On the value of genotyping terminal crossbred pigs for nucleus genomic selection for carcass traits.

A. Sewell¹, H. Li¹, C. Schwab¹, C. Maltecca² & F. Tiezzi²

¹The Maschhoffs LLC., Carlyle, IL, 62231, USA
alysta.sewell@pigsrus.com (Corresponding Author)
²North Carolina State University, Raleigh, NC, 27695, USA

Summary

In this study we investigated the advantages of building a swine crossbred reference population for genomic selection in the nucleus, when selection is performed with the terminal cross performance as a breeding goal. Carcass measures were collected in about 5,000 three-way-cross individuals and live measures of body weight and tissue deposition were collected on about 3,000 terminal Duroc individuals, paternal-half-sibs of the crossbred individuals. Around 400 of these Purebred individuals served as selection candidates in this study, and the goal was to predict their direct genomic value for the crossbred traits. Different models were built, where an increasing number of crossbred genotypes was included (0 to 1,252), while all the purebred candidates and their sires were considered genotyped. We also investigated the advantage of including crossbred genotypes in presence or absence of purebred phenotypic records. Data show that including crossbred genotypes is beneficial to the selection of breeding candidates, especially when purebred phenotypes are not available. However, even in presence of purebred phenotyping, predictions of carcass Average Daily Gain (CADG) and carcass Loin Depth (cLD) are basically null if crossbred genotypes are not included. Prediction of carcass Back Fat (cBF) still benefits from the inclusion of crossbred genotypes, although even with only pedigree information prediction accuracy was non-null. We conclude that crossbred individuals should be considered in building reference population for the improvement of crossbred carcass value.

Keywords: genomic selection, genotyping crossbred, prediction of crossbred performance

Introduction

Genetic improvement of swine populations has led to notable improvements of performance, with strong benefits to the industry. This progress has been in part based on the use of the 3-way terminal cross, where boars from terminal lines (e.g. Duroc) are often used to sire commercial crosses with maternal lines crosses (e.g. Duroc x (Landrace x Large White)). However, nucleus selection has often proceeded using information (phenotypes) collected in the nucleus itself, ignoring the non-unity genetic correlation between purebred and crossbred performance.

The ultimate goal of purebred selection is to improve crossbred performance and genomic selection provides an opportunity to improve prediction. The addition of crossbred genomic information has been shown to be feasible via extension of single-step BLUP method for genomic evaluation incorporating both purebred and crossbred performance (Christensen et al., 2014) as well as the combination of pedigree and marker based
relationships (Christensen et al., 2015). As the genetic correlation between purebred and crossbred performance has been demonstrated to be less than 1 (Hidalgo et al., 2015), there has been exploration of inclusion of non-additive effects to increase the prediction accuracy of crossbred performance (Esfandyari et al., 2016). The purpose of this study was to assess the value of inclusion of crossbred genotype data along with purebred phenotypic information the prediction accuracy of crossbred performance.

Material and methods

Commercial crossbred (CB) and purebred Duroc (PB) individuals were generated for 28 paternal half-sibs families (PHS) and raised at The Maschhoffs LLC (Carlyle, IL, USA) farms. Crossbred individuals were allocated into 334 single-sire single-gender pens, and kept from the first week after weaning to the time they were ready for slaughter. An entire pen of pigs was sent to market when the pen average weight reached 136 kg. Four individuals were genotyped from each pen to represent the within-pen variation. Crossbred carcass data collected included hot carcass weight and the Fat-O-Meator measures Back Fat (cbBF) and Loin Depth (cbLD) and carcass Average Daily Gain (cbADG) was derived. Purebred phenotypes included live Average Daily Gain (pbADG) from weaning to 180 days of age, and the ultrasound measures Back Fat (pbBF) and Loin Eye Area (pbLEA). Phenotypic data was available on 5,124 CB and 3,036 PB and of those 1,252 and 1,071 were genotyped (GeneSeek® Genomic Profiler Porcine), respectively. A phenotypic index was constructed for each genotyped CB individual applying equal emphasis on each of the CB carcass traits. Pen average index was calculated and deviation from pen average was computed for each genotyped pig. Three individuals per pen were assigned A, B, or C in order of increasing deviation from the pen average. Both CB and PB phenotypes were pre-adjusted for the non-genetic effects of contemporary group, gender, pen and litter.

Cross-validation was used to assess the model performance in predicting purebred breeding value for crossbred performance. Scenarios were simulated where PB candidates for selection from 7 out of 28 PHS were to be predicted, thus their phenotypes were masked. There were 404 PB male selection candidates with genotypes, and these individuals were used to test the predictive ability of the models. The direct genomic value for crossbred performance for each of the traits was analyzed and used for ranking the 404 PB selection candidates. This was obtained by multiplying the PB genotypes by the marker effects calculated for cbADG, cbBF and cbLD on the entire set of crossbred individuals to approximate the closest estimate to the average performance of the CB progeny.

