

## Feasibility of a genomic selection approach for gastrointestinal nematodes resistance based on single-step evaluation and a female informative population in Sarda breed sheep

S. Salaris<sup>1</sup>, S. Casu<sup>1</sup>, M.G. Usai<sup>1</sup>, A. Scala<sup>2</sup> and A. Carta<sup>1</sup>

<sup>1</sup>Research Unit: Genetics and biotechnology, AGRIS Sardegna, 07100, Sassari, Italy

<sup>2</sup>Department of Veterinary Medicine - University of Sassari, 07100, Sassari, Italy

Corresponding Author: [acarta@agrisricerca.it](mailto:acarta@agrisricerca.it) ()

Gastrointestinal nematodes are a major health problem in grazing animals. In Sardinia, sheep farming systems are based on grazing natural pastures and forage crops where infection is unavoidable. However, the inclusion of nematode resistance in the current breeding scheme is difficult due to the laboriousness of the fecal egg counting at Herd Book scale. The aim of this work was to assess the feasibility of a genomic selection approach based on a female nucleus as reference population and young Herd Book rams as selection candidates by using ssGBLUP. The female nucleus consisted of 5386 ewes from 10 F1 Sarda x Lacaune and 208 Sarda rams. From 2000 to 2020, 16,977 individual records were collected in 25 dates on 4,304 ewes. All recorded ewes, their sires and available male ancestors were genotyped as well as 734 Herd Book rams. The weights of A and G matrices when building the H matrix in ssGBLUP were varied from 0 to 1 to detect the model showing the highest heritability and predictive ability and then using it for the BV prediction of the selection candidates. The highest heritability estimate (0.27) was obtained when the G weight was between 0.20 and 0.40 whereas predictive ability was more or less constant across the different G weights. The average expected accuracies of 245 young Herd Book rams depended on the amount of information on relatives in the female nucleus. Overall results show that in our population high weights of the pedigree information are needed when building the relationship matrix in ssGBLUP to improve heritability and predictive ability. It is likely that pedigree information adjusts for genomic relationships which are partly due to IBS rather than IBD genome sharing. Moreover, recording pedigree information in the Herd Book is crucial also to calculate the amount of information on relatives in FRP of selection candidates and predict their expected accuracies.

### Abstract

*Keywords: Genomic selection, female nucleus, nematodes resistance, ssGBLUP.*

Gastrointestinal nematodes (GIN) are major health problems in grazing animals (Kaplan *et al.*, 2012). GIN infections determine important yield reductions and increased production costs due to medical treatments and higher culling rates (Mavrot *et al.*, 2015). Moreover, anti-helminthics resistance is more and more an issue in several countries (Sargison *et al.*, 2007). Fecal egg count (FEC), i.e., the number of parasite eggs per gram of faeces has been largely used as proxy trait to measure resistance to GIN. In Sardinia, sheep farming systems are based on grazing natural pastures and forage crops where GIN infection is unavoidable. Thus, most farmers administrate anti-helminthics often with an empirical approach in terms of individual diagnosis, doses

### Introduction

and frequency of treatments (Sechi *et al.*, 2010). Genetic variation between individuals and breeds has been documented (Assenza *et al.*, 2014; Bouix *et al.*, 1998; Sechi *et al.*, 2009). However, the inclusion of GIN resistance in current breeding schemes is difficult due to the laboriousness of the FEC recording at Herd Book (HB) scale.

The aim of this work was to assess the feasibility of a genomic selection approach based on a female nucleus as reference population and young HB rams as selection candidates by:

1. Using single-step GBLUP to improve the genomic predictions accuracy of HB rams by exploiting both genomic and pedigree information;
2. Detecting the optimal weights of **A** and **G** matrices into **H** matrix using heritability and predictive ability as indicators;
3. Evaluating the expected accuracies of selection candidates.

## Material and methods

### Female reference population

The generation of the female reference population (FRP) started in 1999 when 10 Lacaune × Sarda F1 sires were mated to Sarda ewes to produce 928 back-cross female lambs. Subsequently, only HB Sarda rams were used to produce the yearly replacement of FRP. Until 2009, the average size of the sire families was 43 daughters whereas, from 2010, it was reduced to 9 daughters to increase the number of bloodlines represented in the FRP. Sarda sires were chosen among those with high genetic impact in HB from the Artificial Insemination Center of the breed. In total, 5,386 ewes from 218 rams (10 F1 and 208 SA) were generated until 2020. The farming system was similar to that commonly applied in Sardinia with most of the adult ewes lambing in autumn and yearlings lambing between January and March. The feeding regime was based on grazing of natural pastures and forage crops, supplemented by hay and concentrates in winter and late spring.

