Evaluating the accuracy of the genetic ranking of rams in the selected population of the Sarda dairy sheep breed

S. Salaris[1], Sara Casu[1], M.G. Usai[1], P. Fresi[2], A. Carta[1]

[1]*Research Unit: Genetics and Biotechnology, DIRPA-AGRIS Sardegna, S.S. 291 Sassari-Fertilia, Km. 18.6, 07100 - Sassari, Italy*
[2]*Associazione Nazionale della Pastorizia – ASSONAPA, Via Togliatti 1587, 00155 - Rome, Italy*

Abstract

In Sarda breed it is difficult to extend the use of AI to the whole selected population due to specific features of the breeding system. The current rate of AI (8 to 10% of replacement ewes) is not sufficient by itself to assure the genetic connectedness between all selected flocks. However, the breeding scheme is characterized by a large use of the controlled natural mating with an important exchange of rams which may counterbalance the low spread of AI. The aim of this study was to assess the accuracy of the genetic ranking of AI and natural mating rams obtained by BLUP methods. The connectedness correlation (r_{ij}) was used to cluster sires of the last generation in connected subpopulations. The iterative process joins at each step the two sires (or clusters) with the highest r_{ij} in a new cluster. The process stops when the highest r_{ij} equals a given threshold. At 0.005 threshold, a large subpopulation including 1,236 sires was detected. For higher thresholds, two large sub-populations were identified: the largest including most of AI rams and the smallest rams belonging to few elite farms or flocks where farmers were used to buy males from the elite farms. This study showed that the current breeding management allows to reach a sufficient level of reliability of the genetic ranking of animals for most flocks.

Keywords: Connectedness correlation, artificial insemination, dairy sheep.

1.0 Introduction

The main feature of the Sarda breeding program is the large application of the controlled natural mating (CNM) combined with low rates of AI. It is difficult to extend the use of AI to the whole selected population due to specific features of the breeding system. The current rate of AI (8 to 10% of replacement ewes) is not sufficient by itself to assure the genetic connectedness between all selected flocks. However, the large use of CNM with an important exchange of males, mainly sons of AI rams, may counterbalance the low spread of AI. In a previous work, Salaris *et al.* (2008) showed that, although limited by organisational constraints, the combining of CNM with AI allowed to reach a high number of direct genetic links between rams (*i.e.* direct comparisons through lactating daughters in the same management unit). Moreover, the number of completely disconnected flocks have had a strong decrease in the last 20 years. However, the authors concluded that a more precise evaluation of the genetic connectedness of the registered Sarda breed population was needed to assess the accuracy of the genetic ranking of animals. Lewis *et al.* (1999) proposed the correlation of breeding value prediction errors (r_{ij}) as a measure of the connectedness between pairs of sires. Khuen *et al.* (2009) used the average r_{ij} to cluster flocks in different subsets. Salaris *et al.* (2010) proposed to use r_{ij} as a criterion to assign rams to different genetic subpopulation and the square of the correlation between estimated and true differences between breeding values (CD; Laloe *et al.*, 1993) as a measure of the accuracy of the contrast between EBV. The aim of this study was to assess the accuracy of the genetic ranking of AI and CNM rams obtained by BLUP method in the selected population of the Sarda dairy sheep breed by using objective measures of connectedness.

2.0 Material and methods

A dataset including 308,503 first lactation records of yearlings born from known sires (9,330) between 1995 to 2007 was extracted from the 2008 genetic evaluation database managed by the Italian Association of Sheep Breeders (ASSONAPA). The ewes yielded in 10,034 management units (flock-year of

production combinations). Data were analyzed with a mixed model including the management unit as fixed effect and the random sire effect.

The connectedness correlation (r_{ij}) and the coefficient of determination (CD) of the contrast for all possible pairs of 1,404 sires of 21,331 ewes born in 2007 were calculated as:

$$ r_{ij} = \frac{PEC(\hat{u}_i, \hat{u}_j)}{\sqrt{PEV(\hat{u}_i)PEV(\hat{u}_j)}} \; ; \; CD(\mathbf{x}) = \frac{\mathbf{x}'(\mathbf{A} - \lambda \mathbf{C})\mathbf{x}}{\mathbf{x}'\mathbf{A}\mathbf{x}} $$

where PEV is the prediction error variance and PEC is the prediction error covariance of the EBV ($\hat{u}_i$ and $\hat{u}_j$) of the ith and j^{th} sires. PEV and PEC were calculated multiplying the residual variance (σ^2_e) by the diagonal and the off diagonal elements of **C**, the inverse of the fixed effect absorption matrix, [**Z'Z– Z'X(X'X)⁻X'Z**+ **A**⁻¹]⁻¹, respectively, and **x** is the vector of coefficients defining the contrast. **X** and **Z** are the incidence matrices of the management unit fixed factor and the sire random factor respectively, **A** is the numerator relationship matrix, λ was calculated as the ratio of σ^2_e and the sire variance (σ^2_s). σ^2_s and σ^2_e were imposed equal to 80 L^2 and 1,040 L^2.

