# Bayesian segregation analysis of test-day milk yield in Tunisian Sicilo-Sarde dairy sheep

## H. Ilahi<sup>1,3</sup> and M.H. Othmane<sup>2</sup>

<sup>1</sup> Université de Carthage, Institut Préparatoire aux Etudes d'Ingénieurs de Bizerte 7021 Zarzouna, Tunisia <sup>2</sup> Centre Régional des Recherches Agricoles de Sidi Bouzid BP 357, 9100 Sidi Bouzid, Tunisia

(Received 13 December 2010; revised version 1 April 2011; accepted 18 June 2011)

#### ABSTRACT

Segregation analyses were performed using a Monte Carlo Markov Chain (MCMC) method implemented using Gibbs Sampling. A total of 8387 test-day milk records obtained from the Sicilo-Sarde dairy flock of the Tunisian National Institute of Agricultural Research (INRAT) were analysed. The major gene was assumed to be purely additive autosomal biallelic (A and B) locus with Mendelian transmission probabilities and priors used for the variance components were uniform. Based on 50 000 Gibbs samples from a chain of 520 000 cycles, the estimated marginal posterior means  $\pm$  posterior standard deviations of variance components of test-day milk yield were  $94.82 \pm 136.60$ ,  $1892.65 \pm 1275.21$ ,  $5965.72 \pm 1166.13$  and  $15979.40 \pm 251.86$  for major gene variance  $(\sigma_{g}^{2})$ , polygenic variance  $(\sigma_{u}^{2})$ , permanent environmental variance  $(\sigma_{pe}^{2})$  and error variance  $(\sigma_a^2)$ , respectively. Results showed that the postulated major gene was not significant, the 95% highest posterior density regions (HPDRs) of most major gene parameters included 0, and particularly for the major gene variance. The 95% HPDR for Mendelian transmission probabilities showed that probabilities for 3 genotypes were overlapping. Results indicated that segregation of a major gene was unlikely and that the genetic determinism of test-day milk in Tunisian Sicilo-Sarde dairy sheep is purely polygenic. The estimated polygenic heritability and repeatability were  $h^2 = 0.07$  and r = 0.32.

KEY WORDS: Bayesian segregation, test-day milk, genetic parameters, major gene, Sicilo-Sarde dairy sheep

<sup>&</sup>lt;sup>3</sup> Corresponding author: e-mail: houcine.ilahi@ipeib.rnu.tn

#### **INTRODUCTION**

Dairy sheep have been traditionally reared in Mediterranean countries in a wide range of local production systems combined to a great genetic diversity (Boyazoglu and Flamant, 1990). Most Mediterranean dairy sheep are dual purpose, with incomes originating from both meat and milk, with milk being processed into high quality cheese. The most frequent breeding strategy corresponds to the selection of the local breeds for milk yield and composition within their specific area and production system (Barillet, 1997).

Applying marker assisted selection (MAS) or gene-assisted selection (GAS), either to speed up selection of routinely measured traits or to implement selection for traits costly to record, seems to be attractive in dairy sheep. Before implementing MAS, GAS or starting to search for quantitative trait loci (QTL), the identification of a major gene based on statistical segregation analyses using only phenotypic data would be more informative, cheaper, and more useful as compared to MAS or GAS.

Therefore, the existence of major genes has been investigated in several studies in livestock species: Janss et al. (1995) for various traits of Dutch Meishan crossbreds, Ilahi (1999) and Ilahi et al. (2000) for milking speed in dairy goats, Pan et al. (2001) for somatic cell scores in dairy cattle, Hagger et al. (2004) for selection response in laying hens, Ilahi and Kadarmideen (2004) for milk flow in dairy cattle, Ilahi and Othmane (2011) for total milk yield in dairy sheep.

Marker-free segregation analysis is the most powerful statistical method to identify major gene using only phenotypic data and without DNA marker information (e.g., Guo and Thompson, 1992; Janss et al., 1995, 1997). Inbreeding and marriage loops in a typical segregation analysis using pedigreed populations make the exact computations of likelihoods or marginal densities impossible. This problem has been overcome by the development of Bayesian method implemented *via* Gibbs sampling, a Monte Carlo Markov chain (MCMC) methodology (Guo and Thompson, 1992) and the application of this methodology to livestock populations: Sorensen et al. (1994), Janss et al. (1997), Hagger et al. (2004), Ilahi and Kadarmideen (2004) and Ilahi and Othmane (2011).

