

Use of field data in genomic reference populations for pig breeding

M. Lillehammer¹, T.H.E. Meuwissen² and A.K. Sonesson¹.

¹Nofima, Norway, ²Norwegian University of Life Sciences

ABSTRACT: Potential benefits of genotyping production animals with phenotypes for a trait not routinely measured on close relatives of the selection candidates were studied by stochastic simulations. The population structure was similar to a typical pig population structure. The trait under investigation had low heritability, was measured late in life on production animals only and was negatively correlated to other traits in the breeding goal. Under such unfavorable conditions, genotyping production animals could not prevent this trait to get negative genetic gain or reduce the drop in genetic level significantly unless the economic weight of the trait in the nucleus was at least 50 % of the breeding goal. The genotyping had however some impact on the rate of inbreeding. If the traits were uncorrelated traits genetic gain increased for the trait under investigation and the effect of genotyping animals with phenotypes increased.

Keywords: Genomic selection; Reference population; Pig

Introduction

Genomic selection (Meuwissen et al. (2001)) is implemented or about to be implemented in many livestock species, facilitating more effective selection for traits that cannot be measured directly on selection candidates. The success of this method builds on the availability of a reference population with phenotypes for the trait(s) under selection and also genotypes. This is easily available for traits measured late in life, where animals genotyped as selection candidates can enter the reference population later, when they get phenotypes. It has been shown that pig breeding schemes could increase the genetic gain for maternal traits through genomic selection in the nucleus (Lillehammer et al. (2013)). For traits not measured in the nucleus, however, it is not clear which and how many animals need to be genotyped to obtain a good reference population. One benefit of genomic selection could be the possibility to select for traits where phenotypes are only available from production animals with a more distant relationship to the selection candidates. Such traits could be traits recorded under specific production conditions or disease resistance for diseases with occasional outbreaks, very important in certain environments, but less important in others, leading to low economic weight in the breeding nucleus.

The aim of this study was to estimate the benefit of genotyping animals that were grand-offspring of the breeding candidates in the nucleus and with specific phenotypes recorded under a genomic selection breeding scheme. The recorded trait was assumed to have low heritability, be measured late in life and to be uncorrelated or negatively correlated with the other traits in the breeding goal, to mimic a trait that would be very difficult to improve by tradi-

tional selection. Schemes with different numbers of genotyped production animals every generation were compared through stochastic simulations.

Materials and Methods

Simulated population structure. The population structure was chosen to reflect a typical pig breeding structure with a breeding nucleus, a multiplier and a production tier. All three tiers were self-recruiting their females. In the nucleus, 2000 dams were mated to 50 sires every generation to create 2000 male offspring and 18 000 female offspring. The males born in the nucleus were selection candidates both to become sires in the nucleus tier and in the multiplier tier. In the multiplier, 50 males were mated to 200 females to create 200 male and 400 female offspring. Males born in the multiplier tier were selection candidates to become sires in the production tier. In the production tier, 50 sires were mated to 5000 dams which obtained 5000 female offspring. No males offspring were simulated in the production tier, since sires for the production tier were recruited from the multiplier and phenotypes were available from the female offspring.

Trait properties. Two traits (N-trait and P-trait), each representing an index of traits, were simulated. N-trait was measured on all animals born in the nucleus, before selection and had heritability of 0.25. P-trait was measured on animals born in the production tier when the animals were one generation old to represent trait measured late in life and had heritability 0.1. The correlation between N-trait and P-trait was assumed to be -0.3. Selection in the nucleus was based on a breeding goal, consisting of both traits, with 90 % economic weight on N-trait and only 10 % economic weight on P-trait, assuming that the nucleus were to serve several markets, where P-trait is only relevant in a segment of the total market. In the multiplier and production tiers, 40 % weight was put on the P-trait and 60 % on the N-trait. Under these parameters, P-trait was expected to be very difficult to improve, and some alternative scenarios (Table 1) were run to test the sensitivity of the results for these assumptions.

