Epistatic genetic variances for litter size in pigs using genomic models

Z.G. Vitezica¹, H. Crichan¹, W. Herring², A. Reverter³ & A. Legarra¹

¹INRA/INPT, UMR 1388 GenPhySE, F-31326 Castanet-Tolosan, France
zulma.vitezica@ensat.fr (Corresponding Author)
²PIC North America, Hendersonville, USA
³CSIRO Agriculture & Food, 306 carmody Rd., St Lucia, Queensland 4067, Australia

Introduction

Additive, dominance and epistatic genetic variances of quantitative traits are required for estimating breeding values and taking selection decisions. Most of the genetic variance is additive (Hill et al., 2008; Mäki-Tanila & Hill, 2014). However, even if functional epistatic action of the genes is mostly included in the additive variance, epistatic variance should not be neglected. Knowing its magnitude in real data and exploring the predictive ability of a model that accounts for epistatic effect are necessary steps. Equally important is to know how much GxG interaction exists (e.g. the effect of the same allele in different breeds) or to account for the contribution of epistasis to the build-up of “new” additive variance.

Recently, a flexible and general approach to construct “genomic” relationship matrices for populations in Hardy-Weinberg equilibrium (HWE) or not, such as F1 crosses, was proposed (Vitezica et al., 2017). The authors proved that epistatic genomic relationship matrices for two or higher order interactions can be constructed using Hadamard products of additive and dominance genomic orthogonal relationships, regardless of the existence of HWE. However, a standardization based on the trace of the relationship matrices is needed.

The objectives of this work are to estimate additive and non-additive (dominance and epistasis) variance components in a real pig population, and to investigate if the predictive ability of genomic evaluation models increases with the inclusion of dominance and second order epistasis effects.

Material and methods

Data for this study were provided by Genus plc (Hendersonville, TN, USA). The studied trait was litter size (total number of piglets born per litter) with an average value (± SD) of 12.68 ± 3.07. A total of 13,369 records and 38,779 SNPs (88% in HWE, P-value<0.001) were available for 3,619 sows with an average of 3.7 litters per sow.

Genomic evaluation models

Phenotypes collected for the genetic nucleus (pure line) were analyzed using a GBLUP (mixed) model. Parity order, and farm, year and month of farrowing were included as fixed effects. The model also included a permanent environmental effect.

The linear model including additive, dominant and second order interaction terms can be written as:
where is the vector of phenotypic records, is the fixed effect vector, models the inbreeding depression, where is genomic inbreeding coefficient and is the inbreeding depression parameter per unit of inbreeding, is a vector of breeding value of the sows, is a vector of dominance deviations, is a vector of the second order epistatic genetic values, and are design matrices relating records to fixed effects and genetic values respectively, is the permanent environmental effect vector, and is a residual vector. The epistatic genetic effects can be partitioned into additive-by-additive (, additive-by-dominance (, and dominance-by-dominance (. The covariance matrices of genetic effects were and where the covariance matrices and (involved in the construction of , and ) were constructed assuming HWE. The additive relationship matrix was calculated as in VanRaden (2008): , where the matrix has elements that are equal to for genotypes , , and is the frequency of . The dominance relationship matrix was computed as where has elements equal to for genotypes , , and , respectively (Vitezica et al., 2013). Note that under the assumption of HWE, the GDBLUP “breeding” model in Vitezica et al. (2013) is a particular case of the NOIA (Natural Orthogonal Interactions) model in Vitezica et al. (2017). The epistatic covariance matrices were computed using the Hadamard products and traces as , , and (Vitezica et al., 2017).

To take into account directional dominance, the genomic inbreeding coefficient was included in the evaluation models (Xiang et al., 2016, Aliloo et al., 2017) defined as the proportion of genotyped SNPs at which an individual is homozygous (Silio et al., 2013).

Variance components were estimated for five models that included only additive effects (A), additive and dominance effects (A+D), and second-order epistatic effects: additive-by-additive (A+D+AA), additive-by-additive plus additive-by-dominance (A+D+AA+AD) and additive-by-additive plus additive-by-dominance plus dominance-by-dominance effects (A+D+AA+AD+DD).

