprehensive characterization of Creole sheep and to reconstruct the pedigree population structure. Data will be used to find the best and sustainable in situ conservation plan. The sustainability of the conservation plan also relies on applying cost-effective and accurate DNA-based technologies to follow and/or check dam and sire pedigree. Despite the high level of homozygosis in the Creole sheep population, SNPs selected for parentage testing were informative and allowed accurate parentage verification. These preliminary results encourage further studies for a final identification of the SNPs to be used and a comprehensive validation.

Conclusions

The preselected SNPs for paternity testing were informative in a inbreed population of Creole sheep that has high levels of homozygosis according to genomic results obtained with HD SNP chip. A total of 72 and 93 SNPs out of 278 included in this subset reached probability exclusion levels of 0.9999 and 0.99999, respectively, which will ensure accurate identifications of parents. A first validation in a small number of confirmed pairs of parent-progeny indicated that both sets of SNPs successfully identified correct and incorrect ewe-lamb pairs. Although these are preliminary results they show the feasibility of using these SNPs for parentage assignment for confirmation of pedigrees. This information will help to correctly record the genealogy of the population in the conservation plan to be implemented.

Literature Cited

Allendorf, F.W., Hohenlohe, P.A., and Luikart, G. (2010) *Nat. Rev.* 11:697-709.

Anderson, R., McEwan, J., Brauning, R. et al. (2014). Plant and Animal Genome XXII. San Diego, C.A.

FAO. 2013. In vivo conservation of animal genetic resources. FAO Animal Production and Health Guidelines. No. 14. Rome.

Grasso, N.A., Goldberg, V., Navajas, E.A. et al. (2014). *Genet. Mol. Biol.* (Submitted)

Hayes, B. J. (2011) *J. Dairy Sci.* 94:2114–2117

Heaton, M.P., Leymaster, K.A., Kalbfleisch, T.S. et al. (2014). Plant and Animal Genome XXII. San Diego, C.A.

Jamieson, A., and Taylor, S. C. (1997). *Anim. Genet.* 28: 397–400.

Kijas, J.W., Lenstra, J.A., Hayes, B. et al. (2012) *PLoS Biol.* 10:e1001258. doi:10.1371/journal.pbio.1001258

Macedo, F., Grasso, N., Navajas E.A. et al. (2014). *Arch Latinoam. Prod. Anim.* (Accepted).

Rousset, F. (2008). *Mol. Ecol. Resour.* 8: 103-106.