

Application of selection index calculations to determine selection strategies in genomic breeding programs

S. König*¹ and H. H. Swalve†

*Institute of Animal Breeding and Genetics, University of Göttingen, 37075 Göttingen, Germany

†Institute of Agricultural and Nutritional Sciences, University of Halle, 06099 Halle, Germany

ABSTRACT

The availability of genomic estimated breeding values (GEBV) allows for possible modifications to existing dairy cattle breeding programs. Selection index calculations including genomic and phenotypic observations as index sources were used to determine the optimal number of offspring per genotyped sire with a focus on functional traits and the design of cooperator herds, and to evaluate the importance of a central station test for genotyped bull dams. Evaluation criteria to compare different breeding strategies were correlations between index and aggregate genotype (r_{TI}), and the relative selection response percentage (RSR) of an index without single nucleotide polymorphism information in relation to a single nucleotide polymorphism-based index. The number of required daughter records per sire to achieve a predefined r_{TI} strongly depends on the accuracy of GEBV (r_{mg}) and the heritability of the trait. For a desired r_{TI} of 0.8, $h^2 = 0.10$, and $r_{mg} = 0.5$, at least 57 additional daughters have to be included in the genetic evaluation. Daughter records of genotyped sires are not necessary for optimal scenarios where r_{mg} is greater than or equal to r_{TI} . There still is a substantial need for phenotypic daughter records, especially for low-heritability functional traits and $r_{mg} < 0.7$. Phenotypic records from genotyped potential bull dams have no relevance for increasing r_{TI} , even with a low value for r_{mg} of 0.5. Hence, genomic breeding programs should focus on recording functional traits within progeny groups, preferably in cooperator herds. For low-heritability traits and with $r_{mg} > 0.7$, the RSR of conventional breeding programs was only 10% of RSR from genomic breeding strategies. As shown in scenarios including 2 traits in the index as well as in the aggregate genotype, the availability of highly accurate GEBV for production traits and low-accuracy GEBV for functional traits increased the risk of widening the gap between selection responses in production and functionality. Counteractions are possible,

such as via higher economic weights for low-heritability functional traits. Finally, an alternative selection strategy considering only 2 pathways of selection for genotyped male calves and for cow dams was evaluated. This strategy is competitive with a 4-pathway genomic breeding program if the fraction of selected male calves for the artificial insemination program is below 1% and if selection is focused on functionality, thus pointing to substantial insufficiencies caused by low reliabilities of breeding values for cows for such traits in conventional bull dam selection schemes.

Key words: genomic selection, breeding program, selection index

INTRODUCTION

Genetic selection in livestock populations, and especially in dairy cattle breeding programs, can currently be based on estimates of the association between large numbers of SNP and phenotypic information. This new technology is called genome-wide selection (**GS**) and is revolutionizing dairy cattle breeding programs (Hayes et al., 2009). Meuwissen et al. (2001) showed by simulation that this approach would be feasible and would yield correlations between true and SNP-based breeding values high enough to eliminate the necessity for traditional methods of breeding value estimation based on progeny testing in dairy cattle. Schaeffer (2006) outlined forms of a breeding program that apply this new technology. However, as several authors (e.g., Hayes et al., 2009; König et al., 2009; VanRaden et al., 2009) have pointed out, the magnitude of the correlation between the true breeding value and the genomic EBV (**GEBV**) is of crucial interest. In fact, this correlation, also referred to as the accuracy of GEBV (r_{mg}), may determine the extent to which genomic selection will be applied, and hence will be the crucial factor determining future breeding programs in dairy cattle.

The inclusion of genomic information in dairy cattle breeding programs should be viewed within the framework of other improvements. Current ideas for improvements of existing dairy cattle breeding programs in Germany are partly based on the establishment of

Received March 22, 2009.

Accepted July 9, 2009.

