Bayesian Analysis Of Uncertain Dam Assignments In Lambs

C. Esquivelzeta[*†], J. Piedrafita[*] and J. Casellas[*]

Introduction

Lamb’s parents can be misidentified or confounded due to the plausibility of two or more potential sires or dams. As discussed by Sapp et al. (2007), multiple-sire mating systems under pastoral conditions are a typical example where paternities must be elucidated from a list of several candidate sires. Paternity errors are expected to bias estimations of genetic parameters (Van Vleck 1970b), breeding values (Van Vleck 1970a; Israel and Weller 2000) and even genetic progress (Van Vleck 1970a; Israel and Weller 2000; Banos et al. 2001). This consequently creates a problem in sire evaluation because of the uncertainty in the assignment of progeny to sires (Im 1992). Banos et al. (2001) found a considerable impact on national evaluations and international genetic comparisons when introducing paternity errors to the pedigrees of dairy cows. Although substantial efforts were addressed to elucidate the influence of misidentified paternities in genetic evaluations and to develop appropriate statistical models accounting for this phenomenon, little is known about missing and uncertain dam assignments in livestock. In intensive farming where the herd size is large and birth control becomes difficult, we can have multiple births in a day and the correct assignment of lambs to the real mother can be complicate, generating incomplete pedigrees.

Different Bayesian approaches have been used by several authors for animal genetic evaluation under uncertain paternity. Cardoso and Tempelman (2003) and Foulley et al. (1987) worked with prior information that included know dam identification for all animals and unknown sire identification. Sapp et al. (2007) considered the possibility of one (i.e. sire) or both unknown parents. Im (1992) proposed the use of alternative methods to deal with uncertain paternities, an approach simpler than the Bayesian solution, but they just had good results when the data set was small. Bayes approaches were more efficient for large data sets. In all previous cases, statistical developments modeled unknown paternities by appropriately modifying the numerator relationship matrix and without taking advantage of conditioned relationships with the incidence of systematic effects.

When maternity must be elucidated between two candidate ewes (e.g. two ewes deliver at the same time and both of them ignore the newborn lamb), Sapp’s et al. (2007) model could be adapted, although both genetic and systematic information must be accounted for. Taking birth weight as example, systematic effects like the age of the ewe must influence the

---

[*] Grup de Recerca en Remugants, Departament de Ciència Animals i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain.

[†] Supported by CONACyT (Consejo Nacional de Ciencia y Tecnología, México).

phenotype of its offspring and therefore, provide relevant information for the inference of true maternities. Within this context, the objective of this work was to adapt Sapp's et al. (2007) model for unknown maternities, integrating both genetic and systematic information to ascertain maternity in lambs.

Material and methods

Field data set. Although this research focused on simulation procedures, new data sets were generated by mimicking the pedigree and phenotypic records for birth weight from the experimental Ripollesa flock kept in the Universitat Autònoma of Barcelona (Bellaterra, Spain). Among the autochthonous sheep breeds in Spain, the Ripollesa is the most common breed in Catalonia producing "pascual"-type lambs with 22 to 24 kg of live weight at slaughter (~3 months of age). It is a rustic meat-type breed, with a convex profile, white "entrefino"-type wool and a characteristic pigmentation. Our data set comprised 2,443 birth weight records registered between 1986 and 2007. All relevant information at the lamb (e.g. identification, sire, dam, sex, year of birth) and ewe levels (e.g. ewe age, type of parturition) were accurately recorded. The full pedigree file included 2,856 individuals. See Casellas *et al.* (2009) for a detailed description of this data set.

Simulation procedure. A total of 1,000 independent populations were simulated on the basis of the field data set described above. For each new population, birth weight records were simulated from a multivariate normal distribution, $N(\mathbf{Xb}+\mathbf{Z_1p}+\mathbf{Z_2a},\mathbf{I}\sigma_e^2)$, were **b** was the vector of systematic effects, **p** was the vector of permanent effects, **a** was the vector of additive genetic effects, **X** and **Z** were incidence matrices with appropriate dimensions, **I** was an identity matrix with dimensions equal to the number of phenotypic records, and σ_e^2 was the residual variance. Note that **b** stored four systematic effects: sex of the lamb (male or female), birth type (single, twins, triplets), age of the ewe (1, 2, 3, 4, 5 or more than five years), and year of lambing (from 1986 to 2007). Permanent effects were sampled from $N(0,\sigma_p^2)$ and additive genetic effects for founder individuals were obtained from $N(0,\sigma_a^2)$, where σ_p^2 and σ_a^2 were the permanent environmental and additive genetic variances, respectively. The genetic effects for non-founder individuals were simulated as the average of their parents' breeding values plus the Mendelian sampling term ($N(0,\sigma_\mathrm{a}^2/2)$). Values in vector **b** were estimated in the Ripollesa data set by using the Mixed Model procedure of SAS v.9.1 (SAS Institute, Inc., Cary, NC) and variance components were previously obtained by Casellas et al. (2009). For each simulated data set, maternity identification procedures focused on two single-birth lambs (lamb L_A and lamb L_B) and their true dams (ewe E_A and ewe E_B). Those animals were selected at random from the 141 lambs born during year 2007.