The objectives of this study were: 1. to investigate whether a segregating major gene influences the test-day milk yield, and 2. to estimate the genetic parameters of test-day milk yield in Tunisian Sicilo-Sarde dairy sheep using Bayesian segregation analysis methodology.

#### ILAHI H., OTHMANE M.H.

## MATERIAL AND METHODS

#### Data

The analysed data contained 8387 test-day milk records (with mean  $\pm$  standard deviation: 553.5 g  $\pm$  189.2) from 303 ewes and were obtained from the Sicilo-Sarde dairy flock of the Tunisian National Institute of Agricultural Research (INRAT), between December 1989 and June 2002. For experimental purposes according to the institute policy in milk recording, all ewes were on the twice-a-day twice monthly recording plan of testing, with the first test-day beginning at week 7 post-partum and the subsequent records obtained at two-weekly intervals thereafter. All ewes were milked twice a day at 08.00 and 16.00. The mean number of test days per lactation was 9.3, and each ewe averaged 2.9 lactations. The total number of sheep in pedigree was 452, of which 363 were ewes and 89 were rams. All rams were used for natural service under good pedigree control (three generations). The numbers of levels of all effects included in the model are given in Table 1.

Table 1. Number of levels of all effects used in the analyses

Effects	Number of levels		
Stage of lactation	12		
Parity	8		
Type of lambing	2		
Test date	113		
Permanent environmental	303		
Polygenic	452		

## Statistical model for segregation analysis

To detect the existence of major genes for test-day milk yield in Sicilo-Sarde dairy sheep, the following mixed inheritance model was applied:

$$y = X\beta + Zu + Qpe + ZWm + e$$

where: y - the vector of test-day milk records;  $\beta$  - the vector of fixed effects including: stage of lactation, parity, lambing type (single and multiple), and test-date; u - the vector of random additive polygenic effects; pe - the vector of random permanent environmental effects; m - the vector of genotype means (i.e. *a*, 0, -*a*); e - the vector of random residual effects; X<sub>(8387, 135)</sub>, Z<sub>(8387, 452)</sub> and Q<sub>(8387, 303)</sub> are incidence matrices relating the effects to their respective observations; and W<sub>(452, 3)</sub> - a matrix that contains the genotype of each individual (i.e. AA, Aa, aa). W and m are unknown and have to be estimated from data by using segregation analysis.

The major gene was modelled as an additive autosomal biallelic *locus* (*A* and *a*) with Mendelian transmission probabilities. The allele *A* is defined to increase the phenotypic value (favourable allele), and allele *a* is defined to decrease the phenotypic value. With these two alleles *A* and *a*, with frequencies *p* and q = 1 - p, where *p* is the estimate of *A* allele frequency in the founder population in which the Hardy-Weinberg equilibrium was assumed, three genotypes *AA*, *Aa* and *aa* can be encountered, with genotype means m = (*a*, 0, -*a*), where *a* is the additive major gene effect.

Distributional assumptions for random effects were:

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{A}_{\sigma_u^2}), \mathbf{pe} \sim N(\mathbf{0}, \mathbf{I}_{\sigma_{ee}^2}) \text{ and } \mathbf{e} \sim N(\mathbf{0}, \mathbf{I}_{\sigma_{e}^2})$$

where: A - the numerator relationship matrix, I - the identity matrix, and  $\sigma_u^2$ ,  $\sigma_{pe}^2$  and  $\sigma_e^2$  - polygenic, permanent environmental and residual variances, respectively.

The variance attributable to the major gene ( $\sigma_{G}^{2}$ ) was estimated as:

$$\sigma_G^2 = 2p(1-p)a^2.$$

Uniform prior distributions were assumed in the range  $(-\infty, +\infty)$  for nongenetic effects and effects at the major locus, in the range  $(0, +\infty)$  for variance components, and in the range (0, 1) for allele frequencies (Janss et al., 1995).

Gibbs sampling algorithm with blocked sampling of genotypes W was used for inference in the mixed inheritance model and implemented using the 'iBay' software package version 1.46 developed by Janss (2008). A single run of the Monte Carlo Markov Chain (MCMC) consisted of 520 000 samples, with the first 20 000 samples used for burn-in period to allow the Gibbs chains to reach equilibrium. Thereafter each 10<sup>th</sup> sample was collected to obtain 50 000 samples in total.