A clustering iterative process was used to group sires into connected subpopulations using r_{ij} as a measure of distance between sires. The process begins setting each sire in a cluster by itself. Then the iterative procedure joins the 2 sires (or clusters) with the highest r_{ij} in a new cluster. When a cluster includes more than one sire, the average r_{ij} of the included sires with the sires of other clusters was calculated. The procedure stops when the highest r_{ij} equals a given threshold. Accordingly to Khuen *et al.* (2009), two thresholds were chosen: 0.010 ("superior" level of connectedness) and 0.005 ("good" level of connectedness). A further intermediate threshold of 0.007 was introduced. At each threshold, the number and the size of identified clusters were calculated. The following parameters were averaged for relevant (more than 50 sires) clusters: r_{ij}, CD, the total number of daughters per sire (NF); the total number of effective daughters per sire (NFW); the total number of CG in which a sire had daughters (NCG); the total number of directly compared sires per sire (NP); the inbreeding coefficient (IC); the average relationship of a sire within a given CG (PW) calculated as the average relationship coefficient with the other sires in the CG weighted for the effective number of daughters. NF, NFW, NCG and NP affect the accuracy of EBV (Tosh and Wilton, 1994) whereas increasing PW and IC reduce CD of contrast (Kennedy and Trus, 1993).

3.0 Results and Discussion

About 61% of the 2008 recorded yearlings were born from known sires. The average PEV of sires' EBV was 40.8 ± 15.0 L^2. The PEC mean was 0.3 ± 0.4 L^2. Average additive relationship between sires was 0.038 ± 0.055 and ranged between 0 and 0.898. The average r_{ij} of all possible comparisons (984,906) between sires were 0.010 ± 0.011 and ranged between 0 and 0.641. Only 3.5% of all possible pairs of sires showed null r_{ij}.
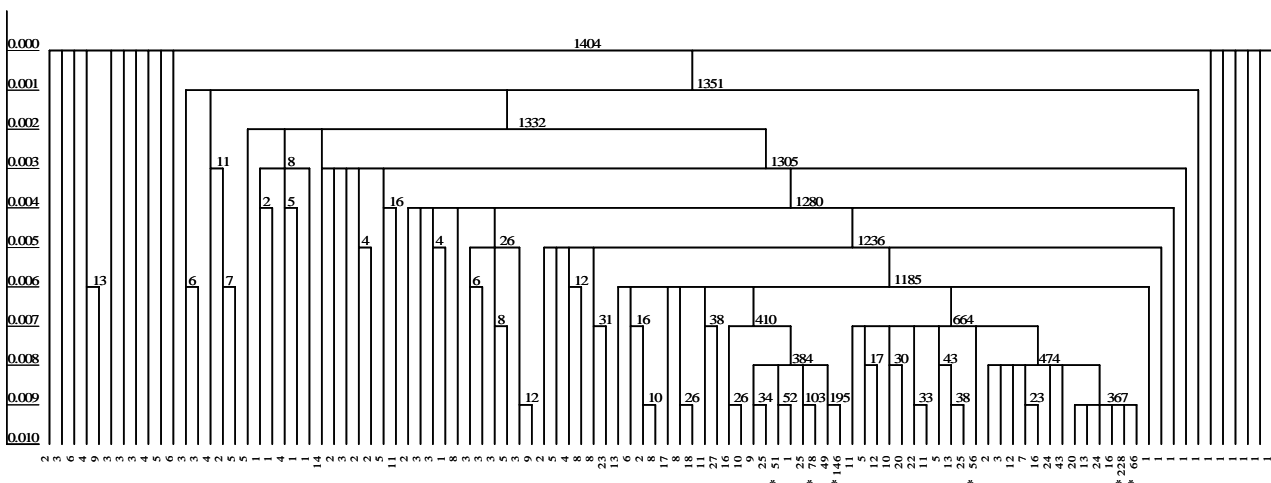

Figure 1. Diagram of clustering procedure results at successively decreasing connectedness correlation (r_{ij}) thresholds (from 0.010 to 0 with a step of 0.001). Horizontal axis represents the number of rams grouped in each cluster at 0.010 r_{ij} threshold level (clusters highlighted with asterisk included more than 50 sires). Vertical axis represents r_{ij} threshold levels. The number of rams grouped in a cluster are reported within the diagram.

The percentage of sires included in the largest cluster increased from 16.2% at 0.010 r_{ij} threshold, to 47.3% at 0.007, and 88.0% at 0.005 (Figure 1). At the 0.010 threshold, 83 different subpopulations including on average 17 ± 31 sires (from 2 to 228 sires) were detected. At 0.005 threshold a large cluster including 1,236 sires and other 26 clusters including more than 2 sires were formed. At 0.007 two large separate subpopulations of rams with 664 and 410 rams respectively were identified. The first included mainly AI sires whereas the second included many rams born in few flocks, located in the central-west part of Sardinia, which represents the historical region where selection of the Sarda breed started. Breeders of these regions usually raise own homebred rams and are the main rams sellers to the registered flocks.

