A compositional genetic analysis of oleic acid content in pig meat

J. ESTANY[1], R. ROS[1], M. TOR[1] and J. REIXACH[2]
[1] Department of Animal Production, University of Lleida, Spain, jestany@prodan.udl.cat
[2]Selección Batallé, Spain

1. Introduction

Intramuscular fat (IMF) content and composition, particularly the oleic fatty acid content (OL), are major quality characteristics of pork fresh and dry-cured products. They are known to be related to nutritional, manufacturing and organoleptic properties, as well as to human health. It is known that IMF content is under genetic control but little evidence is available for IMF composition, namely OL. There are very few estimates in the literature regarding genetic parameters for OL (Suzuki et al., 2006) and, besides, most of them are based on small data sets from experiments designed for other purposes (Ntawubizi et al., 2010; Sellier et al., 2010). However, genetic parameters associated to IMF and OL (i.e. heritability and genetic correlations with other relevant traits) are needed for developing selection criteria and optimum breeding strategies and programmes.

IMF content is usually expressed in percent of dry or wet matter and OL in percent of total fatty acids in IMF. However, all research done in this field was not aware of the compositional nature of these data (Aitchison, 1986). The purpose of the present contribution is to compare results from standard linear with compositional data analyses for IMF and OL. Analyses were compared in terms of genetic parameter estimates, selection efficiency, and predictive capacity.

2. Material and methods

2.1 Animals and experimental data

The data set used for the estimation of the genetic parameters and breeding values consisted of 93,920 pedigreed Duroc pigs, from which 85,253 had at least one recorded trait. Pigs with records were born from 1996 to 2010 and recorded traits were body weight (BW) and backfat thickness (BT) at 180 days, and IMF and OL at 215 days. A description of the records used for each trait is given in Table 1. BT was ultrasonically measured at 5 cm off the midline at the position of the last rib (Piglog 105®, Herlev, Denmark). IMF was determined as the sum of eleven fatty acids (Ros et al., 2011) expressed as triglyceride equivalents (Bosch et al., 2009) and then expressed as percent of wet matter. Individual fatty acids and OL in particular were determined in duplicate on a sample of *gluteus medius* muscle by gas chromatography. Then, OL was expressed as percent of total fatty acids.

Trait	n	Mean	SD	Minimum	Maximum
Body weight,kg	85,002	104.8	12.5	62.0	167.0
Backfat depth, mm	80,687	15.6	3.5	6.5	36.0
IMF, %	943	4.9	1.9	1.5	13.3
OL,% fatty acid	947	44.8	3.1	36.3	55.5

Table 1. Data description of the data set used in the analyses.

Egozcue, J.J., Tolosana-Delgado, R. and Ortego, M.I. (eds.)
ISBN: 978-84-87867-76-7

2.2 Statistical analyses

2.2.1. Estimation of genetic parameters and genetic values

A multiple four-trait animal mixed model was used to analyze the data in Table 1. The model for BW and BT included batch (1,039 classes), sex (male, female, and castrated), litter (32,426 litters), and animal genetic (breeding) value (93,920 animals), while the model for IMF and OL only included batch (13 classes) and animal. Age at measurement was included as a covariate for all traits. IMF and OL were analyzed using either the raw or the transformed scores. As a first approach to convert IMF and OL compositional data to samples in real space, the following isometric logratio (ilr) transformed variables were used (Egozcue et al., 2003):

$$u_1 = \frac{1}{\sqrt{6}} \ln \left(\frac{OL\ x\ NOL}{(1-IMF)^2} \right) \qquad u_2 = \frac{1}{\sqrt{2}} \ln \left(\frac{OL}{NOL} \right)$$

where NOL=IMF-OL is the percent of fatty acids in IMF other than OL, and (1-IMF) is the percent of wet matter in pork other than IMF, such that (1-IMF) + OL + NOL = 100.

