Genetic Evaluation Combining Purebred And Crossbred Data In A Pig Breeding Scheme

N. Ibáñez-Escriche[*], J. Reixach[†], N. Lleonart[†], and J.L. Noguera∗

Introduction

The breeding system of domestic pigs is pyramidal. Nucleus herds supply genetically improved breeding stock to multiplier herds which disseminate genetic gain by supplying hybrid stock to commercial herds. As Dekkers (2007) point out, one of the main limitations of this system is that the pure breeds performance (PB) in nucleus farms can be a poor predictor of future cross breeds performance (CC) in commercial farms, i.e. genetic correlations from 0.4 to 0.7 (Lutaaya et al., 2001; Merks and de Vries, 2002). An alternative to improve this system is combine crossbred and purebred selection (CCPS), in which phenotypic data collected on CC relatives are used for selection of PB (Wei and van der Steen, 1991; Lo et al., 1993). CCPS can increase response to selection (Bijma and van Arendonk, 1998) but requires an efficient system for routine recording of performance and pedigrees at the CC level. For this reason, CCPS has been mostly studied using computer simulations and few works have used real data. The objectives of this study are: 1) perform CCPS for lean meat percentage in a real pig population, and 2) compare the impact of CCPS against a breeding program based on PB in the selection criterion.

Material and methods

Data. Data for this study were obtained from the Selección Batallé S.A. company and spanned a period of 4 yr (2006 to 2009). Data corresponding to three nucleus of purebred populations, Landrace (LD), Duroc (DU) and Pietrain (PI), one multiplying farm with animals from the two way cross F_1 (DU x LD) and commercial farms with animals from the three way cross F_3 (F_1 x PI). The trait analyzed was the lean meat percentage measured in slaughterhouse by AUTOFOM. This system allows obtaining individual records automatically (see Busk et al., 1999). Table 1 shows the distribution of records, means and standard deviations (Sd) of lean mean percentage for the PB and CC used in the analyses.

Table 1. Distribution of records, means and standard deviation of lean meat percentage for each purebred and crossbred population.

	Duroc	Landrace	Pietrain	F_1	F_3
Number	1,178	683	722	766	4,661
Percentage	14.71	8.59	9.01	9.57	58.22
Mean	48.39	55.20	63.42	51.82	59.46
Sd	5.07	3.47	2.48	4.05	2.31

[*] IRTA-Lleida. Genètica i Millora Animal. Av. Rovira Roure, 191. 25198 Lleida. Spain

[†] Selección Batallé S.A., Riudarenes, Spain

Statistical models. One purebred model and two crossbred models were used in the genetic evaluation. The purebred model (PB) was, $\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e}$.

The first crossbred model (CCPS1) considered records from PB (DU, LD, PI) and CC lines (F$_1$, F$_3$) as five correlated traits. This model can be written as an extension of the previous model. Let, $\mathbf{y}' = [\mathbf{y}'_{DU}, \mathbf{y}'_{LD}, \mathbf{y}'_{PI}, \mathbf{y}'_{F1}, \mathbf{y}'_{F3}]$ is now

$$\mathbf{y_i} = \mathbf{X_i b_i} + \mathbf{Z_i a_i} + \mathbf{e_i},$$

where i subscript denotes the vectors/matrices for appropriate breed composition (i=DU, LD, PI, F$_1$, F$_3$).

The second crossbred model (CCPS2) follows the model for crossbreeding schemes described by García-Cortés and Toro (2006), which is equivalent to the additive model developed by Lo et al. (1993). In this model, the covariance matrix of the additive values (G) is split into several parts allowing a simple analysis of the variance. In our case the model becomes

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{DU}\mathbf{a}_{DU} + \mathbf{Z}_{LD}\mathbf{a}_{LD} + \mathbf{Z}_{PI}\mathbf{a}_{PI} + \mathbf{Z}_{DU,LD}\,\mathbf{a}_{DU,LD} + \mathbf{e},$$

where $\mathbf{a}_{DU}$, $\mathbf{a}_{LD}$, $\mathbf{a}_{PI}$ are the vectors of the breeding values split by origin components and $\mathbf{a}_{DU,LD}$ correspond to the segregation term (see García-Cortés and Toro, 2006 for more details).