Genomic information. Each animal was assumed to have 18 chromosomes, each with a length of 1M. Before starting the selection scheme, 2000 generations of random mating using 200 parents each generation were performed to create a base population in mutation-drift balance. Polymorphisms and recombinations were sampled as in Sonesson and Meuwissen (2009). Each chromosome was assumed to hold 100 QTL and 500 neutral markers. These were selected among the polymorphisms (SNPs) created during the simulation of the historical population. The QTL

were selected at random among SNPs with a minor allele frequency (MAF) above 0.05. QTL effects were sampled from a multivariate normal distribution. The SNPs with highest MAF among those not selected as QTL were selected to be neutral markers, i.e. known for genotyped animals. It was assumed that males born in the nucleus were genotyped for the neutral markers and selected on genomic breeding values, calculated using GBLUP (Meuwissen et al. (2001)), while female selection and selection in other tiers were based on conventional BLUP breeding values.

Reference populations. It was assumed that 2300 animals, which were the parents of the nucleus and multiplier animals in the first generation, with one offspring each with phenotypes for both traits were genotyped before the selection scheme started to create an initial reference population. The reference population for the nucleus trait was updated every generation with nucleus-boars with phenotypes and genotypes. For the P-trait, initially, no additional information to update the reference population was available (basic). Alternatively a number x of production animals with phenotypes was genotyped every generation (P_x) and used to update the reference population. The number x was varied from 0 to 5000 in the main scenario and was set to 0 or 1000 in the alternative scenarios. The selection schemes were run for 9 generations and genetic gain was evaluated for the different schemes from generation 4 to 9, omitting the first generations of selection where the initial reference population was strongly related to the selection candidates and hence had a large influence on the results. Each alternative was replicated 30 times.

Results and Discussion

Main scenario. Table 1 shows the genetic gain for the two traits, respectively for different number of genotyped production animals every generation within the main scenario. Genetic gain for the P-trait was negative, and genotyping production animals did not change the genetic gain of any of the traits considerably. This was a consequence of low heritability for the P-trait, unfavorable genetic correlation between the traits and low weight on the production-trait in the nucleus. Further, the production animals were only distantly related to the selection candidates because of the two generations between the nucleus and the production tier and the further generation that passed before the production animals got phenotypes, causing a total of 5 generations between the latest addition to the reference population and the selection candidates. Relationship between the reference population and the selection candidates have previously been shown to have large effects on the reliability of genomic breeding values (Soneson and Meuwissen (2009); Pszczola et al. (2012)). This study confirmed that it is not sufficient to genotype animals with phenotypes if the relationship between these animals and the selection candidates is poor, especially when dealing with a difficult trait with low heritability and which is negatively correlated to other traits in the breeding goal.

Genotyping production animals gave however decreased rate of inbreeding, indicating that there was some gain in selection accuracy. Controlling rate of inbreeding could however be done just as effectively using other strategies than genotyping lots of production animals every generation. It may be noted that our simulation model was perhaps pessimistic in that all QTL were affecting both traits. It was thus little suited to obtain positive responses for both traits, since there were no QTL that broke the genetic correlation (e.g. by affecting one trait but not the other).

Table 1. Alternative scenarios with different correlations between N-trait and P-trait (Corr), different heritability of P-trait (h^2) and different weight of P-trait in the multiplier and production tiers (EW P).

| Scenario | Corr | h^2 | EW P |
|----------|------|-------|-------|
| Main | -0.3 | 0.1 | 40 % |
| A | 0 | 0.1 | 40 % |
| B | 0 | 0.1 | 100 % |
| C | -0.3 | 0.25 | 40 % |
| D | -0.3 | 0.25 | 100 % |
| E | 0 | 0.25 | 40 % |
| F | 0 | 0.25 | 100 % |

Table 2. Genetic gain (σ_g) per generation for a trait measured in the nucleus (N-trait) and a trait measured on production animals (P-trait), and rate of inbreeding for different numbers of genotyped production animals under the main scenario.

| Strategy | Genetic gain | | Rate of inbreeding |
|-----------------------|--------------|---------|--------------------|
| | N-trait | P-trait | |
| Basic ¹ | -0.30 | 1.06 | 0.9 % |
| Prod250 ² | -0.30 | 1.03 | 0.9 % |
| Prod500 ² | -0.25 | 0.99 | 0.8 % |
| Prod1000 ² | -0.28 | 1.02 | 0.7 % |
| Prod2500 ² | -0.29 | 1.08 | 0.7 % |
| Prod5000 ² | -0.30 | 1.11 | 0.6 % |

¹A basic scheme where the reference population for P-trait consisted of only 1 progeny per nucleus sire.