¹Corresponding author: skoenig2@gwdg.de

central station tests for potential bull dams (König et al., 2007a) and on cooperator herds for progeny testing (Gernand et al., 2007). Central station tests for potential bull dams, although creating substantial additional costs, are mainly implemented to prevent biases in EBV caused by preferential treatment (Kuhn et al., 1994), or to avoid the impact of heterogeneous variances in different environments (e.g., Garrick and van Vleck, 1987). Both the central station test of bull dams and progeny testing in selected cooperator herds generate additional advantages when the focus is on detailed recording of functional or health traits (König and Swalve, 2006; Swalve and König, 2007). An example is the application of selection indices to determine the genetic improvement of hoof health traits for different breeding strategies. König and Swalve (2006) found that the selection response for reducing incidences of laminitis per generation as well as the accuracy of EBV of bulls for laminitis could be doubled if direct observations on the laminitis status of 50 daughters were included as index sources. Their suggestion was to install cooperator herds for progeny testing to enable the routine recording of various types of disorders. Swalve (2008) gave an extended overview of the necessity of cooperator herds in dairy cattle breeding programs by considering aspects from data recording of functional traits up to aspects of the logistical advantages in the era of genomic selection. Recording of functional traits on a national scale is difficult to implement when daughters of young bulls are distributed over 49,202 dairy cattle farms (ADR, 2007) that keep registered herdbook cows.

However, both innovations (i.e., the implementation of a station test for bull dams and contracting cooperator herds for progeny testing) do not substantially alter the genetic framework of conventional dairy cattle breeding schemes, as suggested by Henderson (1964) or by Skjervold and Langholz (1964). A substantial modification of current dairy cattle progeny-testing programs seems to be possible through the application of GS. Schaeffer (2006) assumed an accuracy of 0.75 for GEBV and applied the general formula proposed by Rendel and Robertson (1950) for calculating annual genetic gain. Annual genetic gain was doubled because of a substantial reduction in of generation intervals, and the costs of proving bulls were reduced by 92% because of the elimination of progeny testing. König et al. (2009) used the same design for a genomic breeding program, but considered discounted costs and discounted returns over a whole investment period of 15 yr. For accuracies of GEBV above 0.70, a distinct economic advantage of genomic breeding programs, compared with progeny-testing programs, of up to a factor of 2.59 was found.

However, current practical results for estimated accuracies of GEBV (e.g., VanRaden et al., 2009) lagged behind the theoretical expectations obtained from simulation studies (e.g., Meuwissen et al., 2001). Especially for low-heritability functional traits, gain from genomic prediction compared with the parent average was relatively low (VanRaden et al., 2009). In this study, genomic information improved the accuracies of genetic values equivalent to 11 daughter records for traits with moderate heritability. These results imply continued progeny testing to ensure reliable genetic evaluations, as well as greater efforts in recording reliable phenotypes of functional traits for the derivation of SNP effects. Schaeffer (2006) suggested implementing a system of cooperator herds as a substantial tool for generating accurate phenotypes and genotypes, which could be financed by the \$23 million savings generated by using the GS strategy in the Canadian Holstein population.

From the current point of research (Hayes et al., 2009), the most likely breeding strategy remains a combination of both GS and progeny testing. The mixture of SNP information and phenotypic performance implies the need for an evaluation or even a modification of breeding strategies related to bull dam selection and progeny testing. Another even more progressive idea in the context of genomic breeding programs is a shift from bull dam selection schemes toward direct selection of genotyped male calves. This approach implies a 2-pathway selection scheme (i.e., the selection of male calves for AI and dams of cows within farms) instead of the classical 4-pathway strategy (Robertson and Rendel, 1950).

For dairy cattle breeding organizations, several questions arise when implementing genomic selection in breeding programs. In Germany, some of these questions are 1) the optimal size of progeny groups, especially when attempting to focus on the documentation of health traits in cooperator herds, 2) the abundance or the necessity of a central station test for potential bull dams, and 3) the possibility of a substantial shift from bull dam selection (2-step selection strategy) toward the direct selection of genotyped male calves as cow sires for AI. To answer these questions with a focus on functional traits, the aim of our study was to extend the calculations done by König and Swalve (2006) by integrating genomic information via selection index calculations. König and Swalve (2006) compared the genetic gain of direct selection strategies on health traits with conventional indirect approaches via indicator traits. Currently, an extension of such strategies is possible via genomic information. Results from different scenarios were used to recommend a suitable breeding strategy for a genomic breeding program.

