SNP genotyping for parentage identification in a Merino nucleus and in a commercial Highlander flock

V. Goldberg[1], F. Macedo[2] & G. Ciappesoni[1].

[1] Instituto Nacional de Investigación Agropecuaria, INIA Las Brujas, Ruta 48, km 10, Rincón del Colorado, 90200, Canelones, Uruguay.
[2] Facultad de Veterinaria, Universidad de la República, Alberto Lasplaces 1150, 11600, Montevideo, Uruguay.
vgoldberg@inia.org.uy *(Corresponding author)*

Summary

Pedigree information is required to estimate breeding values accurately and to ensure high rates of genetic gain. DNA markers information can be used as an alternative to traditional recording of pedigree, being SNPs the markers of choice for parentage verification. The aim of the present study was to determine sheep parentage by SNPs using a very low density panel in: 1) a stud Merino flock to know the percentage of parentage error in order to correct pedigree misidentification and; 2) a commercial Highlander flock, to study the possibility to not control lambing anymore because for the breeder it is very laborious, time-consuming and disturb the relationship ewe-lambs during parturition. Genomic DNA was isolated from 200 samples of Merino sheep in 2015 and from 108 and 904 samples of Merino and Highlander sheep in 2016, respectively; and were genotyped with a very low density panel containing 507 SNPs. In 2015, for the 91 lambs genotyped, the error in parentage assignment was 16.5%. Although the assigned sire did not match with the declared sire for 15 lambs, the true sire was assigned for 14 of them. Thus, 99% of the lambs genotyped, had a sire assigned by the SNP panel. For the 80 lambs with their dams genotyped, the error rate was 12.5%. For the 10 lambs with mismatches with the declared dams, the true dam was assigned for five of them. In 2016, Merino samples were genotyped to link only lambs with their sires. For the 101 lambs, the error of parentage assignment was 21.8%. For Highlander, 7.5% of samples failed the genotyping and analyses was conducted without knowledge of the relationships between lambs with sires and dams. The 51% of the lambs genotyped had a sire and a dam assigned, 21% had only the sire assigned and 15% had only assigned a dam. Thus, genotyping by SNPs assigned a sire to the 72% of the lambs and a dam to the 66% of the lambs. The main problem was that only 5 of the 11 rams used as sires were genotyped and the high percentage of samples which failed the genotyping. However, taking into account the lack of information related to pedigree, we consider that a high rate of lambs had a sire and a dam assigned by the SNP panel. In conclusion, the development of a SNP panel for parentage assignment at a low price, would provide breeders with the opportunity of making mating management and control lambing easier and more relaxed while improving the known of pedigree information, ensuring high rates of genetic gain.

Keywords: molecular markers, parentage exclusion, parentage verification, sheep.

Introduction

It is well known that efficient genetic evaluation systems require phenotypic information and the collection of pedigree to estimate breeding values (BV) for the traits of interest. Complete pedigree information is required not only to estimate BV accurately but also for inbreeding management, thus, ensuring high rates of genetic gain (Clarke *et al*., 2014). According to the reviewed by Dodds *et al*. (2005) the estimated pedigree error rates in sheep are in the range of 1 and 15%, leading to decrease the genetic progress into a breeding program and ultimately, to profit loss. In sheep production, is very common multi-sire mating, and when artificial insemination is used, those ewes that repeat heat are mated with another ram in the following estrous cycle, unknowing the parentage of the offspring or assigning it incorrectly in other cases. It is usual that the assignment be made controlling the number of days between the insemination and the lamb birth, leading to high rates of errors. Additionally, ewe-lamb pairing is assigned visually, which requires a lot of work and time during lambing season. Consequently, is necessary to establish reliable methodologies for paternity exclusion to improve the quality of genealogical data in populations under genetic evaluation (Macedo *et al.,* 2013).

An alternative to traditional recording of pedigree, is to use DNA marker information to identify parents. SNPs are attractive as parentage markers because they are abundant, genetically stable and amenable to accurate high-throughput automated genotyping platforms (Heaton *et al.*, 2014). Parentage identification by SNPs requires high genotyping accuracy (≥99%) and high minor allele frequency of the markers (MAF≥0.30) (Heaton *et al.*, 2014). This approach also requires the estimation of the probability of exclusion, that is a measure of efficiency in paternity testing and it refers to the a priori ability of a battery of tests to detect paternity inconsistencies. This parameter measures the capacity of the system to detect a false accusation of paternity (Cifuentes *et al.*, 2016).

The aim of the present study was to determine sheep parentage by SNPs using a very low density panel in:

a) a stud flock of Australian Merino sheep in order to know the percentage of error in the parentage assignment and correct pedigree misidentification; and
b) a commercial Highlander flock, with the objective to study the possibility to register pedigree through SNP genotyping and do not control lambing anymore because for the farmer it is very laborious, time-consuming and disturb the relationship ewe-lambs during the first hours of birth causing high rates of miss-mothering.

