Predicted Response of Genomic Selection for New Traits Using Combined Cow and Bull Reference Populations

***Mario P. L. Calus*[1]*, Yvette de Haas*[1]*, Marcin Pszczola*[1,2]*, and Roel F. Veerkamp*[1]**
[1]*Animal Breeding and Genomics Centre, Wageningen UR Livestock Research,*
8200 AB Lelystad, The Netherlands
[2]*Department of Genetics and Animal Breeding, Poznan University of Life Sciences,*
Wolynska 33, 60-637 Poznan, Poland

Introduction

Genomic selection is particularly beneficial for dairy cattle breeding programs, because it allows to significantly reduce generation interval, and cheaply increase selection intensity, while the accuracy of selection is only marginally lower compared to progeny testing schemes. It relaxes the requirement of traditional dairy cattle breeding schemes to measure phenotypes from progeny groups for each male selection candidate. Therefore, genomic selection holds the promise to allow selection for new traits, that are difficult or expensive to measure.

For traits that have been part of the breeding goal, currently, reference populations are composed of up to 16,000 daughter proven sires (Lund, *et al.*, 2010). For new more expensive traits reference populations may only be feasible for perhaps a few thousand cows with a single phenotypic measurement. Such reference population may be formed by combining data from different research herds, such as realized within the RobustMilk project. Examples of new traits are methane emission, energy balance, progesterone profiles as a proxy for fertility, and susceptibility to paratuberculosis (Johne's disease).

Small reference populations comprising only up to a few thousand cows with records, are expected to yield direct genomic values (DGV) with relatively low reliability. Since the generation interval is decreased with genomic selection, the selection response might still be noteworthy. Also, it might be possible to increase reliability by including information from correlated indicator traits in a bivariate analysis (Calus and Veerkamp, 2011), using a reference population with daughter proven bulls. Important questions are whether such approach would help to significantly increase accuracy of genomic selection for new traits, and what the response to genomic selection for those new traits is, even if their DGV reliability is low.

The objectives of this paper, therefore, are 1) to predict DGV reliability for small cow reference populations, that may be supplemented with sire reference populations using an indicator trait, and 2) to predict response to genomic selection for a new trait across the range of the DGV reliabilities.

Material and Methods

Evaluated traits

The evaluated simplified breeding program was based on an index modelled as a single trait with a heritability of 30%. The breeding program was executed using genomic selection, assuming that DGV of selection candidates have a reliability of 0.64 for the index. In this breeding program, a new trait was introduced. We considered a heritability of 30% for the new trait, and genetic correlations with the index (r_g(index, new trait)) of -0.5, 0, or 0.5. Both the index and the new trait had a phenotypic variance of 1.

Predicted DGV reliability

Expected reliabilities of genomic predictions were derived as explained below. First, reliabilities are predicted considering either only cow or bull data. Those reliabilities are then blended afterwards using selection index theory. Reliabilities using cow or bull data are predicted using the formula (Daetwyler, *et al.*, 2008; Daetwyler, 2009):

$$r_{g\hat{g}}^2 = q^2 \frac{\frac{n_p}{n_G}h^2}{\frac{n_p}{n_G}h^2+1} \qquad [1]$$

where q^2 is the proportion of the total genetic variance captured by the markers, n_p is the number of phenotypes used, n_G is the number of effective QTL or chromosome segments, and h^2 is the reliability of the trait. For q^2 we used a value of 0.8, which is reported for the commonly used 50k SNP chip (Daetwyler, 2009). Values for n_G were varied from 0 to 4,000 for cows, and was considered to be 0, 200, 500, 2000, 5000, 20,000 for bulls.

The reliability using cow data only (R_{cow}) was predicted using 0.3 for the heritability. To predict the reliability using bull data (R_{bull}) , the heritability used reflects the reliability of average offspring performances such as daughter yield deviations (VanRaden and Wiggans, 1991) or de-regressed proofs (Sigurdsson and Banos, 1995). We considered that bulls in the reference population had an estimated breeding value based on 100 daughters with records. Combined with the heritability of 0.3 using the formula $r_{IH}^2 = \frac{\frac{1}{4}nh^2}{1+\frac{1}{4}(n-1)h^2}$ (Mrode, 2005), this yields a reliability (r_{IH}^2) of 0.89, that was used as h^2 in the Daetwyler formula.

Since the cow and bull data partly explain the same variance, we used the ‘information source method’ (Harris and Johnson, 1998) to blend R_{cow} and R_{bull}. Considering the two pieces of information (R_{cow} and R_{bull}), the combined reliability (R) for the new trait is

$$R = \frac{R_{cow} + R_{bull} - 2R_{cow}R_{bull}}{1 - R_{cow}R_{bull}}$$

where

$$R_{bull}(new\ trait) =$$

$$r_g^2(index, new\ trait)R_{bull}(index)$$

Predicted response of genomic selection

In the evaluated breeding program, the new trait was considered to have the same economic value as the index. Note that results for a negative economic value for the new trait, while the genetic correlation between the index and the new trait is positive (e.g. 0.5), are the same as those for a genetic correlation of -0.5.

