

Influence of the Accuracies Defined in an Optimal Selection Approach on the Outcomes of a Breeding Program

F.Ytournel^{*}, H. Simianer^{*}

Introduction

The evolution of the molecular and statistical methods provides the opportunity of a more precise knowledge of the genes involved in quantitative traits. This knowledge is now routinely used in breeding nuclei for some traits (Dekkers, 2004). However, the outcomes in the overall population are not considered in this selection. The reference method to define a breeding program is the gene-flow method (Hill, 1974; Elsen and Mocquot, 1974). The central element of this method is the so-called P-matrix which describes the transmission paths between the different cohorts, classically defined by the purpose of the group, the sex and age of the animals. One thus follows in the whole population the genes from a given path over the generations. We here considered each genotype as a subpart of the cohorts and accounted for the changes in the gene frequencies due to the selection in the breeding nuclei. These new frequencies were computed with a deterministic algorithm to optimize the genetic progress (Chakraborty *et al.*, 2002) adapted to overlapping generations (Ytournel and Simianer, 2009). The P-matrix was thus changed from one generation to the next. We considered a design where the breeding population was selected using optimum selection and added a population that was mated to the breeding males. We compared the natural genetic gain obtained with different accuracies of the polygene, obtained with genomic selection or a classical index. We also considered the achieved standard discounted expressions (SDE) as an evaluation of the breeding programs. These two parameters are the keys leading to the returns, which are one of the most important results when evaluating a breeding program.

Material and methods

Optimal selection with overlapping generations (Ytournel and Simianer 2009). The method aims at maximizing the weighted sum of the genetic progress resulting from the selection on a genetic value including known genes over a definite time horizon (thereafter set to 10 generations). The individuals in the breeding nuclei were selected on a combination of a polygenic value and the effects of known genes. We supposed two unlinked bi-allelic genes with effects 1.0 and 0.5 polygenic standard deviations ($\sigma_{P_{ol}}$). The initial frequencies of the favourable alleles were 0.2 and 0.3. The breeding males (BM) were subdivided into three age-classes (AC) while there were six AC among the breeding females (BF). The composition of the selection groups relatively to the age-classes was: (a) for the males: 0.50,

^{*} Animal Breeding and Genetics Group, Georg-August- University, 37075 Göttingen, Germany

0.30, 0.20 and (b) for the females 0.50, 0.23, 0.12, 0.08, 0.05 and 0.02. The selected proportions were 0.30, 0.20 and 0.90 (BM) and 0.80, 0.70, 0.60, 0.90, 0.90 and 0.90 (BF). The accuracies of the polygenic terms were calculated for heritabilities of 0.4 and 0.1. Accuracies obtained from three different sources of information were envisaged. First, we considered a classical breeding scheme (CS) with the performances of the parents and 8 full-sibs contributing to the estimation of the polygenic BV of the animals. We also considered two cases of genomic selection with 500 (GS₅₀₀) or 3000 (GS₃₀₀₀) animals in the training set. Accuracies were computed according to Daetwyler et al. (2008) with supposed 1000 genes responsible for the polygenic part. The resulting accuracies are presented in table 1.

Table 1: Accuracies achieved for the 2 selection indexes

| Heritability | CS | GS ₅₀₀ | GS ₃₀₀₀ |
|--------------|-------|-------------------|--------------------|
| 0.4 | 0.521 | 0.408 | 0.739 |
| 0.1 | 0.358 | 0.218 | 0.480 |

Several mating situations were possible. We assumed that the males would always have at least as good accuracies as the females. The 6 following mating systems were then envisaged: CSxCS, GS₃₀₀₀xGS₃₀₀₀, GS₅₀₀xGS₅₀₀, GS₃₀₀₀xCS, CSxGS₅₀₀ and GS₃₀₀₀xGS₅₀₀ with the breeding scheme for BM first.

For each combination heritability / mating situation, we recorded the total BV and the selection intensity of each genotype from each AC and computed the natural genetic gain.

Gene-flow. The population was composed of the 2 previous populations plus a female population (PF) with one AC. The origin of the genes is described through the synthetic gene-flow matrix presented in table 2. The trait was expressed in both first AC of the females (4% in BM and 96% in PF).

From the previously obtained genotype frequencies, we computed the haplotype frequencies transmitted to the offspring generation which was made possible by distinguishing the parental origin of the haplotypes. For the path PF > PF, we supposed that the frequencies corresponded to the probabilities of obtaining each haplotype from a given genotype and were thus treated as fixed.

We followed the genes originating from the first AC of the BM over the generations. We recorded the Standard Discounted Expressions (SDE) over the generations to evaluate how often discounted returns can be realized. The interest rate was 0.05 per year.

Table 2: Synthetic P-matrix

| | | From | | |
|----|----|---------|---------|---------|
| | | BM | BF | PF |
| To | BM | BM > BM | BF > BM | |
| | BF | BM > BF | BF > BF | |
| | PF | BM > PF | | PF > PF |

Results and discussion

Discounted total genetic gain. The results are presented in figure 1. As expected, the natural genetic gain increased with the heritability. It reached 4.20 σ_{Pol} for the mating scheme

GS₃₀₀₀×GS₃₀₀₀, and a heritability of 0.4, while it was only 0.39 σ_{Pol} for GS₅₀₀×GS₅₀₀ and $h^2=0.1$. When using different selection strategies for both sexes, the result ranked between those of each of both strategies applied to both sexes. However, there was hardly any difference between the designs CS×GS₅₀₀ and GS₃₀₀₀×GS₅₀₀, probably indicating that the lowest accuracy is the limiting factor for the whole selection process. The breeding scheme GS₃₀₀₀×GS₃₀₀₀ was the most sensitive to the reduction in accuracy: it lost 2.74 σ_{Pol} . The other breeding schemes lost between 1.81 σ_{Pol} and 0.81 σ_{Pol} .