### Phenotyping

FEC was the proxy trait used to assess GIN resistance under natural conditions of infection in the experimental flock. Periodically, a sample of around 50 ewes representative of the different management groups was monitored to evaluate the percentage of infected animals and decide whether to sample the whole flock and possibly administrate anthelmintic.

FEC was determined using a copro-microscopic exam conforming to the McMaster technique on individual samples. When the number of infected animals and the level of infestation was considered sufficient to appreciate individual variability, individual FEC were measured on the whole flock. From 2000 to 2020 individual FEC were recorded from 1 to 3 times per year, more frequently in September and July. Due to the low level of infection found, no individual measures were realised between June 2006 and November 2007. FEC recording for a QTL detection experiment was closed in 2012 and restarted in 2015 for the new cohorts with the perspective of implementing MA or genomic selection.

Table 1. Distribution of FEC records (NR), dates (ND) and animals (NA) per year (Y).

Y	2000	2001	2002	2003	2005	2006	2008	2009	2010	2011	2012	2016	2017	2018	2019	2020
NR	1813	1775	1703	1631	717	1803	1270	748	1190	636	588	931	540	599	640	393
ND	2	2	2	2	1	3	2	1	2	1	1	2	1	1	1	1
NA	914	906	853	988	717	879	918	748	792	636	588	740	540	599	640	393

Finally, 16,977 FEC measurements were recorded in 25 dates on 4,304 animals (Table 1).

FEC presented skewed distribution and were log-transformed using  $\ln\text{FEC}=\ln(\text{FEC}+14)$ .

All 4,304 ewes of FRP, 194 sires and 78 available male ancestors (272 rams) as well as 734 HB Sarda rams were genotyped with the Illumina Inc. OvineSNP50 Beadchip. SNP editing was performed using call rate and MAF thresholds of 90% and 1%, respectively. After quality control, 43,293 SNP across 26 autosomes were retained for the analysis.

### Genotyping

Single-step genomic BLUP was used to estimate variance components and breeding values with a repeated animal model including the permanent environment and additive genetic random effects. Environmental fixed effects were the date of sampling, the age of the animal (from 1 to 4 years) and the interaction between the number of lambs and the days from lambing. The additive genetic effect was modeled using a realized **H** relationship matrix, which combines pedigree and genomic relationships (Legarra *et al.*, 2009). Compared to BLUP, in ssGBLUP, the inverse of the numerator relationship matrix **A**<sup>-1</sup> is replaced by **H**<sup>-1</sup> defined as follows:

### Single-step genomic BLUP

in which **G** is the genomic relationship matrix. The **G\*** matrix used for blending was obtained according to VanRaden (2008) and Aguilar (2010) as:

$$\mathbf{G}^* = [w\mathbf{G}^{-1} + (1-w)\mathbf{A}_{22}^{-1}]^{-1}$$

were the weight (*w*) ranged between 0.0 and 1.0 by 0.2. Heritability (*h*<sup>2</sup>) and repeatability (*r*) were computed as

$$\sigma_a^2 / (\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2) \text{ and } (\sigma_a^2 + \sigma_{pe}^2) / (\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2) \text{ respectively}$$

where *s*<sub>a</sub><sup>2</sup> is the additive genetic variance, *s*<sub>pe</sub><sup>2</sup> is the permanent environment variance and *s*<sub>e</sub><sup>2</sup> is the error variance. The parameter *w* scales differences between genomic and pedigree-based information. (Christensen and Lund, 2012; Meyer *et al.*, 2018; Lourenco *et al.*, 2020). In several studies, different values of *w* were used to construct the relationship matrix without estimating variance components and *h*<sup>2</sup> (Aguilar *et al.*, 2010; McMillan and Swan, 2017). Usually variance component estimation has been performed just for traditional AMBLUP, GBLUP and ssGBLUP with a *w* between 0.95 and 0.99 (Aldridge *et al.*, 2020; Cesarani *et al.*, 2019; Forni *et al.*, 2011; Hidalgo *et*

*al.*, 2020). In our study,  $w$  was varied from 0.0 to 1.0 to detect the blending weights of **A** and **G** matrices showing the highest  $h^2$  and PA and then using them for the BV prediction of the selection candidates.