The largest cluster at each threshold included always the largest rate of AI rams and it always showed the higher average CD of contrasts (Table 3). The difference of CD means between the two largest subpopulations at 0.007 threshold was 0.024. The CD mean within the largest cluster at 0.005 threshold was 0.524. This level of connectedness can be ascribed to the strategy of combining a low rate of AI with an important exchange of CNM sires. Actually, Salaris *et al.* (2008) showed that the average number of natural mating groups per flock was 1.9, with 52% of flocks with only 1 sire, 38% with 2 or 3 sires and 10% with more than 3 sires. Over the last twenty years the percentage of CNM sires born by AI increased from 2% to 32%, connecting 19% of flocks not directly involved in the AI program. As far as the genetic links created by CNM rams are concerned, the average annual percentage of sires born in one flock and used in another one was 62% and the average annual percentage of CNM sires with daughters in more than one flock was 23%. Significant higher value of NFW, NCG, and NP and smaller values of PW and IC means were found in clusters with high rate of AI sires. These differences are in the direction of a better accuracy of EBV as confirmed by a significant lower PEV mean.

This study showed that the current breeding management of the Sarda dairy sheep breed allows to reach a sufficient level of accuracy of the genetic ranking of rams involved in the breeding program, including CNM rams which still play a crucial role in the selection program.

Table 1. Number of sires (NS) and AI sires (NAIS) grouped in clusters with more than 50 rams at 0.010, 0.007 and 0.005 r_{ij} threshold levels (TL) and means[1] per cluster of: connectedness correlation (r_{ij}), coefficient of determination o f the contrast (CD), prediction error variance of EBV (PEV), number of daughters (NF) and number of effective daughters (NFw) per sire, number of management units in which the sire has daughters (NCG), number of other sires directly compared to the considered sire (NP), average additive relationship coefficient within CG (Pw), inbreeding coefficient (IC) per sire.

TL	NS	NAIS	r_{ij}	CD	PEV	NF	NFw	NCG	NP	Pw	IC
0.010	51	0	0.028[a]	0.540[a]	40.3[a]	42	17	3[b]	9[b]	0.063[c]	0.158[b]
	56	1	0.040[b]	0.502[b]	39.7[a]	33	18	3[b]	9[b]	0.060[bc]	0.071[a]
	66	5	0.021[c]	0.564[c]	35.3[ab]	38	22	5[b]	24[b]	0.046[bc]	0.064 [a]
	78	0	0.024[d]	0.500[b]	41.9[a]	35	15[b]	2[b]	5[b]	0.058[c]	0.162[b]
	146	1	0.023[c]	0.518[d]	39.3[a]	38	20	2[b]	9[b]	0.106[a]	0.180[b]
	228	74	0.016[e]	0.586[e]	33.2[b]	42	25[a]	11[a]	64[a]	0.036[b]	0.044 [a]
0.007	410	1	0.014[a]	0.519[a]	41.1[a]	39	18[b]	2[b]	7[b]	0.076[b]	0.166[b]
	664	119	0.012[b]	0.543[b]	37.4	37	21[a]	7[a]	39[a]	0.038[a]	0.052[a]
0.005	1236	128	0.010	0.524	39.3	37	19	5	24	0.052	0.089

[1]Means with different upper case letters are significantly different with a p<0.05. Test for multiple comparisons are adjusted with Bonferroni test.

4.0 References

Kennedy, B.W., and D. Trus, 1993. Considerations on genetic connectedness between management units under an animal model. J. Anim. Sci., 71: 2341–2352.

Kuehn, L.A., R.M. Lewis, and D.R. Notter, 2009. Connectedness in Targhee and Suffolk flocks participating in the U.S. National Sheep Improvement Program. J. Anim. Sci., 87: 507-515.

Laloë, D., 1993. Precision and information in linear models of genetic evaluation. Genet. Sel. Evol., 25: 557–576.

Lewis, R.M., R.E. Crump, G. Simm, and R. Thompson, 1999. Assessing connectedness in across-flock genetic evaluations. Proc. Brit. Soc. Anim. Sci. 121.

Salaris, S., Sara Casu, P. Fresi, and A. Carta, 2008. Effect of combining controlled natural mating with artificial insemination on the genetic structure of the flock book of Sardinian breed sheep. In Proc. ICAR 36th Session, Technical Series n. 13, pages 113-121.

Salaris, S., Sara Casu, M.G. Usai, P. Fresi, and A. Carta, 2010. Evaluating the connectedness of a dairy sheep selected population with low rates of artificial insemination and controlled natural mating. In Proc. 9th WGLAP (in press).

Tosh, J.J., and J.W. Wilton, 1994. Effects of data structure on variance of prediction error and accuracy of genetic evaluation. J. Anim. Sci. 72: 2568-2577.