Reliability of Genomic Selection of Reproduction Traits in Finnish Yorkshire Pig Breed

P. Uimari[1]**, M-L Sevón-Aimonen**[2] **and T. Serenius**[1]**.**
[1]University of Helsinki, [2]Agrifood Research Finland, MTT, Jokioinen, [3]Figen Oy, Finland

ABSTRACT: Genomic selection is effective for traits that can be measured only from one sex. These traits include sow reproduction traits in pigs. The aim of this study was to estimate reliability of genomic selection in the Finnish Yorkshire population. The data included 723 genotyped AI-boars. Phenotypic data was divided into two sets: records before 2005 and records up to 2012. Boars having daughter records before 2005 were treated as reference animals and boars having daughters between 2005-2012 were treated as validation animals. Direct genomic values (DGV) were estimated using SNP-BLUP. The Interbull validation method was applied to calculate the validation reliability. The reliability varied from 0.32 to 0.58. It was the highest for pig mortality in later parities and the lowest for the total number of piglets born in the first parity. Corresponding validation accuracy of the parent average was 0.20.
Keywords: pig; reproduction traits; reliability; genomic selection

Introduction

Profitability of pork production relies mainly on production traits such as growth rate and reproduction traits such as number of piglets born alive. Reproduction traits have been part of the national selection program in Finland for several decades and breeding values of different reproduction traits have been estimated using BLUP since the beginning of 1990's. Currently the average number of weaned piglets / sow / year is around 24 in Finland; that is the same as in Sweden and Norway but lower than in Denmark and the Netherlands where the corresponding number is 29 (AgroSoft, (2012)). The average genetic progress in piglet production in Finnish Yorkshire has been 0.1 weaned piglets / year. Thus, the classical breeding program based on litter records, pedigree information, and BLUP has been successful for litter traits.

Selection using SNP-markers i.e. genomic selection is widely implemented in dairy cattle breeding. Genomic selection is also used by commercial pig breeding companies such as Topigs (www.topigs.com) and Danavl (www.danavl.com) among others. The most critical components of successful genomic selection are sufficient size of the reference population, homozygosity of the population and close relatedness between the reference and the commercial populations.

The national breeding program of the Finnish Landrace and Yorkshire is operated by Figen (www.figen.fi). Both breeds are purebred with relatively small effective population size (Uimari and Tapio, (2011)). Typical to small breeds the limiting factor for genomic selection is a relatively small number of AI-boars that have enough daughter records. Fortunately, a repository of hair samples of AI-boars from 1990's to current day has been available for genotyping purposes.

The objective of this study was to predict reliability of genomic selection for sow reproduction traits in the Finnish Yorkshire breed.

Materials and Methods

Animals. Data included 723 genotyped Finnish Yorkshire AI-boars. Genotyping was conducted using Illumina PorcineSNP60 BeadChip. Only animals having call rate higher than 90% were included in the data. SNP that had call rate less than 90% were also removed from the data. For the rest of the SNPs missing genotypes we imputed using FastPhase-program (Stephens et al. (2001)).

EBVs. For each AI-boar EBVs of litter traits were calculated using a mixed model approach with herd-year, year-month, type of insemination, litter breed, and age at farrowing as fixed effects and litter sire, permanent environmental effect and animal itself as random effects (animal model BLUP). Accuracy of genomic selection was estimated for total number of piglets born in first (TNB1) and later parities (TNB2), number of still born piglets in first (NSB1) and later parities (NSB2), pig mortality before weaning in first (PM1) and in later parities (PM2)

Genomic breeding values. Genomic breeding values or direct genomic values (DGV) were calculated using SNP-BLUP approach applied in Mix99 program package (Lidauer and Strandén, (1999), Vuori et al., (2006)).