Marginal posterior densities of the following parameters were directly estimated in each Gibbs cycle from the mixed general model: variance components  $\sigma_u^2$ ,  $\sigma_{pe}^2$ ,  $\sigma_e^2$  and  $\sigma_g^2$  additive effect at the major gene *a*, allele frequency *p*, and the Mendelian transmission probabilities. The heritabilities and repeatabilities were calculated, using polygenic and single locus variance component estimates (Janss, 2008), as:

for heritability and repeatability:

$$h^{2} = \frac{\sigma_{u}^{2}}{\sigma_{u}^{2} + \sigma_{pe}^{2} + \sigma_{e}^{2}} ; r = \frac{\sigma_{u}^{2} + \sigma_{pe}^{2}}{\sigma_{u}^{2} + \sigma_{pe}^{2} + \sigma_{e}^{2}}$$

for total heritability and repeatability:

$$h_t^2 = \frac{\sigma_u^2 + \sigma_G^2}{\sigma_u^2 + \sigma_G^2 + \sigma_{pe}^2 + \sigma_e^2} ; r_t = \frac{\sigma_u^2 + \sigma_G^2 + \sigma_{pe}^2}{\sigma_u^2 + \sigma_G^2 + \sigma_{pe}^2 + \sigma_e^2}$$

#### ILAHI H., OTHMANE M.H.

The genetic parameters under polygenic model for test-day milk yield were estimated using the VARCOMP procedure of SAS (2003) in a preliminary analysis.

## **RESULTS AND DISCUSSION**

Estimates of genetic parameters under polygenic model for test-day milk yield are shown in Table 2.

Table 2. Estimated variance components of test-day milk yield in Sicilo-Sarde dairy sheep using polygenic model

Parameters	Estimates
Error variance	15975.40
Polygenic variance	1739.65
Permanent variance	5782.80
Polygenic heritability	0.07
Repeatability	0.32

Results from segregation analyses of test-day milk yield, using mixed inheritance model to infer polygenic effects and effect of single major gene, are given in Tables 3 and 4. These Tables showed the estimates of posterior means and standard deviations of marginal distributions of all fitted parameters in the model. These estimates are based on 50 000 Gibbs samples. Posterior marginal distributions of all variance components of test-day milk yield are shown in Figure 1.

Table 3. Estimated marginal posterior means and standard deviations for fitted parameters from mixed inheritance model and left and right 95% highest posterior density regions (HPDRs<sub>95%</sub>) for test day milk yield in Sicilo-Sarde dairy sheep, based on 50 000 Gibbs samples

Estimates of parameters	Posterior means	Posterior	HPDRs <sub>95%</sub>	
		standard - deviations	left	right
Error variance	15979.40	251.86	15488.45	16483.64
Polygenic variance	1892.65	1275.21	18.80	4322.06
Permanent environmental variance	5965.72	1166.13	3594.86	8181.42
Major gene variance	94.82	136.60	0.00	358.35
Additive major gene effect	12.52	9.32	0.00	30.40
Frequency of allele A	054	0.22	0.15	0.95
Heritability	0.08	0.06	-	-
Repeatability	0.33	0.02	-	-

Table 4. Estimated posterior means, left and right 95% highest posterior density regions (HPDRs<sub>95%</sub>) for transmission probabilities from mixed inheritance model for test-day milk yield in Sicilo-Sarde dairy sheep, based on 50 000 Gibbs samples

Transmission probability <sup>1</sup>	Posterior means	HPDRs <sub>95%</sub>	
		left	right
$\Pr(A AA)$	0.89	0.66	1.00
$\Pr(A Aa)$	0.71	0.35	0.99
$\Pr(A aa)$	0.42	0.00	0.81

<sup>1</sup> transmission probabilities, presented as the probabilities to inherit *A* allele from *AA*, *Aa*, and *aa* genotypes (Elston and Stewart, 1971)



Figure 1. Marginal posterior distributions of error variance, polygenic variance, permanent environmental variance and major gene variance from mixed inheritance model of test-day milk yield in Sicilo-sarde dairy sheep

According to Box and Tiao (1973), the highest posterior density regions (HPDRs), based on a non-parametric density estimate, using the averaged shifted histogram technique (Scott, 1992), were obtained for all model parameters. These highest regions were constructed to include the smallest possible region of each sampled parameter value. In our analyses, the highest posterior density regions at 95% (HPDs<sub>95%</sub>) of the additive gene effect (a) and the variance at the major locus  $(\sigma_{g}^{2})$  included zero (Table 3). The allele frequencies in the analysed population were intermediate (p=0.54 and q = 1 - p = 0.46). The polygenic variance of test-day milk yield ( $\sigma_u^2 = 1892.65$  $\pm$  1275.21) was significantly higher than the major gene variance ( $\sigma_e^2 = 94.82$  $\pm$  136.60). Janss et al. (1997) and Miyake et al. (1999) suggested the use of the magnitude of the major gene variances as an indicator for the existence of major gene segregation. Following Elston (1980), the evidence of a significant segregating major gene in quantitative trait requires three conditions: statistical significance of the major gene component in the model, statistical differences among the transmission probabilities and these transmission probabilities are significantly different from an environmental model.