Bayesian analysis. Simulated data sets were analyzed by a Bayesian mixed linear model. Bayesian likelihood was defined as mulivariate normal, $N(\mathbf{Xb}+\mathbf{Z_1p}+\mathbf{Z_2a},\mathbf{I}\sigma_e^2)$, a priori distributions for **a** and **p** effects were $N(0,\mathbf{A}\sigma_a^2)$ and $N(0,\mathbf{I}\sigma_p^2)$, respectively (**A** was the numerator relationship matrix and **I** was an identity matrix with appropriate dimensions), and improper flat priors were assumed for the remaining parameters. It is important to note that the a priori probability for lamb L_A to belong to ewes E_A and E_B was assumed 0.5 and 0.5, respectively. The same assumption was made for lamb L_B. During the sampling process, changes in maternity assignment forced rearrangements in matrices **A** (sire and dam identification), **X** (effect of the age of the dam) and $\mathbf{Z}_1$ (permanent environmental effect characterized by the dam). Samples from the marginal posterior distribution of all parameters were obtained by Gibbs sampling with the exception of the dam that required a Metropolis-Hastings step. A total of 25,000 iterations cycles were run after a burn-in period of 1,000 cycles. The probability to correctly assign each lamb to its own dam (E_*) was calculated as the number of iterations where the Metropolis-Hastings sampling returned E_*, divided by the total number of iterations (i.e. 25,000)

The probability of each lamb to be correctly assigned to its own dam was calculated at the end of the Markov chain Monte Carlo process. Differences in this probability linked to the age of the dams involved in the analysis as well as to lambs from the same or different sires were evaluated by using the Student Newman Keuls multiple range test within the Generalized Linear Models procedure of SAS v.9.1 (SAS Institute, Inc., Cary, NC).

Results and discussion

The average probability to correctly assign each lamb to its own dam was 57.5%, although values ranged between 4.8% and 99.9%. When comparing percentages on the basis of the age of competing dams (Table 1), there were not significant differences excepting the most extreme scenario. If a young (1 or 2 years old) and an old ewe (more than four years old) were involved in the analysis, the average probability of right assignment rose up to 61.74%, highlighting the relevance of this systematic source of variation for maternity identification. On the other hand, percentages did not differ when lambs share the same sire (56.96%) or where they were offspring from two different sires (57.75%), although it could be also due to the fact than only three sires contributed during this lambing season.

Table 1: Probability of correct assignment depending on the age of competing ewes

Age of dam	N	Mean (s.e.)
1-2 vs 1-2	89	52.45 ^{b} (1.82)
1-2 vs 3-4	193	56.85 ^{b} (1.24)
3-4 vs 3-4	81	55.88 ^{b} (1.91)
1-2 vs >4	282	61.74 ^{a} (1.02)
3-4 vs >4	192	56.52 ^{b} (1.24)
>4 vs >4	163	55.40 ^{b} (1.35)

Table 2: Probability of correct assignment depending on the sire of the two competing lambs

Sire	N	Mean (s.e.)
Same sire	364	56.96^a (0.91)
Different sire	636	57.75^a (0.69)

Although a low to moderate heritability limits the amount of information available in the phenotypic data to discriminate between candidates sires (Sapp, 2007) our result showed a better probability to infer true paternities (0.575), including information of birth weight (low heritability) and age of dam, than those reported by Cardoso and Tempelman (2003) and Sapp (2007). Cardoso and Tempelman (2003) reported posterior probabilities for paternity identification within the context of multiple-sire matings ranging from 0.28 to 0.52 depending on the number of competing sires and for a highly heritable trait ($h^2 = 0.5$). Also Sapp et al. (2007) showed smaller probabilities when only two candidate sires were competing for paternity, fluctuating from 0.51 to 0.54 on the basis of the assumed heritability (from 0.3 to 0.7).

Current development must not be viewed as a statistical tool for massive recovery of unknown maternities. Indeed, the percentage of unknown maternities in current extensive sheep industry is moderate to low. Nevertheless, this approach provides a statistical methodology for inferring unknown maternities at a very low-cost, it being useful when the lamb could become a replacement ewe or ram depending or their true parents. Although any result must be corroborated by genotyping the lamb and their potential parents, the preliminary screening for unknown maternities must discard several candidates to selection, reducing unnecessary genotyping and the consequent economic waste.

Conclusion

Our analyses showed that uncertain maternities could be appropriately modeled under Bayesian approaches by taking advantage of systematic (e.g. age of the dam) and genetic (e.g. breeding value) sources of information. This method prevents information lost and enhances the accuracy of ascertain maternities in livestock populations.

References

Banos, G., Wiggans, G. R., and Powell, R. L. (2001). *J. Dairy Sci.*, 84:2523-2529.

Casellas, J., Piedrafita, J., Caja, G., *et al.* (2009). *J. Anim. Sci.*, 87:72-79.

Cardoso, F. and Tempelman, R. (2003). *Genet. Sel. Evol.*, 35:469-487.

Foulley, J.L., GIanola, D., Planchenault, D. (1987). *Genet. Sel. Evol.*, 19:83-102.

Henderson, C.R. (1976). *Biometrics*, 32:69-83.

Im, S. (1992). *Appl. Statist.*, 41:109-116.

Israel, C., and J. I. Weller. (2000). *J. Dairy Sci.*, 83:181–187.

Quaas, R.L. (1976). *Biometrics*, 32:949-953.

Sapp, R.L. *et al.* (2007). *J. Anim. Sci*., 85:2391-2400.

Van Vleck, L.D. (1970a). *J. Dairy Sci.*, 53:1469-1475.

Van Vleck, L.D. (1970b). *J. Dairy Sci.*, 53:1697-1702.