Genetic parameters and animal genetic effects were estimated in a Bayesian framework using Gibbs sampling (Legarra et al., 2008). Flat priors were used for variance components and systematic effects. For animal effects, prior distribution was a multivariate normal distribution with mean zero and variance $\mathbf{G} \otimes \mathbf{A}$, where **A** is the numerator relationship matrix and **G** is a 4×4 genetic relationship matrix between traits. Prior distributions for litter and residual effects were multivariate normal with mean zero and variances $\mathbf{C} \otimes \mathbf{I}$ and $\mathbf{R} \otimes \mathbf{I}$, where **C** is a 2×2 (co)variance matrix between litter effects of BW and BT and **R** is a 4×4 (co)variance matrix between residuals. Statistical inferences were derived from the samples of the marginal posterior distribution using a unique chain of 500,000 iterations, where the first 100,000 were discarded and one sample out of 100 iterations retained. The heritability of each trait was calculated as the ratio between animal to total variance. The genetic correlation between two traits was calculated as the correlation between breeding values for such traits. The correlation between raw (IMF, OL) and ilr-transformed (u_1, u_2) estimates of the breeding values was calculated as a measure of selection efficiency of the estimates inferred using raw scores.

2.2.2. Predictive capacity

A cross-validation approach using real and simulated data based on a full-sib design was used to assess the predictive capacity for future records of each analysis. Thus, (case 1) from data in Table 1, 350 paired full-sibs with records on IMF and OL were taken with the purpose of predicting the record of one sib (predicted set) by regressing on the record of the other sib (observed set). Robustness of the prediction was assessed using as regression coefficients values taken at equal distance from the interval 0.1 to 1. The predictive capacity of each case (ilr-transformed or not by the regression coefficient value) was evaluated in terms of the mean square error of the prediction, expressed as a proportion of the total variance (MSE). This process was repeated 5,000,000 times placing sibs at random either in the observed or in the predicted set. Average and standard error of MSE across repetitions were calculated. On the other hand, (case 2) simulated compositional data mimicking observed data were obtained by repeated sampling of (u_1, u_2, $u^{FS}{}_1$, $u^{FS}{}_2$) from a multivariate distribution with mean (-2.8, -0.1, -2.8, -0.1) and (co)variance V, where the superscript FS refers to full sib. Matrix V was derived according to standard quantitative genetics theory using the estimates obtained in 2.2.1. A full-sib design consisting of 1,000,000 was repeated 10 times. Then, predictive capacity was assessed as with real data.

3. Results

3.1 Genetic parameters and breeding values

The genetic parameter estimates associated to IMF and OL hardly differed between the standard and the compositional analysis (Table 2). The maximum absolute divergence occurred for the genetic correlation between OL and BT, where the posterior mean (standard deviation) was 0.23 (0.11), for the standard analysis, and 0.19 (0.11), for the ilr-compositional analysis. The genetic correlation between IMF and OL (0.48; SD: 0.11) was similar to that between the transformed variables (0.50; SD: 0.12).

	h^2		r_G with BW		r_G with BT	
Trait	Mean (SD)	$HPD_{95}{}^{\alpha}$	Media (DT)	$HPD_{95}{}^{\alpha}$	Media (DT)	$HPD_{95}{}^{\alpha}$
IMF	0.57 (0.09)	0.43;0.75	0.29 (0.11)	0.08;0.47	0.38 (0.10)	0.18;0.56
OL	0.51 (0.08)	0.36;0.68	0.12 (0.11)	-0.11;0.34	0.23 (0.11)	0.01;0.43
u_1	0.57 (0.10)	0.38;0.78	0.27 (0.10)	0.08;0.46	0.36 (0.10)	0.16;0.54
u_2	0.51 (0.08)	0.36;0.68	0.13 (0.11)	-0.08;0.37	0.19 (0.11)	-0.02;0.37

Table 2. Heritability (h^2) and genetic correlation (r_G) of IMF and OL, and u_1 and u_2, with BW and BT .$^{\alpha}$ HPD95: Highest Probability Density Interval at 95%.