In order to check the statistical significance of the major gene component in the model Janss (1998) proposed, to check the 95% highest posterior density region (HPDR<sub>95%</sub>) of the postulated major gene variance: if the HPDR<sub>95%</sub> doses not include zero (the postulated major gene is statistically significant) or includes zero (not significant).

The Mendelian transmission (probabilities 1, 1/2, and 0) was tested by checking if the highest posterior density regions at 95% (HPDRs<sub>95%</sub>) were overlapped or not. Mendelian transmission probabilities for the 3 genotypes were estimated (Table 4) as suggested by Elston and Stewart (1971). These probabilities were parameterised to indicate the Mendelian transmission of the favourable allele, with probabilities of *A* allele transmission of 1, 1/2, and 0 for genotypes *AA*, *Aa*, and *aa*, respectively.

Table 4 showed that the three estimated posterior means of Mendelian transmission probabilities were not significantly different, and as well, their highest posterior density regions at 95% (HPDRs<sub>95%</sub>) for the three genotypes were overlapped.

Furthermore, the density of marginal posterior distribution for the major gene variance (Figure 1) was unimodal marginal density with mode = 0, suggested the absence of a major gene for the analysed trait (Janss et al., 1995; Pan et al., 2001).

Based on these results obtained from Bayesian segregation analysis using only phenotypic data sets, we can conclude that the postulated major gene was not significant and the inheritance of test-day milk yield in Sicilo-Sarde dairy sheep is purely polygenic. The estimates of heritability and repeatability are consistent across models (polygenic and mixed models; Tables 2 and 3). This finding confirmed again that the postulated major gene is not significant on test-day milk yield in Sicilo-Sarde dairy sheep.

The heritability estimated in the present study for test-day milk yield trait  $(h^2=0.07)$  is on the lower limit of those reported by Gutierrez et al. (2007) in Assaf breed and Serrano et al. (2003) in Manchega ewes and Othmane et al. (2002a) in Churra breed. This low heritability for test-day milk yield in Sicilo-Sarde dairy sheep in comparison with others dairy breeds could be explained by the existence of selection and breeding programmes well established and more advanced in European breeds and their high level of milk production (Othmane et al., 2002b).

Repeatability estimated for milk yield (r = 0.32) was consistent with those reported for the same breed (Othmane, 2004) and other dairy ewes (Gutierrez et al. (2007) in Assaf breed, Baro et al. (1994) and Othmane et al. (2002a) in Churra breed. This estimate, together with the low heritability value, suggested that milk yield in the studied flock is more influenced by some permanent environmental factors rather than by genetic effects; perhaps because of a lack in genetic connection with other flocks of the same breed (Othmane, 2004).

### CONCLUSIONS

This study was the first attempt on this topic in Tunisian Sicilo-Sarde dairy sheep. Test day milk yield had low heritability and a repeatability value within the range of those reported for other dairy sheep breeds. No existence of a major gene was observed and the mode of inheritance for test day milk yield in Sicilo-Sarde breed is then polygenic. However, more research using both molecular information and a more consistent phenotypic data set would be useful to investigate the evidence of segregating major genes affecting milk yield in such a breed, since segregation analysis using molecular information is more powerful.

#### ACKNOWLEDGMENTS

Authors thank Dr. Luc L.G. Janss for supplying the 'iBay" software used for this research.

