



Complex Segregation Analysis of Total Milk Yield in Churra Dairy Ewes

Houcine Ilahi* and M. Houcine Othmane¹

Université 7 Novembre de Carthage, Institut Préparatoire aux Etudes d'Ingénieurs de Bizerte, 7021 Zarzouna, Tunisia

ABSTRACT : The mode of inheritance of total milk yield and its genetic parameters were investigated in Churra dairy sheep through segregation analyses using a Monte Carlo Markov Chains (MCMC) method. Data which consisted of 7,126 lactations belonging to 5,154 ewes were collected between 1999 and 2002 from 15 Spanish Churra dairy flocks. A postulated major gene was assumed to be additive and priors used for variance components were uniform. Based on 50 000 Gibbs samples from ten replicates chains of 100,000 cycles, the estimated marginal posterior means±posterior standard deviations of variance components of milk yield were 23.17±18.42, 65.20±25.05, 120.40±42.12 and 420.83±40.26 for major gene variance (σ_G^2), polygenic variance (σ_u^2), permanent environmental variance (σ_{pe}^2) and error variance (σ_e^2), respectively. The results of this study showed the postulated major locus was not significant, and the 95% highest posterior density regions (HPDs_{95%}) of most major gene parameters included 0, and particularly for the major gene variance. The estimated transmission probabilities for the 95% highest posterior density regions (HPDs_{95%}) were overlapped. These results indicated that segregation of a major gene was unlikely and that the mode of inheritance of total milk yield in Churra dairy sheep is purely polygenic. Based on 50,000 Gibbs samples from ten replicates chains of 100,000 cycles, the estimated polygenic heritability and repeatability were $h^2 = 0.20 \pm 0.05$ and $r = 0.34 \pm 0.06$, respectively. (**Key Words :** Segregation Analysis, Bayesian Method, Major Gene, Genetic Parameters, Milk Yield, Dairy Sheep)

INTRODUCTION

A genetic improvement program for milk yield and composition in dairy sheep is an important component toward the development of a viable industry. To remain competitive, the industry needs to increase its productivity and a genetic program is still the best alternative for improving the biological efficiency of producing milk for cheese.

Marker-assisted selection (MAS) has been shown to increase genetic improvement via reduction of generation interval or by increasing the selection intensity in an "outbred" population (Schrooten et al., 2005). Before applying MAS, or starting to search for quantitative trait loci (QTL), identification of a major gene based on statistical segregation analyses using only phenotypic data would be informative, cheap and useful, as compared to MAS. Therefore, the existence of major genes has been

investigated in several studies in livestock species: by Janss et al. (1995) for various traits of Dutch Meishan crossbreds; by Ilahi (1999) and Ilahi et al. (2000) for milking speed in dairy goats; by Pan et al. (2001) for somatic cell scores in dairy cattle; by Hagger et al. (2004) for selection response in laying hens; and by Ilahi and Kadarmideen (2004) for milk flow in dairy cattle.

Segregation analysis using pedigreed animal populations is impossible by analytical approaches due to the existence of many (inbreeding) loops and due to the family sizes, which do not allow to sum and integrate out genotypes and polygenic effects from the likelihood or posterior density. This problem has been simplified by the development of Gibbs sampling, a Monte Carlo Markov chain (MCMC) methodology (Guo and Thompson, 1992), and its applications to livestock populations by Sorensen et al. (1994), Janss et al. (1995), Janss et al. (1997) and Ilahi and Kadarmideen (2004).

The aim of this study was to investigate whether a segregating major gene affects the total milk yield trait, and to estimate its genetic parameters in Churra dairy sheep using Bayesian segregation analysis methodology.

* Corresponding Author: Houcine Ilahi. Tel: +216-96-800-331, Fax: +216-72-593-265, E-mail: houcine.ilahi@ipeib.rnu.tn

¹ Centre Régional des Recherches Agricoles de Sidi Bouzid, BP 357, 9100 Sidi Bouzid, Tunisia.