Material and methods

SNP panel

In 2012, the National Research Institute of Agricultural (INIA) of Uruguay, carried out a research project in Merino and Corriedale breeds with the main objective to find SNPs associated to gastrointestinal parasite (GIP) resistance and to develop a very low density panel containing also SNPs for parentage identification. Animals were genotyped with the OvineSNP50K BeadChip (Illumina®) and whole genome sequencing was carried out (Geneseek, Nebraska, California) in order to select candidate markers.

At the final of the project, by an agreement between INIA and Eureka Genomics (currently is part of Affimetrix, Thermo Fisher Scientific), a very low density panel was developed, containing a subset of 507 SNPs as is shown in Table 1. A requirement for choosing the candidate markers was that they were well distributed over the 26 autosomal chromosomes. Within each chromosome was tried to maintain a considerable distance between SNPs. A total of 250 SNPs were selected for parentage identification, from a pool of SNPs associated to GIP resistance, based on two principal criteria: MAF in each breed and the physical location in the chromosome. Actually, these markers are included in the 15K panel (Illumina®) developed by the International Sheep Genome Consortium (ISGC). In addition, 69 markers from 89 SNPs described by Kijas *et al*. (2012) were also selected for paternity exclusion, which were used in a large number of breeds distributed all around the world.

Population study

Merino flock

It is an Australian Merino nucleus selected for fine and ultra-fine wool fiber, belonging to "Glencoe", a research experimental station of INIA, and that also participate in the National Genetic Evaluation of the breed. In this flock, lambing is controlled 24 hours a day during the peaks of parturition in order to identify the pair dam-lamb. During the mating season, all ewes are mainly artificially inseminated with fresh semen, and in the case they repeat heat, they are mated with another ram in the following estrous cycle.

Highlander flock

This is a nucleus and commercial flock, where animals are evaluated through an intra-flock genetic evaluation. The breeder of this flock was very interested in parentage identification by SNPs because Highlander is a prolific breed and during parturition they have a lot of work identifying dam-lamb pairs visually and they have also observed that intensive recording often causes stress to dams, which increase miss-mothering and cross-mothering. In this flock, most of ewes are multiple-sire mated, and a low percentage are artificially inseminated with frozen semen. Since 2015, the breeder decided to not control lambing and complete pedigree information by SNPs genotyping.

Samples

In 2015, 200 blood samples from Merino sheep were collected into heparinized tubes (100 lambs, 91 dams and 9 rams). In 2016, blood samples from 108 Merino (101 lambs and 7 rams) and 904 Highlander sheep were collected (rams, dams and lambs).

Genomic DNA was isolated and genotyped with the 507 SNPs panel (Affimetrix Inc.). The MAF was calculated for each SNP for both breeds. The parentage exclusion was calculated with Cervus© package (Kalinowski *et al*., 2007).

Results and discussion

In Tables 2 and 3 are the values for call rates of the samples, of the SNPs and MAF for all the markers in the panel, and for the 319 SNPs for parentage identification; respectively.

In 2015, 13 of the 200 Merino samples failed in the genotyping (9 lambs and 4 dams, which one was the dam of twins). For 76 of the 91 lambs genotyped, the assigned sire

matched with the declared sire (83.5% agreement). For the remaining 15 lambs, the assigned sire did not match the declared sire but the true sire was assigned for 14 of them. Thus, 99% of the lambs genotyped, had a sire assigned by the SNP panel. For the total of the lambs genotyped, only 80 had the dam genotyped because some of them were not in the flock at the moment of the blood sampling and four of them failed in the genotyping. For 70 lambs, the assigned dam matched with the declared dam (87.5% agreement). For the 10 lambs with mismatches with the declared dams, to five of them was assigned the true dam. In 2016, Merino samples were genotyped to link lambs only with their sires. For the 101 lambs, the parentage verification had an agreement of 77.2% and a lamb that had not a sire declared, the SNPs genotyping could assign a sire.

The Merino flock is a selection nucleus which participate in the National Genetic Evaluation and is one of the main flock through Merino breeders connect for the evaluation. These results showed that the percentages of error assignment were higher than expected. In the case of the pair lamb-dam the error rate was 12.5% and in this flock lambing control is carried out during 24 hours. On the other hand, the error rate for the pair lamb-sire was 16.5 and 21.8% in 2015 and 2016, respectively. In this flock artificial insemination is carried out with the mating with a new ram for those dams that repeat heat, leading to errors in the sire assignment. The SNPs genotyping allowed to correct pedigree misidentification leading to estimate BV accurately and consequently, increase the genetic gain. Misidentification rates of 7 to 15% have been reported which have led to a 2.5% to 15% decrease in genetic gain (in Clarke *et al.* 2014), and ultimately lead to profit loss.