Validation. The data were divided into two sets: full and reduced following the example of Koivula et al. (2012). The full data included all observations up to 2012 and were used for calculation of EBVs of all AI-boars. The reduced data included observations up to 2005 and were used for calculation of EBVs of AI-boars that had daughters records available up to 2005. These AI-boars comprised the reference population (588 boars) that was used in SNP-BLUP to form the prediction equation for genomic

selection. The validation population included AI-boars that did not have daughters in the reduced data but had in the full data (134 boars). Thus, validation animals had EBVs from the full data, parent averages (PA) from the reduced data and DGVs based on own genotypes and the prediction equation from SNP-BLUP. Prior to analysis EBVs were deregressed (DRP) using the same approach as for the bulls by Interbull.

Validation reliability (R^2) of DGV and PA were assessed using the Interbull protocol (Mäntysaari et al. (2010)): $y_i = b_0 + b_1 g_i + e_i$, where for each validation animal i y_i is the DRP, g_i is either DGV or PA and e_i is the residual term. Validation reliability was obtained by dividing the coefficient of determination of the validation model by the average reliability of DRPs of the validation animals using the full data.

Results and Discussion

Table 1 gives the estimates of the validation reliabilities of SNP-BLUP for different sow reproduction traits. In general, DGVs had better reliabilities than PAs. As an example, an average reliability of PA for TNB2 was only half of that obtained for DGVs. Intercepts (b_0) were not different for zero. Also, intercept for parent average was close to zero indication that there were no bias in EBVs of the parents of the boar selected for AI. Regression coefficients (b_1) varied from 0.56 to 0.99 indicating general overdispersion in DGVs (see also Figure 1 for plots of DRP vs. DGV or PA).

Table 1. Validation measures (intercept b_0, regression coefficient b_1, and validation reliability (R^2) from SNP-BLUP and PA

Trait	b_0	b_1	R^2	r^2_{REF}	r^2_{VAL}
TNB1	-0.06	0.74	0.32	0.74	0.66
NSB1	-0.09	0.56	0.34	0.68	0.63
PM1	-0.08	0.72	0.37	0.61	0.56
TNB2	-0.07	0.88	0.44	0.75	0.68
NSB2	-0.10	0.74	0.55	0.70	0.65
PM2	-0.07	0.99	0.58	0.63	0.58
PA_TNB2	-0.16	0.70	0.20		

TNB1: Total number of piglets born in the first parity and in later parities (TNB2); NSB1: Number of still born piglets in the first parity and in later parities (NSB2); PM1: Piglet mortality before weaning in the first parity and in later parities (PM2); PA: parent average

Validation reliabilities of DGVs varied from 0.32 to 0.58. These values are relatively high given the small number of reference boars. A clear trend in these results is that the validation reliabilities were lower for reproduction traits related to first parity than those related to later parities. The difference was from 0.12 to 0.21 (Table 1). Also, the order of the validation reliabilities was the same in both traits groups (first parity vs. later parities): PM >