MATERIALS AND METHODS

Selection index methodology was proven as a valid approach for calculating the selection response for different genomic selection criteria in the study by Dekkers (2007). In the present approach, this method was extended to specific scenarios relevant in current dairy cattle breeding programs when including GS. To answer the above-mentioned topics, 3 different forms of selection index formulae were developed. For all calculations, the SNP genotype was defined as “marker” ($\mathbf{m}$) and the observed phenotype as “observation” ($\mathbf{y}$), and were considered different but correlated traits. Additive genetic effects were denoted as $\mathbf{a}$. According to Lynch and Walsh (1998), the heritability for m was fixed to 1. This implies an equal genetic and phenotypic standard deviation for m , resulting in $\sigma_m = r_{mg} \cdot \sigma_a$ (Lande and Thompson, 1990), with r_{mg} denoting the correlation between m and the true genetic value ($\mathbf{g}$). The covariance between m and y can be written as $\sigma_{am} = a_{ij} \cdot r_{mg}^2 \cdot \sigma_a^2$ (Falconer and Mackay, 1996), where a_{ij} is the coefficient of additive genetic relationships between m on animal i and y on animal j . The breeding goal is the additive genetic value of the trait itself, resulting in an economic value of 1 for the performance trait, and is zero for the SNP information in the univariate case (e.g., one trait in the index, and the same trait in the aggregate genotype).

Scenario I for a Genotyped Bull Dam and the Bull Dam's Own Performance

The coefficient of relationship a_{ij} is 1 when the animals used in the index and in the aggregate genotype are identical, resulting in $\sigma_{am} = r_{mg}^2 \cdot \sigma_a^2$. This is the case when using a genotyped bull dam's own performance to predict its genetic value. Hence, the (co)variance matrix $\mathbf{P}$ for observations for this scenario was

$$\mathbf{P} = \begin{bmatrix} \sigma_y^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix}.$$

The (co)variance matrix $\mathbf{G}$ between observations and breeding values to be predicted was

$$\mathbf{G} = \begin{bmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix}.$$

Matrix $\mathbf{G}$ contains the (co)variances between information traits and breeding values.

Finally, the vector $\mathbf{w}$ of economic values has to be defined:

$$\mathbf{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The equation $b = \mathbf{P}^{-1} \mathbf{G} \mathbf{w}$ has to be solved. The coefficients for b describe the regression of the bull dam's true breeding value on its phenotypic performances and markers, resulting in b -values for y and m . As known from selection index theory, the variance in the aggregate genotype ($\mathbf{T}$) is $\sigma_T^2 = \mathbf{w}' \mathbf{G} \mathbf{w}$, and the variance for the index ($\mathbf{I}$) is $\sigma_I^2 = b' \mathbf{G} \mathbf{w}$. The ratio $r_{TI} = \sigma_I / \sigma_T$ is the correlation between the index and aggregate genotype, and is generally referred to as the accuracy of EBV.

For different correlations between the GEBV and the true breeding value (r_{mg} ; in the range from 0.5 to 0.9 in increments of 0.1) and for different heritabilities of the trait (0.01, and from 0.05 to 0.5 in increments of 0.05), r_{TI} and the selection response per generation ($\Delta \mathbf{G}$) with and without the inclusion of SNP information in the index were calculated. For comparison, the relative selection response in percentage (**RSR**) of an index without SNP information in relation to an SNP-based index was calculated as

$$RSR = \frac{\Delta G_{\text{without SNP information in the index}}}{\Delta G_{\text{including SNP information in the index}}} \cdot 100.$$

The main results (i.e., the gain in r_{TI} and the decrease in RSR when including SNP information of a genotyped bull dam) allow for assessment of the necessity of a central station test for bull dams when GS has been implemented.

Scenario II for a Genotyped Sire and a Different Number of Daughters

In scenario II, the number of daughters as an additional source of information for an already genotyped young sire is varied when estimating the breeding value. Hence, information sources in the index were the phenotypic observations of the daughters and the SNP genotype of the sire itself. These index sources were used to predict the aggregate genotype of the sire. Hence, the (co)variance matrix $\mathbf{P}$ of index sources, and again defining $\sigma_{am} = r_{mg}^2 \cdot \sigma_a^2$, was

$$\mathbf{P} = \begin{bmatrix} \frac{[1 + (n-1) \cdot 0.25 \cdot h^2]}{n} \sigma_y^2 & 0.5 \sigma_{am} \\ 0.5 \sigma_{am} & \sigma_m^2 \end{bmatrix},$$

where n is the number of daughters per bull. The (co) variance matrix $\mathbf{G}$ was

$$\mathbf{G} = \begin{bmatrix} 0.5\sigma_a^2 & 0.5\sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix}.$$

In this case, $\mathbf{G}$, the (co)variance matrix between observations and breeding values to be predicted, consists of 2 lines. The first line pertains to information from the bull's daughters, whereas the second line refers to genomic information on the bull itself.