²Prodx denotes a strategy where x production animals with P-trait records are genotyped per generation to obtain and update a reference population.

Table 3. Genetic gain (σ_g) per generation for a trait measured in the nucleus (N-trait) and a trait measured on production animals (P-trait) for varying economic weight (EW $_n$) of P-trait in the nucleus.

| EW $_n$ | N-trait | | P-trait | |
|---------|--------------------|--------------------|--------------------|--------------------|
| | Basic ¹ | P1000 ² | Basic ¹ | P1000 ² |
| 10 % | 1.03 | 0.98 | -0.29 | -0.29 |
| 20 % | 1.02 | 0.96 | -0.31 | -0.22 |
| 30 % | 1.03 | 0.97 | -0.25 | -0.24 |
| 40 % | 1.03 | 0.90 | -0.25 | -0.09 |
| 50 % | 0.95 | 0.82 | -0.25 | 0.03 |

¹A basic scheme where the reference population for P-trait consisted of only 1 progeny per nucleus sire.

²1000 production animals with P-trait records are genotyped per generation to obtain and update a reference population.

Alternative scenarios. To increase the economic weight of P-trait in the nucleus could be a way of obtaining positive gain for P-trait even under very difficult trait parameters. Table 3 shows that increased economic weight of P-trait shifted the genetic gain towards this trait and increased the effect of genotyping production animals on genetic gain. However, under these challenging trait parameters, 50 % weight had to be put on P-trait in the nucleus to obtain positive genetic gain for P-trait. This would only be realistic if P-trait was relevant for a large segment of the market, especially since the increase in genetic gain for P-trait resulted in a decrease in genetic gain for N-trait. Moreover, some simple selection-index calculations revealed that, if both traits could be measured on the nucleus animals and had 50/50 economic weights, still a negative response for the lowly heritability would be obtained. In this light, a positive response for the P-trait in the genomic selection scheme seems quite good.

The results for more favorable trait parameters are summarized in table 4, showing that if P-trait and N-trait were uncorrelated, positive genetic gain was obtained for P-trait. Economic weights in the multiplier and production tiers did not affect results, since the strongest selection happened in the nucleus. Heritability of P-trait was shown to be less important for the genetic gain and for the effect of genotyping.

Table 4. Genetic gain (σ_g) per generation for a trait measured in the nucleus (N-trait) and a trait measured on production animals (P-trait) for alternative trait parameters. The alternative scenarios are described in table 1.

| Scenario | N-trait | | P-trait | |
|----------|--------------------|--------------------|--------------------|--------------------|
| | Basic ¹ | P1000 ² | Basic ¹ | P1000 ² |
| A | 1.04 | 1.02 | 0.02 | 0.07 |
| B | 1.02 | 0.99 | 0.00 | 0.03 |
| C | 1.05 | 1.03 | -0.31 | -0.28 |
| D | 1.04 | 0.89 | -0.32 | -0.27 |
| E | 1.04 | 1.01 | 0.01 | 0.05 |
| F | 1.05 | 1.01 | 0.01 | 0.04 |

¹A basic scheme where the reference population for P-trait consisted of only 1 progeny per nucleus sire.

²1000 production animals with P-trait records are genotyped per generation to obtain and update a reference population.

Conclusion

To compensate for negative genetic gain for a trait where phenotypes were not routinely recorded in nucleus or closely related animals through genotyping production animals was shown to be strongly affected by the correlation between the this trait and the other traits in the breeding goal. For negatively correlated traits, obtaining positive genetic gain was shown to be difficult and require a high economic weight in the nucleus. In addition, a larger P-trait reference population and a denser marker map combined with BayesB type of estimation methods are expected to further improve selection accuracies of the P-trait.

Literature Cited

- Lillehammer, M., Sonesson, A. K., and Meuwissen, T. H. E. (2013). *J. Anim. Sci.* 91:3079-3087.
- Pszczola, M., Strabel, T., Mulder, H. A. et al. (2012). *J. Dairy Sci.* 95:389-400.
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). *Genetics* 157:1819-1829.
- Sonesson, A. K., and Meuwissen, T. H. E. (2009). *Genet. Sel. Evol.* 41:37.