Received August 7, 2010; Accepted November 15, 2010

MATERIALS AND METHODS

Data

The Churra sheep is an autochthonous breed raised in Castile and Leon, in north-western Spain. It is a milk production breed of great hardiness, well suited to the continental climate of Castile and Leon, with long, severe winters, very short springs, and hot dry summers.

The data available for this study consisted of 7,126 records of total milk yield (with mean±standard deviation: 94.1 kg±35.6) from 5,154 ewes, and were collected between 1999 and 2002 from 15 Spanish Churra dairy flocks. All pedigree information available was included in the analyses. Thus, the pedigree included 8,053 animals with 448 different sires used within the pedigree.

Statistical model

To analyze the presence of a major gene for milk yield in Churra dairy sheep, the following mixed inheritance model was used:

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

where *y* is the vector of observations, β is a vector of non-genetic fixed effects including: flock-year-season, age at the first lambing, lactation number (parity) and type of lambing, *u* is a random vector of individual polygenic effects, *pe* is a random vector of permanent environmental effects, *W* is a design matrix that contains the genotype of each individual (i.e., AA, AB, BB), *m* is the vector of genotype means (i.e., -*a*, 0, *a*), *e* is a random vector of residual effects, and *X*, *Z* and *Q* are incidence matrices relating the observations to their respective effects. In the term modeling the single gene, both *W* and *m* are unknown and have to be estimated from data by using segregation analysis.

The number of levels of all effects included in the model and the number of animals in the pedigree are summarized in Table 1.

The major gene was modelled as an additive autosomal biallelic (*A* and *B*) locus with Mendelian transmission probabilities. Allele *A* is defined to decrease the phenotypic value and allele *B* is defined to increase the phenotypic

value (or favourable allele). With these two alleles *A* and *B*, 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*, *AB* or *BA* and *BB* can be encountered, with genotype means *m* = (-*a*, 0, *a*), where *a* is the additive major gene effect.

Distributional assumptions for polygenic effects were, *u* ~ *N*(0, *A* σ_u^2), where *A* is the numerator relationship matrix. The distribution of the permanent environmental effects were, *pe* ~ *N*(0, *I* σ_{pe}^2). Residual effects were assumed to be distributed as *e* ~ *N*(0, *I* σ_e^2). σ_u^2 , σ_{pe}^2 and σ_e^2 are polygenic, permanent environmental and residual variances, respectively. The relationship matrix of the full pedigree *A* was used in the analyses. The variance attributable to the major gene (σ_G^2) was calculated as: $\sigma_G^2 = 2p(1-p)a^2$.

Uniform prior distributions were assumed in the range (-∞, +∞) for non-genetic effects and effects at the major locus, in the range (0, +∞) for variance components, and in the range [0, 1] for allele frequencies (Janss et al., 1995).

A 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).

Ten replicates of the Monte Carlo Markov Chains (MCMC) of 100,000 Gibbs samples were run, using a spacing of 20 cycles, obtaining 5,000 Gibbs samples per chain and 50,000 samples in total. A burn-in period of 20,000 cycles was used to allow the Gibbs chains to reach equilibrium.

From the mixed general model, marginal posterior densities of the following parameters were directly estimated in each Gibbs cycle: variance components σ_G^2 , σ_u^2 , σ_{pe}^2 , and σ_e^2 , additive effect at the major gene *a*, allele frequency *p*, and the Mendelian transmission probabilities. Using variance components for polygenes and major genes, following Janss (2008), the heritabilities and repeatabilities were calculated 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:

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

Table 1. Description of data sets used for the analysis

Effects	Number of levels
Flock-year-season	72
Age at first kidding	8
Parity	6
Type of lambing	2
Permanent environmental effects	5,154
Animal in the pedigree	8,053

Table 2. Estimated marginal posterior means and marginal posterior standard deviations for variance components from a polygenic model and left and right 95% highest posterior density regions (HPDs_{95%}) for total milk yield in Churra dairy sheep, based on 50,000 Gibbs samples