Different statistical models were tested to assess the value of including: 1) an increasing number of CB genotypes, 2) phenotypic information for purebreds (i.e. nucleus phenotypes). Predictions were obtained using single-step Genomic BLUP (ssGBLUP), which included only the additive genetic effect of the pig and the residual. For each trait (ADG, BF and LD) different models were tested, all models included CB phenotypic information as well as sire and PB genomic information.

As for the inclusion of CB genotypes, we tested:
- PED: no CB genotypes included
- A: included genotypes for CB individuals A (n=334)
- AB: included genotypes CB individuals A and B (n=668)
- ABC: included genotypes for CB individuals A, B and C (n=1,002)
- ALL: included genotypes for all CB individuals (n=1,252)

While for the inclusion of PB phenotypes, we tested:
- Scenario 1: no PB phenotypes included
- Scenario 2: PB phenotypes included only for the 21 training PHS

In the nucleus scenario, we assumed that PB selection candidates had not yet completed performance test.

Relationship matrices A and H were created using the preGSf90 program (Aguilar et al., 2014) and breeding values were estimated using BLUP as implemented in the BLUPf90 family of programs (Misztal et al., 2002). Model predictive ability was computed as the correlation between the predicted and the breeding value of the PB selection candidates as produced by ssGBLUP. Four values were obtained for the four validation sets, then averaged.

**Results and discussion**

The list of traits with respective descriptive statistics is listed in Table 1. Model performance is shown in Figure 1. Accuracy of prediction of crossbred performance was in fact negative for ADG and LD in the PED model for both scenarios. For ADG and LD, with the inclusion of CB genotypes, accuracy increased up to 0.3 for Scenario 1 and 0.1 for Scenario 2. For BF, accuracy was comparatively larger even using model PED (0.25 both scenarios) and increased up to 0.6 in Scenario 1 and 0.35 in Scenario 2.

The inclusion of an increasing number of CB genotypes showed an increase in prediction accuracy both when nucleus phenotypes were excluded (Scenario 1) and included (Scenario 2). However, in the latter case, the advantage of using CB genotypes was smaller.

**Conclusions**

The inclusion of crossbred genotype data did increase the predictive ability of the crossbred carcass traits ADG, LD and BF. Each additional increase in the amount of crossbred genotype data resulted in an incremental increase in prediction accuracy whether purebred phenotypic data was included or not. While the inclusion of purebred phenotypic data in the nucleus resulted in higher accuracy when pedigree models were compared, the addition of crossbred genotype data had a larger impact on predictive ability in the absence of purebred phenotypic data. Further exploration is necessary with larger sample sizes and with other traits of importance such as mortality.
Table 1. Descriptive statistics of the phenotypic data used.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Metric</th>
<th>N</th>
<th>mean</th>
<th>stddev</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossbred FOM(^1) Loin Depth</td>
<td>cm</td>
<td>4,894</td>
<td>6.68</td>
<td>0.69</td>
<td>3.6</td>
<td>8.9</td>
</tr>
<tr>
<td>Crossbred FOM(^1) back-fat depth</td>
<td>cm</td>
<td>4,893</td>
<td>2.26</td>
<td>0.49</td>
<td>1</td>
<td>5.1</td>
</tr>
<tr>
<td>Crossbred carcass average daily gain</td>
<td>kg/d</td>
<td>5,124</td>
<td>0.55</td>
<td>0.07</td>
<td>0.31</td>
<td>0.96</td>
</tr>
<tr>
<td>Purebred ultrasound Loin Eye Area</td>
<td>cm(^2)</td>
<td>3,036</td>
<td>50.69</td>
<td>6.26</td>
<td>26.45</td>
<td>71.9</td>
</tr>
<tr>
<td>Purebred ultrasound back-fat depth</td>
<td>cm</td>
<td>3,036</td>
<td>1.62</td>
<td>0.40</td>
<td>0.56</td>
<td>3.43</td>
</tr>
<tr>
<td>Purebred live average daily gain</td>
<td>kg/d</td>
<td>2,967</td>
<td>0.89</td>
<td>0.12</td>
<td>0.41</td>
<td>1.29</td>
</tr>
</tbody>
</table>

\(^1\) = Fat-O-Meator

Figure 1. Cross-validation prediction accuracy of purebred genomic value for crossbred performance. On the left: scenario 1, phenotypic data in the nucleus is not available. On the right: scenario 2, phenotypic data in the nucleus is available only on the previous rounds of performance test (adg= carcass average daily gain, bf=back-fat depth at market weight, ld=loin depth at market weight).

List of References