The residual parameter estimates are given in Table 3. The maximum absolute divergence occurred for the residual correlation between IMF and BT, where the posterior mean (standard deviation) was 0.14 (0.08), for the standard analysis, and 0.20 (0.08), for the ilr-compositional analysis. The residual correlation between IMF and OL (0.17; SD: 0.11) was lower than that encountered between the transformed variables (0.30; SD: 0.13).

	h_E^2		r_E with BW		r_E with BT	
Trait	Mean (SD)	$HPD_{95}{}^{\alpha}$	Media (DT)	$HPD_{95}{}^{\alpha}$	Media (DT)	$HPD_{95}{}^{\alpha}$
IMF	0.43(0.09)	0.25;0.57	0.06 (0.08)	-0.08;0.21	0.14 (0.08)	-0.02;0.31
OL	0.49 (0.08)	0.32;0.64	0.19 (0.08)	0.04;0.35	0.21 (0.08)	0.06;0.37
u_1	0.43 (0.10)	0.22;0.62	0.08 (0.07)	-0.06;0.23	0.20 (0.08)	0.06;0.38
u_2	0.49 (0.08)	0.32;0.64	0.19 (0.07)	0.05;0.32	0.24 (0.08)	0.09;0.40

Table 3. Residual variance relative to total variance (h_E^2) and residual correlation (r_E) of IMF and OL, and u_1 and u_2, with BW and BT .$^{\alpha}$ HPD95: Highest Probability Density Interval at 95%.

The Spearman correlation between the estimated breeding values from both analysis was high (0.97, for IMF and u_1, and 0.95, for OL and u_2), particularly when only pigs with data on IMF and OL where considered (r>0.99 for both correlations). Posterior standard deviations associated to breeding values were proportionally higher in the standard as compared to ilr-transformed analysis, particularly for IMF. If all pigs are included, the coefficient of variation associated to the estimates of the breeding values for IMF was almost doubled. As a result, the correlation of the standard deviations associated to raw breeding values and transformed breeding values were consistently lower than that observed between the corresponding posterior means of the breeding values (0.67, for IMF and u_1, and 0.59, for OL and u_2).

3.2 Predictive capacity

The ilr-estimate associated to IMF (u_1) showed better predictive ability, in terms of MSE, than the raw score of IMF (Figure 1), with differences around 2%. No relevant differences for MSE between OL and u_2 were observed, but, in comparison with IMF, values for OL were much lower (0.56 and 0.82 at minimum MSE for OL and IMF, respectively). Estimates based on transformed variables, particularly for u_1, were slightly more robust across regression values as well. These results were confirmed using simulated data (Figure 2). Interestingly, u_1 also displayed lower variance than IMF, both in the real (0.04 vs. 0.06, at minimum MSE) and in the simulated (0.18 vs. 0.33, at minimum MSE) case, and more robust across repetitions. Transformed variables u_1 and u_2 resulted to be less sensible to occasional outliers.

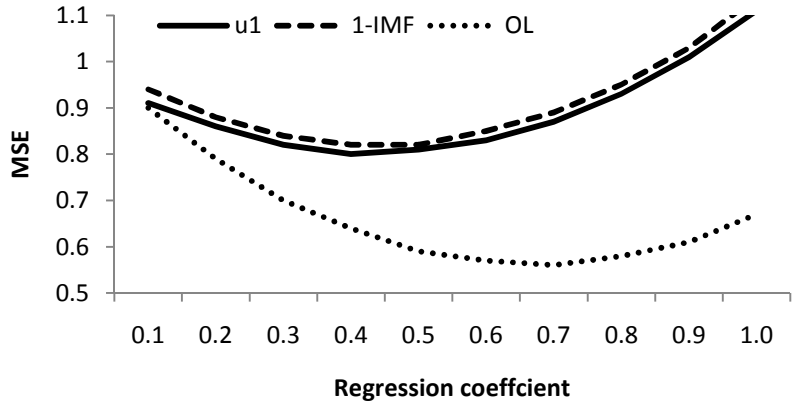

Figure 1. Standarized mean square error of prediction by the regression coefficent used in predicting the performance of an individual for u1, (1-IMF), and OL with a full sib's record. Results for u_2 are the same as in OL. Case 1: Real data.