Parameters	Posterior means	Posterior standard deviations	Left HPD _{95%}	Right HPD _{95%}
Error variance (σ_e^2)	424.19	38.43	303.48	565.13
Polygenic variance (σ_u^2)	129.17	39.38	49.44	289.52
Permanent variance (σ_{pe}^2)	91.50	48.75	3.20	260.95
Polygenic heritability (h^2)	0.20	0.05	-	-
Repeatability (r)	0.34	0.06	-	-

RESULTS AND DISCUSSION

In a preliminary analysis based on the same statistical model used in the Bayesian segregation analysis described above, the variance components under a polygenic model of total milk yield in Churra dairy sheep were estimated by Bayesian analysis, using the 'iBay' software package version 1.46 as well (Janss, 2008). These estimates were based on 50,000 Gibbs samples from 10 replicated chains of 100,000 cycles. The effective sample size was 33,602 Gibbs samples. These estimates are presented in Table 2.

Marginal posterior means and standard deviations of parameter estimates of total milk yield in Churra dairy sheep, using Bayesian segregation analyses implemented by Gibbs sampling, are presented in Table 3 and 4. These estimates were based on 50,000 Gibbs samples from 10 replicated chains. The effective sample size was 27,319 Gibbs samples. Posterior marginal distributions of all variance components of total milk yield are shown in Figure 1.

According to Box and Tiao (1973), the highest posterior density regions (HPD), 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_{95%}) of the additive gene effect (a) and the variance at the major locus (σ_G^2) included zero (Table 3). The allele frequencies in the analysed population were intermediate ($p = 0.56$ and $q = 1-p = 0.44$). The polygenic variance ($\sigma_u^2 = 65.20 \pm 25.05$) was significantly higher than the major gene variance ($\sigma_G^2 = 23.17 \pm 18.42$). Janss et al. (1997) and Miyake et al. (1999) also suggested the use of the magnitude of the major gene variances as an indicator for the existence of segregating a major gene. Following Elston (1980), the evidence of a significant segregating major gene in a 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 (HPD_{95%}) of the postulated major gene variance: if the 95% HPD does not include zero (the postulated major gene is statistically significant) or includes zero (not significant).

Table 3. Estimated marginal posterior means and marginal posterior standard deviations for fitted parameters from a mixed general model and left and right 95% highest posterior density regions (HPDs_{95%}) for total milk yield in Churra dairy sheep, based on 50,000 Gibbs samples from 10 replicated chains

Parameters	Posterior means	Posterior standard deviations	Left HPD _{95%}	Right HPD _{95%}
Error variance (σ_e^2)	420.83	40.26	348.90	504.58
Polygenic variance (σ_u^2)	65.20	25.05	1.70	156.90
Permanent variance (σ_{pe}^2)	120.40	42.12	3.41	228.70
Major gene variance (σ_G^2)	23.17	18.42	0.00	102.25
Additive effect (a)	6.45	3.90	0.00	17.83
Frequency of major gene (p)	0.56	0.20	0.08	0.98
Heritability (h^2)	0.10	0.02	-	-
Repeatability (r)	0.30	0.04	-	-

Table 4. Estimated marginal posterior means, left and right 95% highest posterior density regions (HPDs_{95%}) for transmission probabilities using a mixed general model for total milk yield in Churra dairy sheep, based on 50,000 Gibbs samples from 10 replicated chains

Transmission probability*	Posterior means	Left HPD _{95%}	Right HPD _{95%}
Pr(B BB)	0.71	0.33	1.00
Pr(B BA)	0.41	0.07	0.77
Pr(B AA)	0.20	0.00	0.53

* Transmission probabilities, presented as the probabilities to inherit a B allele from BB, BA, and AA genotypes (Elston and Stewart, 1971).