Single step GBLUP was performed using the package AIREMLf90 included in the BLUPf90 suite programs (Misztal *et al.*, 2002). The following options were provided to AIREMLf90 to built **H** matrix and blending **A** and **G**:

- the pedigree file was provided to built the  $\mathbf{A}^{-1}$  internally including inbreeding;
- the genomic relationship matrix was created as in VanRaden (2008);

The pedigree file included 6,242 animals tracing back to three generations of FRP ewes. Further 2,323 animals were added to the pedigree file when HB Sarda rams were included in the analyses.

The predictive ability (PA) of future phenotypes was calculated for each tested  $w$  as the correlation between the breeding values predictions and the individual average yield deviation of 235 last generations ewes. Breeding values and their expected accuracies ( $r_{BV}$ ) were calculated with a training dataset in which the 421 records of the 235 ewes were masked.

### Expected accuracy of selection candidates HB rams

The model showing the highest  $h^2$  and PA estimates was used for the GEBV estimation of 1006 genotyped rams born from 1993 to 2020 collected in different frameworks of genotyping activities and classified according to their age and level of relationship with FRP:

- 272 sires or ancestors of FRP ewes;
- 489 old HB rams born from 1993 to 2015 with daughters in HB flocks
- 245 young HB rams (including 27 son of FRP ewes) born from 2016 and 2020 without daughters in HB flocks and considered as selection candidates.

Expected accuracy of breeding values of animal  $i$  were estimated as:

$$r_{BV_i} = \sqrt{1 - \frac{SEP_i^2}{h_{ii} \times \sigma_a^2}}$$

where  $SEP_i$  is the standard error of prediction derived from the diagonal element of the left-hand side inverse of the mixed model equations and  $h_{ii}$  is the diagonal element for the animal  $i$  of the **H** matrix.

The overall results for FEC were compared with those obtained for milk yield (MY) in the same population using the same approach (unpublished results).

### Results and discussion

The overall mean and standard deviation of FEC and InFEC were  $332 \pm 727$  and  $4.82 \pm 1.43$ , respectively.

Estimated variance components are shown in Table 2. The highest values of  $\sigma_a^2$  were reached when  $w$  is between 0.2 and 0.4 and then decreases as  $w$  increases. Moreover,  $\sigma_{pe}^2$  decreases as  $\sigma_a^2$  increases while  $\sigma_e^2$  is quite constant across the models.

Table 2. Estimated variance components ( $\sigma^2_a$  : additive genetic variance;  $\sigma^2_{pe}$ : permanent environment variance;  $\sigma^2_e$ : error variance;  $\sigma^2_{tot}$ : total variance) based on single-step GBLUP varying the weight ( $wG$ ) of genomic information in blending H.

	$wG=0.0$	$wG=0.2$	$wG=0.4$	$wG=0.6$	$wG=0.8$	$wG=1.0$
$\sigma^2_a$	$0.416 \pm 0.035$	$0.460 \pm 0.035$	$0.459 \pm 0.034$	$0.438 \pm 0.032$	$0.407 \pm 0.030$	$0.383 \pm 0.029$
$\sigma^2_{pe}$	$0.067 \pm 0.022$	$0.029 \pm 0.021$	$0.029 \pm 0.020$	$0.045 \pm 0.019$	$0.070 \pm 0.018$	$0.082 \pm 0.017$
$\sigma^2_e$	$1.189 \pm 0.015$	$1.190 \pm 0.015$	$1.191 \pm 0.015$	$1.192 \pm 0.015$	$1.192 \pm 0.015$	$1.193 \pm 0.015$
$\sigma^2_{tot}$	1.672	1.680	1.678	1.675	1.669	1.658

Table 3. Heritability ( $h^2$ ), repeatability ( $r$ ), predictive ability (PA) and mean of expected accuracy ( $r_{BV}$ ) of single-step GBLUP ranging the weight ( $wG$ ) of genomic information when blending A and G into H.