Vector $\mathbf{w}$ was identical compared with scenario I:

$$\mathbf{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Because different animals are used in the index and in the breeding goal, the (co)variance matrix $\mathbf{C}$ for breeding values is different compared with $\mathbf{G}$ and was, for this scenario,

$$\mathbf{C} = \begin{bmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix}.$$

Matrix $\mathbf{C}$ needs to be defined for the calculation of $\sigma_T^2 = \mathbf{w}'\mathbf{C}\mathbf{w}$. Relative to scenario I, r_{mg} and heritabilities were varied in the same range and increments. Another variation considered the number of daughters (1, 10, 50, and 100) per bull. Ultimate evaluation criteria were r_{TI} , ΔG , and RSR. Equations from scenario II were used to determine the required number of daughters per bull to reach a predefined breeding value accuracy. These results should help to demonstrate the optimal number of daughters for genotyped sires by varying r_{mg} and the respective heritabilities. Hence, the results should be an aid in planning of the number and size of cooperator herds.

Scenario III for 2 Traits in a Combined Breeding Goal

Scenario III is an extension of scenario I because it includes 2 traits. This scenario uses the bull dam's own performances plus SNP information for both traits for the individual in the aggregate genotype. For a combined breeding goal including production and functionality, the traits considered were 305-d milk yield (**305MY**), and the interval from calving to first service (**CTFS**). For the SNP genotypes, the notations were **305MY_M** and **CTFS_M**, respectively.

According to Dekkers (2007), covariances for matrix $\mathbf{P}$ were derived as follows:

$$\sigma_{305MY:305MY_M} = r_{mg_{305MY}}^2 \cdot \sigma_{a_{305MY}}^2,$$

$$\sigma_{CTFS:CTFS_M} = r_{mg_{CTFS}}^2 \cdot \sigma_{a_{CTFS}}^2,$$

$$\sigma_{305MY:CTFS} = r_{p_{305MY:CTFS}} \cdot \sqrt{\sigma_{305MY}^2} \cdot \sqrt{\sigma_{CTFS}^2},$$

$$\sqrt{h_{305MY}^2} \cdot r_{mg_{CTFS}} \cdot r_{g_{305MY:CTFS}} \cdot \sqrt{\sigma_{305MY}^2} \cdot r_{mg_{CTFS}} \cdot \sqrt{\sigma_{a_{CTFS}}^2},$$

$$\sigma_{305MY_M:CTFS} =$$

$$\sqrt{h_{CTFS}^2} \cdot r_{mg_{305MY}} \cdot r_{g_{305MY:CTFS}} \cdot \sqrt{\sigma_{CTFS}^2} \cdot r_{mg_{305MY}} \cdot \sqrt{\sigma_{a_{305MY}}^2}, \text{ and}$$

$$\sigma_{305MY_M:CTFS_M} =$$

$$r_{g_{305MY:CTFS}} \cdot r_{mg_{305MY}} \cdot \sqrt{\sigma_{a_{305MY}}^2} \cdot r_{mg_{CTFS}} \cdot \sqrt{\sigma_{a_{CTFS}}^2},$$

resulting in

$$\mathbf{P} = \begin{bmatrix} \sigma_{305MY}^2 & \sigma_{305MY:305MY_M} & \sigma_{305MY:CTFS} & \sigma_{305MY:CTFS_M} \\ \sigma_{305MY:305MY_M} & \sigma_{305MY_M}^2 & \sigma_{305MY_M:CTFS} & \sigma_{305MY_M:CTFS_M} \\ \sigma_{305MY:CTFS} & \sigma_{305MY_M:CTFS} & \sigma_{CTFS}^2 & \sigma_{CTFS:CTFS_M} \\ \sigma_{305MY:CTFS_M} & \sigma_{305MY_M:CTFS_M} & \sigma_{CTFS:CTFS_M} & \sigma_{CTFS_M}^2 \end{bmatrix}.$$