Numbers of cows and bulls included, were similar to the study of Schrooten *et al.* (2005). Each generation, 2,000 male and 2,000 female selection candidates were available for selection. For the males, 2% were selected, while for the females 20% were selected for use in the breeding program. In our study, we adopted single stage genomic selection. For the DGV of the new trait, the accuracy of selection was varied from 0.0 to 1.0 with steps of 0.1. Generation intervals of sires to breed bulls and cows, were considered to be 2 years. Generation intervals of dams to breed bulls and cows, were considered to be 3.33 and 5 years, respectively. Those values were adopted from Pryce *et al.*, (2010). All analyses were performed using SelAction (Rutten, *et al.*, 2002), following the guidelines from Dekkers (2007).

Results

Predicted DGV reliability

Predicted DGV reliabilities for the new trait are given in Figure 1, when an r_g(index, new trait) of 0.5 was assumed. Having 1,000 cows in the reference population, yielded a DGV reliability for the new trait of 0.18. Increasing the number of cows to 2,000, yielded a DGV reliability of 0.3. The same was achieved by adding 5,000 bulls with a DGV for the index to a cow reference population of 1,000 animals. The marginal contribution of additional bulls decreases quickly, when the number of already included bulls becomes larger. Note that a genetic correlation between the index and the new trait of -0.5 gives the same results as those in Figure 1. When r_g(index, new trait) = 0.0, the obtained reliability follows the same curve as for ‘0 bulls’ in Figure 1.

Predicted response of genomic selection

Predicted responses for genomic selection of the new trait are shown in Figure 2, in genetic standard deviations per year, considering simultaneous selection with the overall index. The vertical line in Figure 2 at a reliability of 0.3 indicates a reference population of 2,000 cows, or one of 1,000 cows supplemented with 5,000 bulls considering a genetic correlation of 0.5 (see Figure 1). Figure 2 indicates that a DGV reliability of 0.3 can already have an important impact on the genetic response of a new trait. For instance, with r_g(index, new trait) = -0.5, a DGV reliability of 0.3 is sufficient to prohibit a negative genetic trend in the new trait. When r_g(index, new trait) = 0.0, a DGV reliability of ~0.3 yields a genetic response that is half the genetic response for the index (results for the index are not shown).

Conclusion

Our results indicate that response to genomic selection for new traits can be important, even when the DGV reliability is much lower than commonly accepted in dairy cattle breeding programs. This relatively low DGV reliability is offset by a decrease in generation intervals. The response for the new trait depends heavily on the genetic correlation with the new trait.

Acknowledgement

The authors acknowledge funding for the study from the Dutch Dairy Board (PZ; Zoetermeer, the Netherlands) and AgentschapNL (The Hague, the Netherlands). The RobustMilk project is also acknowledged for financial support. The RobustMilk project is financially supported by the European Commission under the Seventh Research Framework Programme, Grant Agreement KBBE-211708. Marcin Pszczola gratefully acknowledges financial support of the Koepon Stichting (Arnhem, the Netherlands) and GreenHouseMilk. The GreenHouseMilk project is financially supported by the European Commission under the Seventh Research Framework Programme, Grant Agreement KBBE-238562. The content of this paper is the sole responsibility of the authors, and it does not necessarily represent the views of the Commission or its services.

References

Calus, M.P.L. & Veerkamp, R.F. 2011. Accuracy of multi-trait genomic selection using different methods. *Genet. Sel. Evol. 43,* 26.

Daetwyler, H., Villanueva, B. & Woolliams, J. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE. 3:e3395.*

Daetwyler, H.D. 2009. Genome-wide evaluation of populations. *PhD thesis,* Wageningen University, Wageningen.

Dekkers, J. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. *J. Anim. Breed. Genet. 124,* 331 - 341.

Harris, B. & Johnson, D. 1998. Approximate reliability of genetic evaluations under an animal model. *J. Dairy Sci. 81,* 2723- 2728. 2729.

Lund, M.S., De Roos, A.P.W., De Vries, A.G., Druet, T., Ducrocq, V., Fritz, S., Guillaume, F., Guldbrandtsen, B., Liu, Z., Reents, R., Schrooten, C., Seefried, M. & Su, G. 2010. Improving genomic prediction by eurogenomics collaboration. *Proceedings of 9th World Congress Genetics Applied to Livestock Production;* Leipzig.

Mrode, R. 2005. *Linear models for the prediction of animal breeding values.* 2nd edition ed. CABI Publishing.