Genetic variances (,, and ) were estimated by Bayesian methods using Gibbs Sampling using the software gibbs2f90 (Misztal et al., 2002). A total of 200,000 iterations were run, discarding the first 10,000 and keeping every 100th sample. Convergence was checked by visual inspection of the chains.

**Predictive ability**

Estimated breeding values (EBVs) were obtained with their respective variance components estimated with Gibbs Sampling. EBVs were computed for all the animals with the “whole” data set, included sows farrowed from 2000 to 2014. In addition, another data set included sows farrowed until 2010 was defined and it was called the “partial” data set. EBVs were also obtained for all the animals with the “partial” data set. The additive model (A) was used to estimates, while were obtained based on three models: A, A+D, and A+D+AA+AD+DD.

The models were compared for the selection candidates (1,179 sows farrowed after 2010) using the approach proposed by Legarra & Reverter (2017). We compute the three statistics of cross-validation: refers to as the prediction bias and should be equal to 0; is the slope of the regression of selection candidates’ EBVs obtained with whole data on EBVs estimated with partial data, measures bias in variance components and should be 1; and is the correlation of partial on whole and measures the accuracy.

**Results and discussion**
Estimates of additive and dominance genetic variances for litter size ranged from $0.81 \pm 0.12$ to $0.83 \pm 0.12$, and from $0.17 \pm 0.11$ to $0.20 \pm 0.11$, respectively for all the models (A, A+D, A+D+AA, A+D+AA+AD and A+D+AA+AD+DD). Variance component estimates do not differ among models (Figure 1), which is expected from the empirical orthogonality in the partition of the total genetic variance, a property that holds under HWE (which holds in this data set) and assuming linkage equilibrium (which does not hold). Under orthogonality, the substitution effect contributes to the additive variance, the dominance deviation contributes to the dominance variance, etc. and there is no covariance between the genetic effects.

Epistatic variance estimates were from $0.12 \pm 0.10$ to $0.14 \pm 0.12$ for the additive-by-additive component, $0.11 \pm 0.09$ for the additive-by-dominant component and $0.11 \pm 0.09$ for the dominant-by-dominant component. Epistatic variance estimation had a large standard error in all the models (A+D+AA, A+D+AA+AD, A+D+AA+AD+DD). These results point out the difficulties in obtaining a good estimate of epistatic variances also from genomic information even when these are only second order interactions.

Narrow-sense heritability estimates for litter size across models were similar, close to $0.09$ and consistent with those reported by Varona et al. (2007) and Guo et al. (2015). Dominance variance as percentage of the total phenotypic variance, was about $0.02$ as in Misztal et al. (1998). The total epistatic variance (additive-by-additive plus additive-by-dominance plus dominance-by-dominance) expressed as percentage of the total phenotypic variance, was of $0.04$. These results need to be taken with caution due to the low precision in the epistatic estimates. The broad-sense heritability for litter size was therefore $0.15$, almost twice the narrow-sense heritability.

For the selection candidates, which means that genetic evaluation is unbiased and selection does not affect the comparisons across old and young animals. The statistic was equal to $0.91$, $0.95$ and $1.02$ in models A, A+D and A+D+AA+AD+DD respectively. The genetic evaluation is slightly inflated with the additive model and this inflation vanishes with the model included additive, dominance and epistatic effects. The inclusion of non-additive components in the evaluation model possibly takes better into account the selection process and the changes in variances. The accuracy was around $0.68$ for all models.

Conclusions
Estimates of additive genetic variance don’t differ when the model is expanded from additive to include dominance and epistatic effects. The broad-sense heritability for litter size was almost twice the narrow-sense heritability.

All models have the same ability to rank the animals. Using additive plus dominance plus epistatic effects rather than only additive effects did not improve the accuracy of prediction of breeding values but reduced the bias of genetic evaluation. The inclusion of non-additive components in the evaluation model possibly takes better into account the selection process and the changes in variances.

**List of References**