Above, $\mathbf{y}$ is the vector of observations; $\mathbf{b}$, $\mathbf{a}$ and $\mathbf{e}$ are the vectors of systematic, genetic and residual effects, respectively; $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices. In the three models, the vector of systematic effects included contemporary group, sex (male or female) and weight at end of a test period nested within sex.

Statistical analysis. Bayesian inference via Markov chain Monte Carlo (McMC) methods was used to analyze the data. Improper priors were assumed for the dispersion parameters and systematic effects, whereas genetic and residual effects were assumed to be distributed as N($\mathbf{0}$, $\mathbf{G_i}$) and N($\mathbf{0}$, $\mathbf{R_i}$), where $\mathbf{G_i}$, and $\mathbf{R_i}$ are, respectively, the additive and residual covariance matrices and i subscript denotes the model i=(PB, CCPS1, CCPS2). $\mathbf{G}_{PB} = \mathbf{A}\sigma^2_a$, $\mathbf{R}_{PB} = \mathbf{I} \otimes \sigma^2_e$, $\mathbf{G}_{CCPS1} = \mathbf{A} \otimes \mathbf{G_0}$ ($\mathbf{G_0}$=5x5 (co)variance matrix), $\mathbf{R}_{CCPS1} = \mathbf{I} \otimes \mathbf{R_0}$, $\mathbf{R_0}$=diag (σ^2_{eDU}, σ^2_{eLD}, σ^2_{ePI}, σ^2_{eF1}, σ^2_{eF3}), $\mathbf{G}_{CCPS2} = \mathbf{I} \otimes \mathbf{G_0}$ ($\mathbf{G_0} = \mathbf{A}_{DU}\sigma^2_{aDU} + \mathbf{A}_{LD}\sigma^2_{aLD} + \mathbf{A}_{PI}\sigma^2_{aPI} + \mathbf{A}_{DU-LD}\sigma^2_{aDU-LD}$) and $\mathbf{R}_{CCPS2} = \mathbf{I} \otimes \sigma^2_e$, where $\mathbf{A}$ is the additive genetic relationship matrix, $\otimes$ is the Kronecker product, σ^2_a is the additive genetic variance, σ^2_e is the residual variance and subscripts DU, LD, PI, F$_1$, F$_3$, correspond to the breed population and DU-LD to the segregation terms. Fully conditional distributions for location effects were multivariate normal, and inverted Wishart and Chi-squared distributions for dispersion parameters.

The comparison between PB and CCPS was based on reliabilities and rank correlations of additive values within PB individuals. Reliabilities were calculated as $r^2_{ij} = 1 - pev_{ij}/\sigma^2_{aj}$, where r^2_{ij} is reliability for animal i and breed j, pev_{ij} is the corresponding prediction error variance, and σ^2_{aj} is the additive variance for breed j (Lutaaya et al., 2002).

Results and discussion

Features of the marginal posterior distributions of additive genetic variances and heritabilities are presented in table 2 and 3, respectively. Posterior means of the additive variances and heritabilities within breed were similar across models, except for the

posterior additive variance of Landrace obtained with the CCPS2 model. Posterior means of the additive variances and heritabilities showed differences between breeds. Duroc and Landrace purebreds had similar and greater additive variances and heritabilities than Pietrain. The additive variance and heritability of the segregation between Duroc and Landrace were low and close to cero. Differences between additive variances for the CCPS models may be due to the fact that the CCPS2 model assumes a common residual variance for all breed populations and this can produce biased variance estimates. An extended model with different residual variances between breeds was developed but the McMC did not converge.

Table 2. Posterior means (standard deviations) of the additive and segregation variances for each breed and crossbred population

Model	Duroc	Landrace	Pietrain	F_1	F_3	S_{DU-LD}
PB	6.24 (1.49)	7.00 (1.54)	0.95 (0.40)			
CCPS1	7.76 (0.88)	6.05 (0.51)	1.10 (0.20)	5.92 (0.90)	1.83 (0.16)	
CCPS2	6.42 (0.50)	3.30 (0.36)	0.81 (0.14)			0.07 (0.02)

S_{DU-LD}: segregation variance of Duroc and Landrace purebreds.