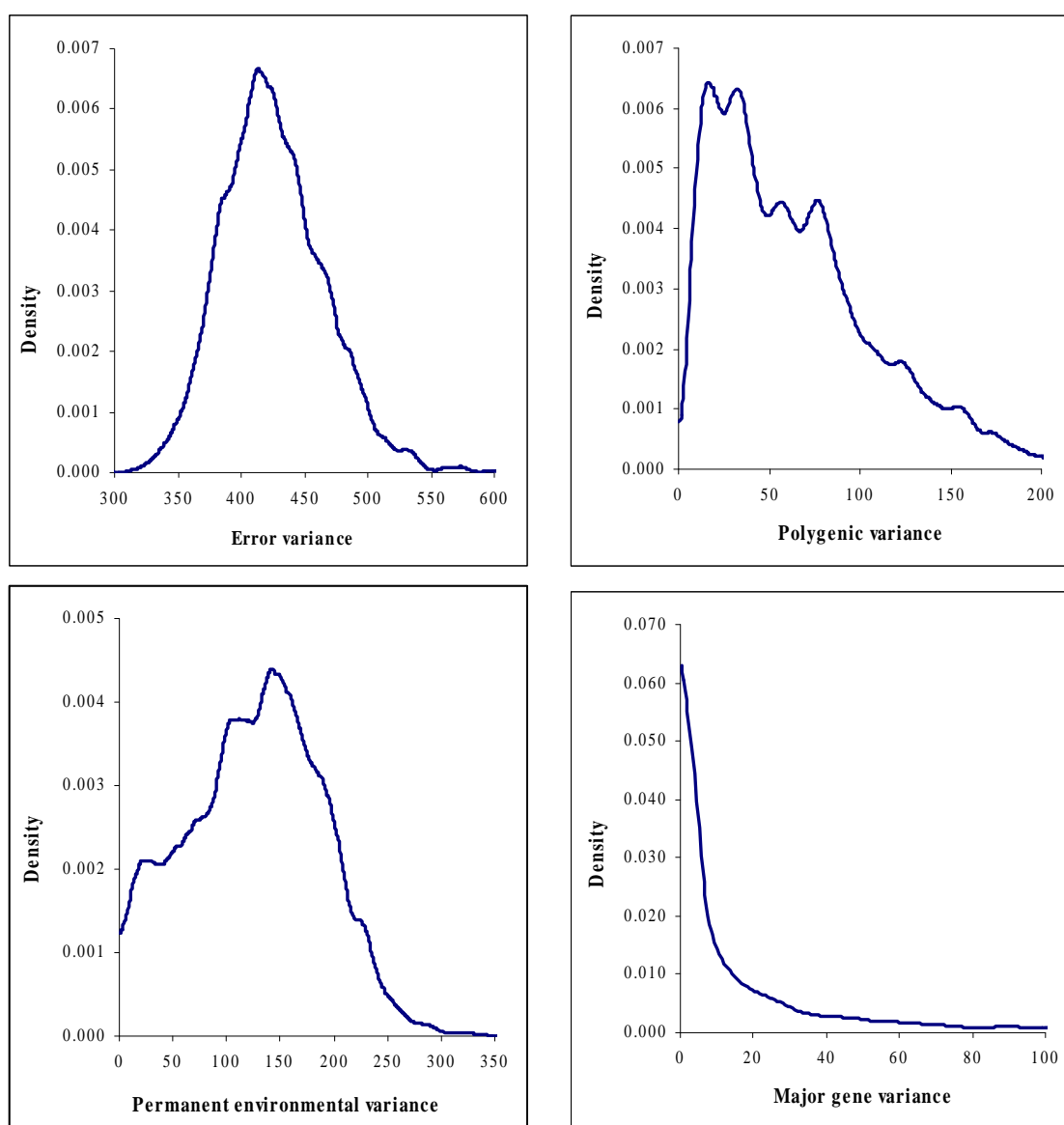


Figure 1. Marginal posterior distributions of error variance, polygenic variance, permanent environmental variance and major gene variance from a mixed general model of total milk yield in Churra dairy sheep.

The Mendelian transmission (probabilities 1, 1/2, and 0) was tested by checking if the highest posterior density regions at 95% (HPDs_{95%}) 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 B allele transmission of 1, 1/2, and 0 for genotypes BB, BA, and AA, respectively.