Setting up matrix $\mathbf{G}$ for this bivariate problem requires genetic covariances to be known between 305MY and CTFS, 305MY_M and CTFS, 305MY_M and CTFS_M, and 305MY and CTFS_M. Additional definitions necessary were

$$\sigma_{a_{305MY:CTFS_M}} = r_{mg_{CTFS}} \cdot r_{g_{305MY:CTFS}} \cdot \sqrt{\sigma_{a_{305MY}}^2} \cdot \sqrt{\sigma_{a_{CTFS}}^2},$$

$$\sigma_{a_{305MY_M:CTFS}} = r_{mg_{305MY}} \cdot r_{g_{305MY:CTFS}} \cdot \sqrt{\sigma_{a_{305MY}}^2} \cdot \sqrt{\sigma_{a_{CTFS}}^2}, \text{ and}$$

$$\sigma_{a_{305MY:CTFS}} = r_{g_{305MY:CTFS}} \cdot \sqrt{\sigma_{a_{305MY}}^2} \cdot \sqrt{\sigma_{a_{CTFS}}^2}.$$

Matrix $\mathbf{G}$ was

$$\mathbf{G} = \begin{bmatrix} \sigma_{305MY}^2 & \sigma_{305MY:305MY_M} & \sigma_{305MY:CTFS} & \sigma_{305MY:CTFS_M} \\ \sigma_{305MY:305MY_M} & \sigma_{305MY_M}^2 & \sigma_{305MY_M:CTFS} & \sigma_{305MY_M:CTFS_M} \\ \sigma_{305MY:CTFS} & \sigma_{305MY_M:CTFS} & \sigma_{CTFS}^2 & \sigma_{CTFS:CTFS_M} \\ \sigma_{305MY:CTFS_M} & \sigma_{305MY_M:CTFS_M} & \sigma_{CTFS:CTFS_M} & \sigma_{CTFS_M}^2 \end{bmatrix}$$

Phenotypic $\left(r_{p_{305MY:CTFS}}\right)$ and genetic correlations $\left(r_{g_{305MY:CTFS}}\right)$ between 305MY and CTFS were assumed to be identical (-0.20) and were kept constant for all runs in scenario III. The heritability used was 0.30 for 305MY, and 0.05 for CTFS.

The vector $\mathbf{w}$ for economic weights was

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_{305MY} \\ \mathbf{w}_{305MY_M} \\ \mathbf{w}_{305MY} \\ \mathbf{w}_{305MY_M} \end{bmatrix}$$

Economic weights for the markers were set to be 0, and varied in 2 steps for 305MY and CTFS according to different ratios of economic values per genetic standard deviation. The ratios of economic values per standard deviation were 1.0, implying equal economic weights for both traits, and 0.5, meaning that the economic weight per genetic standard deviation for CTFS was 50% of the economic weight per standard deviation for 305MY. As an evaluation criterion for different breeding strategies, selection response in both single traits was calculated by altering selection intensities, r_{mg} , and economic values.

2-Pathway Selection Strategy

Correlations between indices and aggregate genotypes from different scenarios for the different pathways of selection were used to compare the selection response from a 2-step approach with the selection response from a 1-step approach. The 2-step approach consisted of preselection of bull dams according to their “conventional” EBV and a final selection step of bull dams according to GEBV. In the 1-step approach, male calves are selected directly based on their genotypes. Both breeding programs and the characteristics of their pathways of selection are outlined in Figure 1. Type a) is comparable to the design developed by Schaeffer (2006) but considers a 2-step selection procedure for

bull dams. Preselection of bull dams is based on pedigree BLUP-EBV. These preselected bull dams need to be genotyped, and their final selection (step 2) is based on GEBV. The selection response for this 2-step selection strategy was calculated by applying the formula $\Delta G = i_1 \cdot r_{TI_1} \cdot \sigma_{a_1} + i_2 \cdot r_{TI_2} \cdot \sigma_{a_2}$. Parameters i_2 , σ_{a_2} , and r_{TI_2} for selection step 2 were corrected for the selection in step 1 as described by Ronningen (1969). Selection intensity at step 2 (i_2) was derived from the number of final bull dams selected according to their GEBV in relation to all preselected bull dams, and was corrected for the skewness of the distribution function (Niebel and Fewson, 1976). Type b) of a genomic breeding program (Figure 1) is an alternative to bull dam selection strategies, specifically the direct selection of genotyped male calves for AI, thus abandoning all efforts regarding the traditional bull-dam path of selection. The background for developing such a scenario is based on a special situation in Germany and is discussed later. Breeding program type a) considers the 4 traditional pathways of selection as introduced by Robertson and Rendel (1950), whereas breeding program type b) is a simplification toward a 2-pathway strategy that includes the selection of male calves and the pathway for dams of cows.