#### REFERENCES

- Barillet F., 1997. Genetics of milk production,. 20. In: I. Piper, A. Ruvinsky (Editors). The Genetics of Sheep. CAB International, pp. 539–564
- Baro J.A., Carriedo J.A., San Primitivo F., 1994. Genetic parameters of test-day measures for somatic cell count, milk yield and protein percentage of milking ewes. J. Dairy Sci. 77, 2658-2662
- Box G.E.P., Tiao G., 1973. Bayesian Inference in Statistical Analysis. Reading Addison-Wesley
- Boyazoglu J.G., Flamant J.C., 1990. Mediterranean Systems of Animal Production. In: J.G. Galaty, D.L. Johnson (Editors). The World of Pastoralism, pp. 353-393
- Elston R.C., 1980. Segregation Analysis. Current Developpements in Anthropological Genetics, vol. 1. J.H. Mielke, M.H. Crawford (Editors). Plenum Publishing Corporation, pp. 327-354
- Elston R.C., Stewart J.M., 1971. A general model for the genetic analysis of pedigree data. Hum. Hered. 21, 523-542
- Guo S.W., Thompson E.A., 1992. Monte Carlo method for combined segregation and linkage analysis. Amer. J. Hum. Genet. 51, 1111-1126
- Gutiérrez J.P., Legazb E., Goyache F., 2007. Genetic parameters affecting 180-days standardised milk yield, test-day milk yield and lactation length in Spanish Assaf (Assaf.E) dairy sheep. Small Ruminant Res.70, 233-238
- Hagger C., Janss L.L.G., Kadarmideen H.N., Strazinger G., 2004. Bayesian inference on major loci in related multigeneration selection lines of laying hens. Poultry Sci. 83, 1932-1939
- Ilahi H., 1999. Variabilité génétique de débit de traite chez les caprins laitiers. Thèse de doctorat. ENSA- Rennes (France)
- Ilahi H., Kadarmideen H.N., 2004. Bayesian segregation analysis of milk flow in Swiss dairy cattle using Gibbs sampling. Genet. Sel. Evol. 36, 563-576
- Ilahi H., Manfredi E., Chastin P., Monod F., Elsen J.M., Le Roy P., 2000. Genetic variability in milking speed of dairy goats. Genet. Res. 75, 315-319
- Ilahi H., Othmane M.H., 2011. Complex segregation analysis of total milk yield in Churra dairy ewes. Asian-Austr. J. Anim. Sci. 24, 330-335
- Janss L.L.G., 1998. "MAGGIC" a package of subroutines for genetic analyses with Gibbs sampling. In: Proceedings of 6<sup>th</sup> World Congress on Genetics Applied to Livestock Production (WCGALP). University of New England, Armidale (Australia) 27, 459-460
- Janss L.L.G., 2008. "iBay Manual Version 1.46". Janss Bioinformatics. Lieden (The Netherlands)
- Janss L.L.G., Thompson R., Van Arendonk J.A.M., 1995. Application of Gibbs sampling for inference in a mixed major gene-polygenic inheritance model in animal populations. Theor. Appl. Genet. 91, 1137-1147
- Janss L.L.G., Van Arendonk J.A.M., Brascamp E.W., 1997. Bayesian statistical analyses for presence of single major genes affecting meat quality traits in crossed pig population. Genetics 145, 395-408
- Miyake T., Gaillard C., Moriya K., Sasaki Y., 1999. Accuracy of detection of major genes segregating in outbred population by Gibbs sampling using phenotypic values of quantitative traits. J. Anim. Breed. Genet. 116, 281-288
- Othmane M.H., 2004. Estimation des paramètres génétiques de la production laitière avec le modèle animal de jour de contrôle "test-day model" chez la brebis Sicilo-Sarde. Ann. l'INRAT 77, 171-190
- Othmane M.H., Carriedo J.A., San Primitivo F., De La Fuente L.F., 2002b. Genetic parameters for lactation traits of milking ewes: protein content and composition, fat, somatic cells and individual laboratory cheese yield. Genet. Sel. Evol. 34, 581-596

- Othmane M.H., Fuente L.F., Carriedo J.A., San Primitivo F.S., 2002a. Heritability and genetic correlations of test day milk yield and composition, individual laboratory cheese yield, and somatic cell count for dairy ewes. J. Dairy Sci. 85, 2692-2298
- Pan Y., Boettcher J., Gibson J., 2001. Bayesian segregation analyses of somatic cell scores of Ontario Holstein cattle. J. Dairy Sci. 84, 2796–2802
- SAS, 2003. SAS/STAT User's Guide, Version 9.1. Statistical Analysis System. SAS Institute. Cary, NC
- Scott D.W., 1992. Multivariate Density Estimation. Wiley and Sons, New York
- Serrano M., Perez-Guzman M.D., Montoso V., Jurado J.J., 2003. Genetic analysis of somatic cell counts and milk traits in Manchega ewes. Mean lactation and test-day approaches. Livest. Prod. Sci. 84, 1-10
- Sorensen D., Anderson S., Jensen J., Wang C.S., Gianola D., 1994. Inferences about genetic parameters using the Gibbs sampler. In: Proceedings of the 5<sup>th</sup> World Congress on Genetics Applied to Livestock Production, Guelph (Canada), Vol. 18, 321-328