Table 3 . Posterior means (standard deviations) of the heritability for each breed and crossbred population

Model	Duroc	Landrace	Pietrain	F_1	F_3	S'_{DU-LD}
PB	0.49 (0.10)	0.67 (0.11)	0.17 (0.07)			
CCPS1	0.57 (0.05)	0.63 (0.04)	0.20 (0.03)	0.52 (0.07)	0.37 (0.03)	
CCPS2	0.64 (0.03)	0.48 (0.03)	0.19 (0.03)			0.02 (0.01)

S'_{DU-LD}: ratio between the segregation variance and the residual variance.

The mean reliability of predicted purebred breeding values obtained for each model is presented in Table 4. The two CCPS models showed greater mean reliabilities than the PB model for all breeds. The improvement in reliability was around 75%, 50% and 140%, for Duroc, Landrace and Pietrain, respectively, when breeding values were estimated using the CCPS1 model, and 57%, 0.5% and 43 %, when the CCPS2 model was used. These results indicate that using crossbred data in the genetic evaluation increases considerably the reliability of the estimated purebred breeding values. In terms of reliability, the CCPS1 model was superior to the CCPS2 model, particularly regarding the Landrace breed. However, reliability depends on the additive variance and CCPS1 provided greater additive variances than CCPS2 for the three purebreds (table 3). It would be interesting to apply more criteria of model comparison to evaluate properly the goodness of fit and predictive ability of CCPS models.

Table 4. Mean reliability of predicted purebred breeding values for lean meat percentage.

Models	Duroc	Landrace	Pietrain
PB	0.21	0.41	0.16
CCPS1	0.37	0.62	0.39
CCPS2	0.33	0.43	0.23

The rank correlations of predicted breeding values are shown in table 5. The rank correlations were different across purebreds. The highest rank correlation between models was obtained by Landrace followed of Duroc, both purebreds with the greatest heritabilities. The rank correlation between PB and CCPS models for Pietrain was smaller than 0.80, suggesting an important re-ranking when information from crossbreds animals is taken into account. These results reveals that selection decisions based only on purebred information could be inappropriate and thus to slow genetic progress. To test whether this re-ranking would affect selection decisions on Pietrain purebred, we compared the top 15 boars out of 45 candidates to selection using PB or CCPS1. Of the top 15 boars selected using PB or CCPS1, there were disagreement in 5 boars, which represented 33% of the total boars selected. This outcome would show that selection decisions were depending on whether the crossbred and purebred information were combined or not in the genetic evaluation.

Table 5. Rank correlations of predicted breeding values obtained for the models.

Breed	N	Models		
		PB,CCPS1	PB,CCPS2	CCPS1,CCPS2
Duroc	5,055	0.91	0.87	0.94
Landrace	1,299	0.97	0.96	0.94
Pietrain	1,780	0.79	0.79	0.85

Conclusion

An increase of reliability was obtained when crossbred and purebred information was combined to perform the genetic evaluation. Moreover an important re-ranking of animals with a corresponding change in the selection was also shown when genetic evaluation was based on PB or CCPS models. These results support the idea of using crossbred models to evaluate lean meat percentage in this pig breeding scheme. However, more research is needed to develop and properly compare crossbred models.

References

Bijma, P., and van Arendonk, J.A.M. (1998). *J. Anim. Sci*. 66: 529–542.

Busk, H., Olsen, E.V., and Brødum J. (1999). *Meat Sci.* 52: 307-314.

Dekkers, J.C.M. (2007). *J. Anim. Sci.* 85: 2104-2114.

Lo, L.L., Fernando, R.L., and Grossman, M. (1993). *Theor. Appl. Genet*., 87: 423-430 .

Lutaaya, E., Misztal, I., Mabry, J.W. *et al.* (2001). *J. Anim. Sci.* 79: 3002-3007.

Lutaaya, E., Misztal, I., Mabry, J.W. *et al.* (2002). *J. Anim. Sci.* 80: 2263-2266.

Merks, J., and de Vries, A.W. (2002). In *Proc 7th WCGALP,* Electronic Comm. No. 03-01.

Wei, M., and Van der Steen, H.A.H. (1991). Anim. Breed. Abstr. 59:281.