Table 4 showed that the three estimated posterior means of Mendelian transmission probabilities were not significantly different, and their highest posterior density regions at 95% (HPDs_{95%}) for the three genotypes were also overlapped.

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 genetic determinism of total milk yield in Churra dairy sheep is polygenic.

Estimated heritability and repeatability of total milk yield obtained by Bayesian segregation under mixed inheritance model analysis were lower than those obtained under the polygenic model by Bayesian analysis approach (Tables 2 and 3). This might be explained by: first, there is a proportion of total genetic variance explained by the major gene variance, but it is not statistically significant and then was not taken into account in the estimation of heritability and repeatability. Second, our current results showed that the genetic determinism of total milk yield in Churra dairy sheep is polygenic and we have to use and apply polygenic models to estimate its genetic parameters.

Heritability estimated using a polygenic model $h^2 = 0.20 \pm 0.05$ was within the range of those reported by several studies (Ugarte et al., 1996; Pollott et al., 1998; Othmane et al., 2002a; Serrano et al., 2003; Nikolau et al., 2004; Afolayan et al., 2008), while it was slightly lower than that reported by Othmane et al. (2002b); El-Saied et al. (2005) and Jawasreh and Khasawneh (2007). Repeatability estimated using a polygenic model $r = 0.34 \pm 0.06$ was in accordance with those reported by Othmane et al. (2002b) and Gutiérrez et al. (2007) and lower than the estimates reported by El-Saied et al. (1999) and Gootwine and Pollott (2000).

CONCLUSIONS

Bayesian segregation analysis based only on phenotypic data was applied in order to investigate the mode of inheritance and to estimate the genetic parameters of total milk yield in Churra dairy sheep, by using standard statistical significance testing based on the 95% highest posterior density regions (HPDs_{95%}) and by checking

Mendelian transmission probabilities. The results of this study showed no existence of a major gene and the mode of inheritance for total milk yield in Churra breed is polygenic. The estimated polygenic heritability and repeatability of this trait in Churra are similar to those estimated and reported in the literature in other dairy sheep breed.

Further genetic analyses using both molecular information and phenotypic data to search for evidence of segregating major genes affecting total milk yield in Churra dairy sheep are supported, since segregation analysis using molecular information is much more powerful and can also separate effects of multiple major genes.

ACKNOWLEDGMENTS

The authors thank Dr. L. L.G. Janss for supplying iBay software for the analysis and PhD students in the Animal Production Department of Leon University, Spain, for collecting the data during the experimentation period.