	$wG = 0.0$		$wG = 0.2$		$wG = 0.4$		$wG = 0.6$		$wG = 0.8$		$wG = 1.0$	
	FEC	MY										
$h^2$	0.249	0.413	0.274	0.441	0.273	0.424	0.262	0.392	0.244	0.354	0.231	0.328
$r$	0.289	0.585	0.292	0.587	0.290	0.585	0.288	0.581	0.286	0.577	0.281	0.573
PA	0.304	0.244	0.317	0.262	0.320	0.260	0.320	0.251	0.319	0.238	0.316	0.221
$r_{BV}$	0.450	0.479	0.470	0.496	0.493	0.514	0.519	0.535	0.547	0.557	0.575	0.582

In Table 3 are reported the results in terms of  $h^2$ ,  $r$ , PA and  $r_{BV}$  in comparison with those obtained for milk yield on the same population with the same approach. For FEC, in agreement with the variance components estimates,  $h^2$  shows the highest value when  $w$  is between 0.2 and 0.4 and gradually decreases as  $w$  increases. When  $w$  assumes values between 0.2 and 0.4,  $h^2$  is 18% higher than when just the genomic information is considered ( $w=1.0$ ). On the other hand, repeatability is quite similar across models. PA is more or less constant across the models whereas the average  $r_{BV}$  of the female validation sample increases as  $w$  increases.

The same pattern of  $r_{BV}$  is found for MY. However, in this case, PA shows a pattern similar to  $h^2$  i.e. the model showing the highest  $h^2$  showed also the highest PA estimate.

The  $r_{BV}$  estimates of HB rams, obtained using the model with the highest  $h^2$  and PA ( $w = 0.4$ ), ranged between 0.95 and 0.19 with a mean of  $0.71 \pm 0.18$  for sires or ancestors of FRP,  $0.33 \pm 0.12$  for young HB rams and  $0.30 \pm 0.08$  for old HB rams. Among the selection candidates, the highest values of  $r_{BV}$  were obtained for the progeny of FRP ewes as shown in Figure 1.

As shown in previous studies (Pszczola *et al.*, 2012; Usai *et al.*, 2018),  $r_{BV}$  of selection candidates can be easily and accurately predicted by parameters derived by the numerator relationships matrix which measure their amount of information on relatives in the reference populations. This result is important for designing the flow of animals from and toward the nucleus in order to achieve good accuracies for most selection candidates.

Overall heritability results show that selective breeding may be an option to increase resistance to infection of gastro-intestinal nematodes in sheep. Both traditional progeny testing and genomic selection with a female reference population are realistic options. However, the laboriousness of fecal egg counting at Herd Book scale makes genomic selection based on a female nucleus potentially more profitable in terms of costs-

## Conclusion

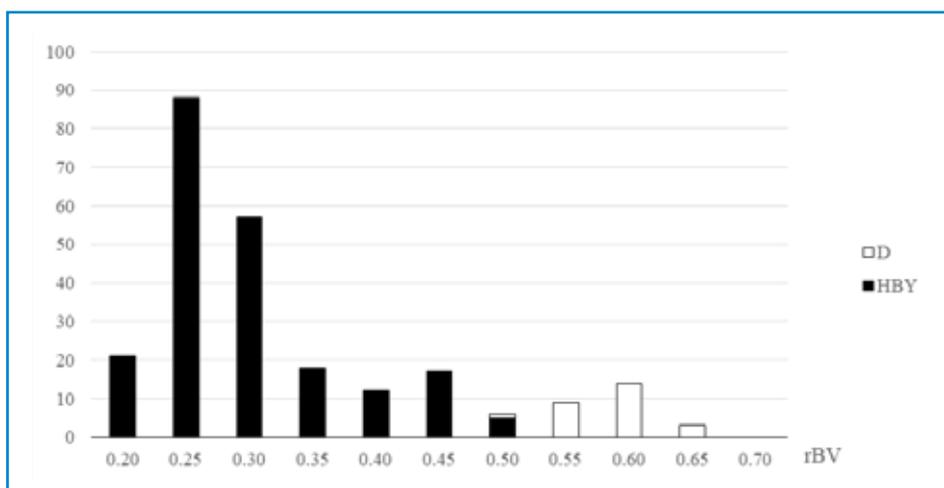


Figure 1. Distribution of the expected accuracies ( $r_{BV}$ ) of 245 selection candidates (D: descendants of FRP; HBY: young HB rams).

benefits. Single-step GBLUP is essential to manage genomic selection based on female nuclei. In our population, high weights of the pedigree information are needed when building the relationship matrix to improve heritability and predictive ability. It is likely that pedigree information adjusts for genomic relationships which are partly due to IBS rather than IBD genome sharing. Moreover, recording pedigree information in HB is crucial also to calculate the amount of information on relatives of selection candidates in FRP and predict their expected accuracies. A better understanding of genomic predictions accuracies is needed to optimize the flow of animals from and toward the nucleus in order to achieve good accuracies for most selection candidates. As a perspective, accuracies of genomic predictions may be improved including in the evaluation model 10 highly significant SNP identified in QTL detection analysis (Casu *et al.*, 2021, submitted). In the short term, genomic predictions for resistance to gastro-intestinal nematodes will be delivered just for the herd book selection candidates well genetically connected with the female nucleus.