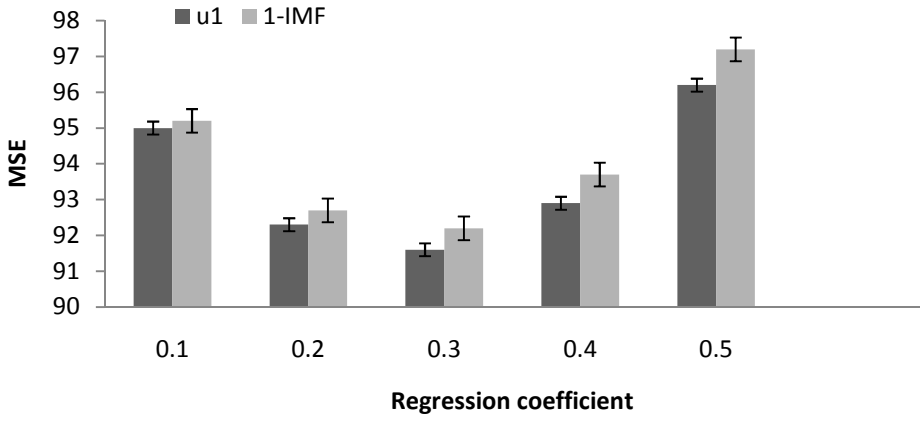

Figure 2. Standarized mean square error of prediction by the regression coefficent used in predicting the performance of an individual for $u_1$and (1-IMF) with a full sibs' record. Bar graphs indicate standard deviation. Case 2: Simulated data.

In conclusion, little differences for genetic parameter and selection efficiency were observed when inferences were based on raw or ilr-transformed data on IMF and OL. This may be attributed to the relatively low variability of IMF and OL (Table 1; Ros et al., 2011). However, even in this simple case, although mostly for IMF rather than for OL, predictions of future records based on ilr-transformed variables performed better than those on raw data. Because biological and economical interpretation of transformed results is not straightforward, specifically for the breeding values,

compositional data analysis can be refrained from being used in practice unless expected returns are clearly shown. Further analyses need to be undertaken using the whole fatty acid compositional profile and, in particular, to assess specifically the behavior of other relevant fatty acids displaying higher compositional variation than OL.

4. Implications

Log-ratio transformation is expected to affect more intramuscular fat content than oleic content in pig meat, and their predicted rather than their realized selection responses.

References

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London (UK). (Reprinted in 2003 with additional material by The Blackburn Press). 416 p.

Bosch, L., M. Tor, J. Reixach, and J. Estany (2009). Estimating intramsucular fat content and fatty acid composition in live and psot-mortem samples in pigs *Meat Science 82*:432-437

Egozcue, J.J., V. Pawlowsky-Glahn, G. Mateu-Figueras, and C. Barceló-Vidal (2003). Isometric logratio transformations for compositional data analysis *Mathematical Geology 35*:279-300

Legarra, A., L. Varona, and E. López-Maturana (2008) (http://cat.toulouse.inra.fr/~alegarra/).

Ntawubizi, M., E. Colman, S. Janssens, K. Raes, N. Buys, and S. De Smet (2010). Genetic parameters for intramuscular fatty acid composition and metabolism in pigs**.** *Journal of Animal Science 88*:1286‑1294

Ros, R., J. Reixach, M. Tor, and J. Estany (2011). Exploratory data analysis for fatty acid composition in pig meat. 4th International Workshop on Compositional Data Analysis, Sant Feliu de Guíxols, Girona, Spain. Ref: 29.

Sellier, P., L. Maignel, and J.P. Bidanel (2010). Genetic parameters for tissue and fatty acid composition of backfat, perirenal fat and longissimus muscle in Large White and Landrace pigs. *Animal 4*: 497–504.

Suzuki, K., M. Ishida, H. Kadowaki, T. Shibata, H. Uchida, and A. Nishida (2006). Genetic correlations among fatty acid compositions in different sites of fat tissues, meat production, and meat quality traits in Duroc pigs. *Journal of Animal Science 84*:2026-2034.

Acknowledgements. Work funded by Ministry of Science and Innovation, Spain (grant AGL2009-09779). R. Ros acknowledges a FPI grant from MICINN (BES-2010-034607).