NSB > TNB. The size of the reference population was the same for all traits thus it did not have any effect on the obtained validation reliabilities. Even though the reliabilities of the DRPs were moderate they did not differ significantly between first parity and later parity traits. Also, reliabilities of the DRPs of the validation animals were similar between the first parity and later parity traits e.g. TNB1 and TNB2.
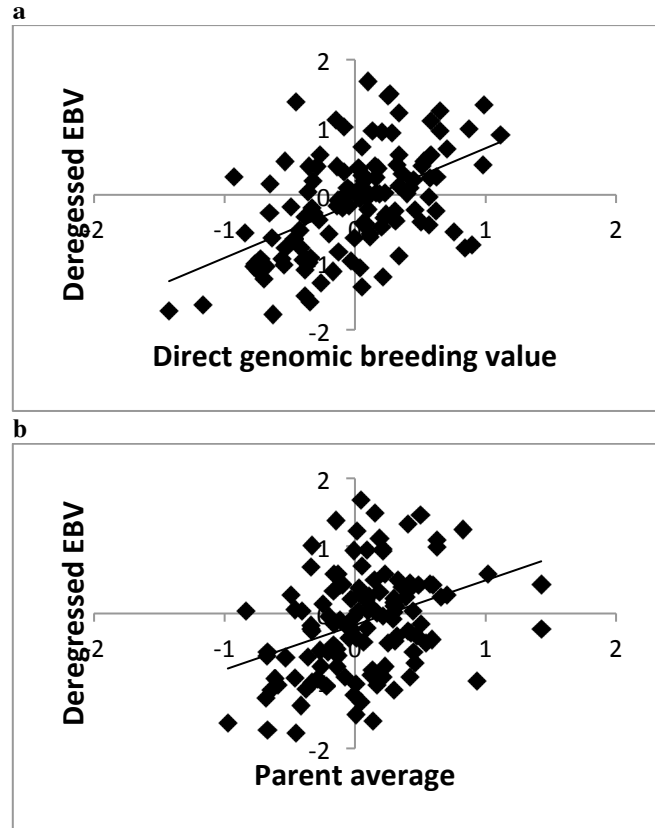
a

b

Figure 1. Plots of deregressed EBVs vs direct genomic values (a) or parent averages (b) for validation boars (Total number of piglets born in later parities, TNB2)

The validation procedure used in this study has some weaknesses that may have given too optimistic validation reliabilities. The prediction equation obtained from SNP-BLUP is based on deregressed proofs of the direct ancestors of the validation animals. Also, the estimated DGVs were compared to deregressed proofs. Thus there is a possibility for autocorrelation between the DGVs and DRPs of the validation animals that artificially can improve the validation reliabilities. The effect of the reliability of the reference population was not able to be studied because removing the boars with the lowest reliabilities would have reduced the size of the reference population that would have had a negative effect on the validation reliabilities. As a matter of fact, the effect of reliabilities of the reference boars on the prediction equation is already taken into account in SNP-BLUP where

reliabilities are used as weights. Using only the validation boars that had the highest reliability of DRP did not have an effect on the obtained estimates of the validation reliabilities.

The obtained validation reliabilities for the first parity traits were similar to those obtained for daily gain in Danish pig population but higher than those obtained for feed conversion ratio (Ostersen et al. (2011)). The number of genotyped pigs was higher in the Danish study and the average reliability of EBVs of the reference animals for daily gain was similar (0.62) to those reported in this study. The statistical method for estimating DGV was the same in this and in the Danish study but the validation method was different. Compared to validation results of the Nordic red breeds (Koivula et al. (2012)) our estimated validation reliabilities were higher than those reported for milk production traits.

Conclusion

Reliability of genomic selection (validation reliability) was estimated for sow reproduction traits using relatively small reference population. The obtained results were promising indicating that genomic selection can be applied in Finnish pig breeding scheme. A simple SNP-BLUP (GBLUP) with a two-step process for genomic selection should be beneficial especially for reproduction traits. An alternative option is to implement a one-step approach to include both genotyped and un-genotyped animals into one breeding value evaluation round (Misztal et al. (2009)).

Literature Cited

Koivula, M., Strandén, I., Su, G., and Mäntysaari, E. (2012) J. Dairy Sci. 95:4065-4073.

Ostersen, T., Christensen, O., Henryon, M. et al. (2011). Genet. Sel. Evol. 43:38.

Lidauer M. and Strandén, I. (1999). Interbull Bulletin No. 20:20-25.

Miztal, I., Legarra, A., and Aguilar, I. (2009) J. Dairy Sci. 92:4648-4655.

Stephens, M., Smith, N., Donelly, P. (2001) Am. J. Hum. Genet. 68:978-989.

Uimari, P. and Tapio, M. (2011) J. Anim. Sci. 89:609-614.

Vuori, K., Strandén, I., Lidauer M., and Mäntysaari, E.A. (2006) Proc 8th WCGALP Proceedings CD.