## List of references

- Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta and T.J. Lawlor.** 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93, 743–752.
- Aldridge, M.N., J. Vandenplas, R. Bergsma, and M.P.L. Calus.** 2020. Variance estimates are similar using pedigree or genomic relationships with or without the use of metafounders or the algorithm for proven and young animals. *J. Anim. Sci.* Vol. 98, No. 3, 1–9
- Assenza F, J-M. Elsen, A. Legarra, C. Carré, G. Sallé, C. Robert-Granié, et al.** 2014. Genetic parameters for growth and faecal worm egg count following *Haemonchus contortus* experimental infestations using pedigree and molecular information. *Genet Sel Evol.* Feb 14;46(1):13.

- Bouix J, J. Krupinski, R. Rzepecki, B. Nowosad, I. Skrzyzala, M. Roborzynski, *et al.*** 1998. Genetic resistance to gastrointestinal nematode parasites in Polish long-wool sheep. *Int J Parasitol.* 1998 Nov;28(11):1797–804.
- Cesarani, A., G. Gaspa, F. Correddu, M. Cellesi, C. Dimauro, and N.P.P. Macciotta.** 2019. Genomic selection of milk fatty acid composition in Sarda dairy sheep: Effect of different phenotypes and relationship matrices on heritability and breeding value accuracy. *J. Dairy Sci.* 102:3189–3203.
- Christensen, O.F., and M.S. Lund.** 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.* 42, 2.
- Forni, S., I. Aguilar, I. Misztal.** 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree, and genomic information. *Genet. Sel. Evol.* 43, 1.
- Hidalgo, J., S. Tsuruta, D. Lourenco, Y. Masuda, Y. Huang, K.A. Gray, and I. Misztal.** 2020. Changes in genetic parameters for fitness and growth traits in pigs under genomic selection. *J. Anim. Sci.* 1–12.
- Legarra A, I. Aguilar and I. Misztal.** 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92, 4656–4663.
- Kaplan RM and A.N. Vidyashankar.** 2012. An inconvenient truth: Global worming and anthelmintic resistance. *Vet Parasitol.* 186(1–2):70–8.
- Mavrot F, H. Hertzberg and P. Torgerson.** 2015. Effect of gastro-intestinal nematode infection on sheep performance: A systematic review and meta-analysis. *Parasites and Vectors.* Oct 24;8(1):557.
- McMillan, A.J., and A.A. Swan.** 2017. Weighting of genomic and pedigree relationships in single step evaluation of carcass traits in Australian sheep. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 22:557-560.
- Meyer, K., B. Tier and A. Swan.** 2018. Estimates of genetic trend for single step genomic evaluations. *Genet Sel Evol.* 50:39.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee.** 2002. Blupf90 and related programs (BGF90). Commun. No. 28–07 in Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France.
- Pszczola M, T. Strabel, H.A. Mulder, M.P.L. Calus.** 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J Dairy Sci.* 95:389–400.
- Sargison N.D., F. Jackson, D.J. Bartley, D.J. Wilson, L.J. Stenhouse and C.D. Penny.** 2007. Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. *Vet Parasitol.* Apr 10;145(1–2):65–76.
- Sechi S, M. Giobbe M,G. Sanna, S. Casu, A. Carta and A. Scala.** 2010. Effects of anthelmintic treatment on milk production in Sarda dairy ewes naturally infected by gastrointestinal nematodes. *Small Rumin Res.* Feb;88(2–3):145–50.
- Sechi S, S. Salaris, A. Scala, R. Rupp, C. Moreno, SCSC Bishop, *et al.*** 2009. Estimation of (co)variance components of nematode parasites resistance and somatic cell count in dairy sheep. *Ital J Anim Sci.* 8(SUPPL. 2):156–8.



**Usai M.G., S. Salaris, S. Casu, T. Sechi, S. Miari, P. Carta, and A. Carta,** 2018. Feasibility of genomic predictions of Sarda breed rams using a female reference population. Proceedings, 11th WCGALP. Auckland, New Zealand, 11–16 February.

**VanRaden, P.M.** 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414–4423.