REFERENCES

- Afolayan, R. A., M. N. Fogarty, J. E. Morgan, G. M. Gaunt, L. J. Cummins, A. R. Gilmour and S. Nielsen. 2008. Genetic analysis of milk production and composition in crossbred ewes from different maternal genotypes. *Anim. Prod. Sci.* 49:24-31.
- Box, G. E. P. and G. Tiao. 1973. *Bayesian Inference in Statistical Analysis*, Reading Addison-Wesley.
- Elston, R. C. 1980. Segregation analysis. Current developments in anthropological genetics, vol. 1. Edited by Mielke J. H and Crawford M. H Plenum Publishing Corporation. 327-354.
- Elston, R. C. and J. M. Stewart. 1971. A general model for the genetic analysis of pedigree data. *Hum. Hered.* 21:523-542.
- El-Saied, U. M., J. A. Carriedo and L. F. de la Fuente. 1999. Genetic parameters of lactation cell counts and milk and protein yields in dairy ewes. *J. Dairy Sci.* 82:639-644.
- El-Saied, U. M., L. F. de la Fuente, J. A. Carriedo and F. S. San Primitivo. 2005. Genetic and phenotypic parameter estimates of total and partial lifetime traits for dairy ewes. *J. Dairy Sci.* 88:3265-3272.
- Guo, S. W. and E. A. Thompson. 1992. Monte Carlo method for combined segregation and linkage analysis. *Am. J. Hum. Genet.* 51:1111-1126.
- Gutiérrez, J. P., E. Legazb and F. Goyache. 2007. Genetic parameters affecting 180-days standardised milk yield, test-day milk yield and lactation length in Spanish Assaf (Assaf.E) dairy sheep. *Small Rumin. Res.* 70:233-238.
- Gootwine, E. and G. E. Pollott. 2000. Factors affecting milk production in improved Awassi dairy ewes. *Anim. Sci.* 24:77-83.
- Hager, C., L. L. G. Janss, H. K. Kadarmideen and G. Strazinger. 2004. Bayesian inference on major loci in related multigeneration selection lines of laying hens. *Poult. 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., E. Manfredi, P. Chastin, F. Monod, J. M. Elsen and P. Le Roy. 2000. Genetic variability in milking speed of dairy goats. *Genet. Res.* 75:315-319.
- Ilahi, H. and H. N. Kadarmideen. 2004. Bayesian segregation analysis of milk flow in Swiss dairy cattle using Gibbs sampling. *Genet. Sel. Evol.* 36:563-576.
- Janss, L. L. G., R. Thompson and J. A. M. Van Arendonk. 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., J. A. M. Van Arendonk and E. W. Brascamp. 1997. Bayesian statistical analyses for presence of single major genes affecting meat quality traits in crossed pig population. *Genetics* 145:395-408.
- Janss, L. L. G. 1998. "MAGGIC" a package of subroutines for genetic analyses with Gibbs sampling, in: Proc. 6th World Congr. Genet. Appl. Livest. Prod., 11- 6 January 1998, Vol. 27, University of New England, Armidale, Australia, pp. 459-460.
- Janss, L. L. G. 2008. "iBay manual version 1.46". Janss Bioinformatics, Lieden, Netherlands.
- Jawasreh, K. I. Z. and A. Z. Khasawneh. 2007. Genetic evaluation of milk production traits in Awassi sheep in Jordan. *Egyptian J. Sheep Goat Sci.* 2:83-100.
- Miyake, T., C. Gaillard, K. Moriya and Y. Sasaki. 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.
- Nikolau, M., A. P. Kominakis, E. Rogdakis and S. Zampitis. 2004. Effect of mean and variance heterogeneity on genetic evaluations of Lesbos dairy sheep. *Livest. Prod. Sci.* 88:107-115.
- Othmane, M. H., L. F. Fuente, J. A. Carriedo and F. S. San Primitivo. 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.
- Othmane, M. H., J. A. Carriedo, F. S. San Primitivo and L. F. De La Fuente. 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.
- Pan, Y., J. Boettcher and J. Gibson. 2001. Bayesian segregation analyses of somatic cell scores of Ontario Holstein cattle. *J. Dairy Sci.* 84:2796-2802.
- Pollott, G. E., O. Gursoy and K. Kirk. 1998. The genetics of meat and milk production in Turkish Awassi sheep. Proc 6th World Congress on Genetics Applied to Livestock Production. 12-16, Jan. Armidale, NSW Australia. 24:177-180.
- Schrooten, C., H. Bovenhuis, J. A. M. Van Arendonk and P. Bijma. 2005. Genetic progress in multistage dairy cattle breeding schemes using genetic markers. *J. Dairy Sci.* 88:1569-1578.
- Scott, D. W. 1992. Multivariate density estimation. Wiley and Sons, New York.
- Serrano, M., M. D. Perez-Guzman, V. Montoso and J. J. Jurado. 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., S. Anderson, J. Jensen, C. S. Wang and D. Gianola. 1994. Inferences about genetic parameters using the Gibbs sampler, in: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Guelph, Canada, Vol. 18: 321-328.
- Ugarte, E., E. Urarte, F. Arrese, J. Arranz, L. Silio and C. Rodríguez. 1996. Genetic parameters and trends for milk production of Blond-Faced Latxa sheep using Bayesian analysis. *J. Dairy Sci.* 79:2268-2277.