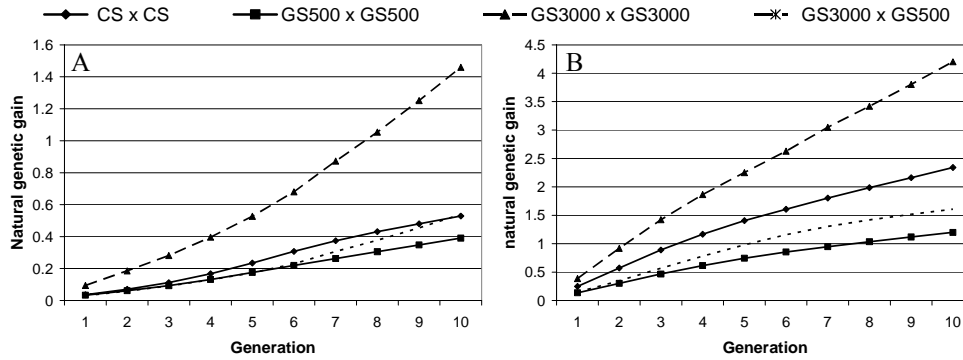


Figure 1: Natural genetic gain for the two heritabilities (A: 0.1 and B: 0.4) and 4 of the mating schemes

Gene-flow. We here only present the results for the paths BM>BM and BM>BF as the results were the same for the reciprocal paths (BM>BM and BM>BF). The achieved SDE in the breeding nucleus were comprised between 0.07 (GS₃₀₀₀×GS₃₀₀₀ with $h^2=0.4$) and 0.16 (GS₃₀₀₀×CS with $h^2=0.1$) (figure 2).

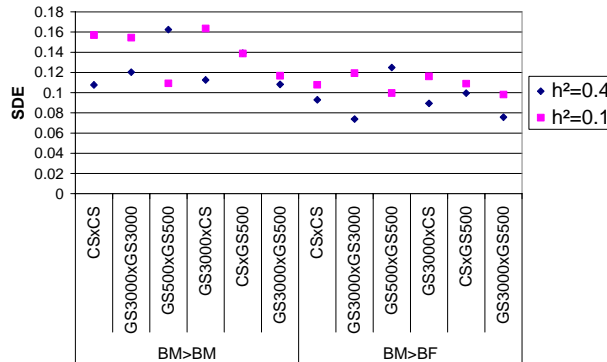


Figure 2: SDE-values achieved in the breeding nucleus for the paths BM>BM and BM>BF

It is not surprising to see lower SDE values in the cases where the most discounted profit was achieved. They correspond to the cases where the favourable genotype is fixed earlier, leading to a sparse P-matrix and thus reducing the amount of genes of the initial population still present after 10 generations. This is comforted by the schemes GS₅₀₀×GS₅₀₀ for both

paths where the highest SDE were reached with the highest heritability while, as previously shown, the achieved discounted total genetic gain was the lowest.

Studies classically attend at evaluating the achieved accuracy when using marker-assisted selection (Villanueva et al., 2005) or genomic selection (Meuwissen et al., 2001). However, the optimal selection method requires defining at the beginning the accuracies which can be estimated via simulations or previous studies, or calculated via deterministic formulae (e.g. Daetwyler et al., 2008). The present study shows the importance of correctly estimating this parameter relatively to the possibly reachable genetic gain.

The classical gene-flow method does not account for genotypes, and thus for changes in gene frequencies that are the consequence of selection (Latter, 1965). Our study shows the influence of these changes on the SDE, which are a key parameter when investigating a breeding program. Multiplied with the genetic gain and the economic values, they finally provide the discounted returns which are very important to each animal breeder.

Conclusion

The present study aimed at determining the influence of the accuracy of the polygenic breeding values provided to an optimal selection method (a) on the natural genetic gain, and (b) on the SDE-values. The results show that accuracies are one of the important parameters when evaluating different breeding programs when accounting for changes in the gene frequencies as both parameters are affected.

Acknowledgments

This study was conducted as part of the FUGATO+brain project and was financially supported by the German ministry for education and research (BMBF), Förderverein Biotechnologieforschung e.V., (FBF Bonn), Lohmann Tierzucht GmbH (Cuxhaven), Verband hannoverscher Warmblutzüchter e.V., Trakehner Verband, Verband des Oldenburger Pferdes e.V., Verband der Züchter des Holsteiner Pferdes e.V. and Westfälische Pferdestammbuch e.V..

References

- Chakraborty R., Moreau L., and Dekkers J.C.M. (2002). *Genet. Sel. Evol.*, 34:145-170.
- Daetwyler H.D., Villanueva B. and Woolliams J.A. (2008). *PLoS ONE*, 3(10):e3395.
- Dekkers J.C.M. (2004). *J. Anim. Sci.*, 82:E313-E328.
- Elsen J.M., and Mocquot, J.C. (1974). *Bull. tech. Dépt. Génét. Anim., INRA*, 17:30-54
- Hill W. (1974). *Anim. Prod.*, 18:117-139.
- Latter B.D. (1965). *Aust. J. Biol. Sci.*, 18:585-598.
- Meuwissen T.H.E., Hayes B.J., Goddard M.E. (2001). *Genetics*, 157(4): 1819–1829.
- Villanueva B., Pong-Wong R., Fernández J., et al. (2005). *J. Anim. Sci.*, 83:1747-1752.
- Ytournal F., and Simianer H. (2009). In *Book of abstracts of the 60th annual meeting of the European Association for Animal Production*, book of abstracts No.15, page 3.