For Highlander breed, 68 of the 904 samples failed the genotyping (one ram, 56 dams and 11 lambs) and analyses of the genotype data was conducted without knowledge of the relationships between lambs with sires and dams. The failing samples were correlated with low DNA input. The 51% of the lambs genotyped had a sire and a dam assigned, 21% had only the sire assigned and 15% had only assigned a dam. Thus, only having partial information of the mating, genotyping by SNPs assigned a sire to the 72% of the lambs and a dam to the 66% of the lambs. The reasons that 28% of the lambs did not have a sire assigned were that of the 10 rams used in 2015 as sires, four of them were not sampled because at the moment to collect blood, they were death or sold, and additionally, one of the samples that failed the genotyping was a ram. The main causes that 34% of lambs had not a dam assigned were that 56 samples which failed the genotyping were dams and in the other hand, there were dams which were not sampled, thus, their DNA were not genotyped. In concordance, Clarke *et al.* (2014) published that the use of pure exclusion methods to assign parentage may fail if information is limiting (where not all parents have been sampled) and genotyping errors are present.

We think that the main problems in Highlander flock were the high percentage of samples which failed the genotyping, and mostly, the high percentage of the sires that were not genotyped. However, in spite of this and taking into account the lack of information related to pedigree, we consider that a high rate of lambs had a sire and a dam assigned by the SNP panel.

Conclusion

The development of a dedicated SNP panel for parentage assignment at a low price, would provide breeders with the opportunity of making mating management and control

lambing easier and more relaxed, while improving the accuracy of the animal's known pedigree, helping to accelerate genetic improvement. Our results showed that in the Merino stud flock, where control parturition is carried out during 24 hours, and sire mating is also controlled, the rates of error in parentage assignment were higher than expected. In the Highlander flock, where control lambing was not carried out and with ignorance about genealogy, the percentage of sire and dam assignment was higher than expected. The most important point to take into account is to ensure that all sires be sampled and genotyped, and the same for dams if full pedigrees are required. As it was showed, this very low density panel can be utilized to infer paternity in the absence of dam information.

References

Cifuentes, L.O., E. H. Martínez, M. P. Acuña & H. G. Jonquera, 2016. Probability of Exclusion in Paternity Testing: Time to Reassess. L. Forensic Sci. 51 (2): 349-350.

Clark, S. M., H.M. Henry, K.G. Dodds, T.W.D. Jowett, T.R. Manley, R.M. Anderson & J.C. McEwan, 2014. A high throughput single nucleotide polymorphism multiplex assay for parentage assignment in New Zealand sheep. PLoS ONE 9(4): e93392.

Dodds, K.G., M.L. Tate & J.A. Sise, 2005. Genetic evaluation using parentage information from genetic markers. J. Anim. Sci. 83: 2271-2279.

Heaton, M. P., K. A. Leymaster, T. S. Kalbfleisch, J. W. Kijas, S. M. Clarke, *et al.*, 2014. SNPs for Parentage Testing and Traceability in Globally Diverse Breeds of Sheep. PLoS ONE 9(4): e94851.

Kalinowski, S. T., M .L. Taper & T. C. Marshall, 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol., 16(5): 1099-1106.

Kijas, J. W., J. McEwan, S. M. Clarke, J.F. Maddox, R. Mcculloch, F. Driver, K. Ilic, K. & M. P. Heaton, 2012. Development of a SNP panel for parentage assignment in sheep. In: Plant and Animal Genome XX Conference. San Diego, California. Poster#P0577.

Macedo, F., N. Grasso, E.A. Navajas, D. Gimeno & G. Ciappesoni, 2013. Exclusión de paternidad mediante un panel de 89 marcadores SNP en una muestra de ovinos Corriedale y Merino de Uruguay. Arch. Lat. Prod. Anim. 21 (4): 215-218.

Sise, J. A., K. G. Dodds, A. G. S. Ramsden & M. L. Tate. 2001. Optimising DNA parentage testing in sheep. Proc. Assoc. Adv. Anim. Breed. Genet. 14:317–320.

Table 1. List of SNPs contained in the very low density panel developed by INIA & Eureka genomics.

n of SNPs	function	Reference
250	Parentage identification	INIA Project
69	Parentage identification	Kijas *et al.* (2012)
2	Horn/poll	Dominik *et al.* (2011) & Johnston *et al.* (2011)
174	Resistance/susceptibility to gastrointestinal parasites	INIA Project
7	Merino breed specific	INIA Project
5	Corriedale breed specific	INIA Project

Table 2. Values of call rate of samples (CR_{sample}), call rate of SNP (CR_{snp}) and Minor allele frequency (MAF) of the three genotyping for the 507 SNPs.

		Merino 2015	Merino 2016	Highlander 2016
CR_{sample}	Mean	0.99	0.94	0.95
	Minimum	0.93	0.84	0.80
	Maximum	1	0.99	0.99
CR_{snp}	Mean	0.99	0.94	0.95
	Minimum	0.30	0.34	0.45
	Maximum	1	1	0.99
MAF	Mean	0.33	0.33	0.31
	Minimum	0	0	0
	Maximum	0.50	0.50	0.50

Table 3. Values of call rate of SNP (CR_{snp}) and Minor allele frequency (MAF) of the three genotyping for the 319 SNPs for parentage identification.

		Merino 2015	Merino 2016	Highlander 2016
CR_{snp}	Mean	0.99	0.94	0.95
	Minimum	0.30	0.34	0.45
	Maximum	1	1	0.99
MAF	Mean	0.37	0.37	0.34
	Minimum	0	0.01	0
	Maximum	0.50	0.50	0.49