Pryce, J.E., Goddard, M.E., Raadsma, H.W. & Hayes, B.J. 2010. Deterministic models of breeding scheme designs that incorporate genomic selection. *J. Dairy Sci. 93,* 5455-5466.

Rutten, M.J.M., Bijma, P., Woolliams, J.A. & van Arendonk, J.A.M. 2002. Selaction: Software to predict selection response and rate of inbreeding in livestock breeding programs. *J. Heredity 93,* 456-458.

Schrooten, C., Bovenhuis, H., van Arendonk, J.A.M. & Bijma, P. 2005. Genetic progress in multistage dairy cattle breeding schemes using genetic markers. *J. Dairy Sci. 88,* 1569-1581.

Sigurdsson, A. & Banos, G. 1995. Dependent-variables in international sire evaluations. *Acta Agric. Scand. Section A-Animal Science 45*, 209-217.

VanRaden, P.M. & Wiggans, G.R. 1991. Derivation, calculation, and use of national animal-model information. *J. Dairy Sci. 74*, 2737-2746.

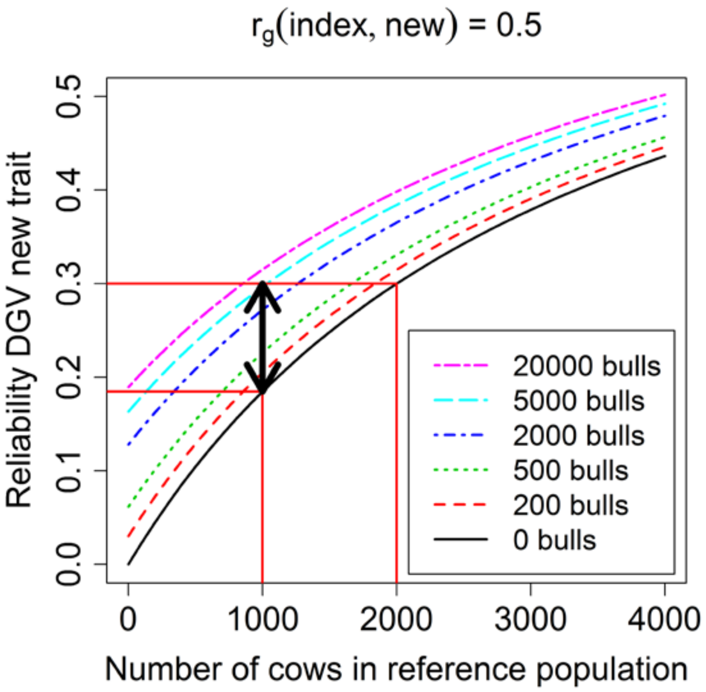

Figure 1. DGV reliability for a new trait, depending on the number of cows and bulls included in the reference population.

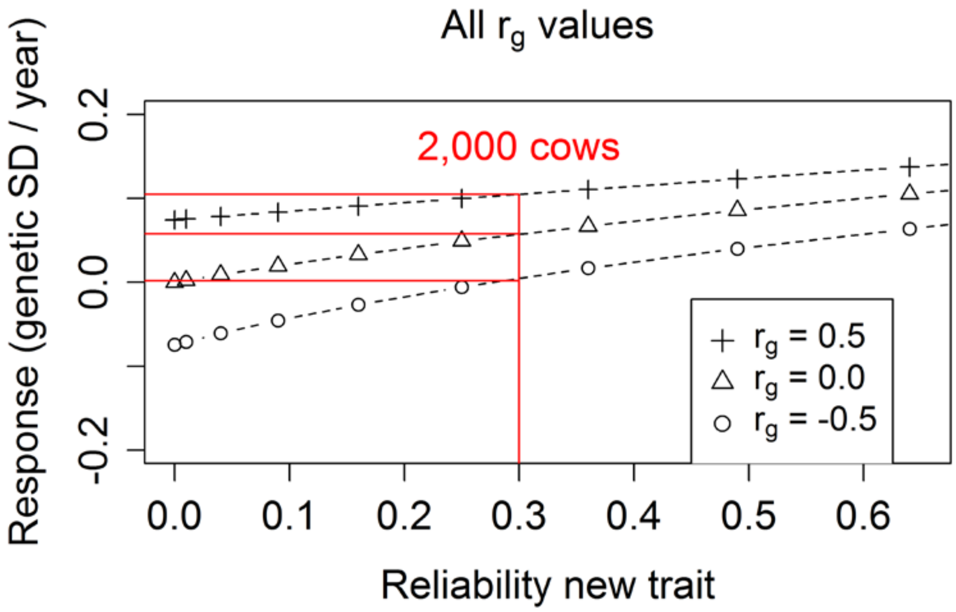

Figure 2. Response to genomic selection for the index and the new trait, depending on the genetic correlation with the overall index (r_g).