RESULTS AND DISCUSSION

Scenario I for a Genotyped Bull Dam and the Bull Dam's Own Performance

As in other countries, attempts have been made in Germany over the past 2 decades to improve traditional selection strategies for bull dams. Several model calculations were undertaken to set up and optimize a central station test for potential bull dams (e.g., Kandzi, 1988; König and Swalve, 2003). These authors focused on the elimination of preferential treatment, currently an irrelevant topic when highly reliable GEBV for cows are available. Other arguments for supporting the implementation of an expensive central station test are summarized by Junge and Stamer (2006) and emphasize the possibility of recording additional functional traits. As shown in Figure 2, the availability of GEBV for cows also erodes this latter argument, because there is limited additional gain in increasing r_{TI} obtained from a cow's own phenotypic performance. Especially for functional traits with low heritabilities in the range from 0 to 0.10 without genomic information, r_{TI} is below 0.25. Breeding schemes such as central station tests for potential bull dams with a focus on the cows' own performances are advantageous only for highly heritable traits. This statement was valid in the past and is even more valid when applying genomic selection.

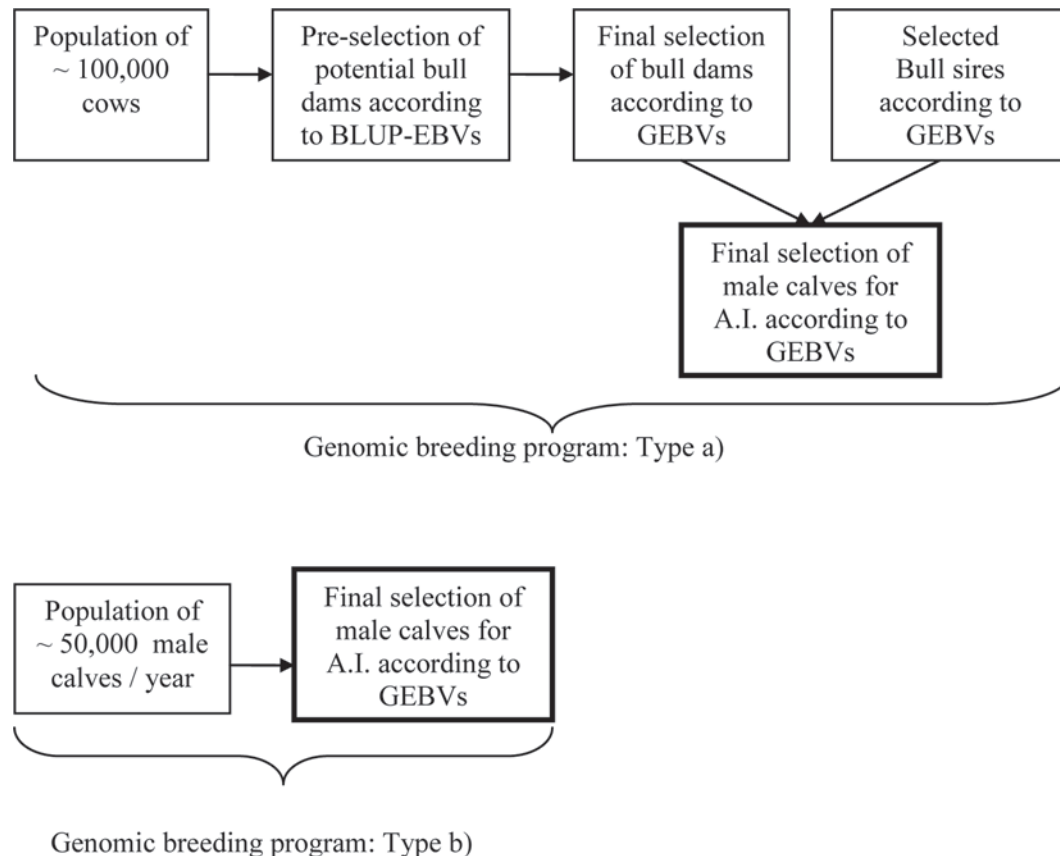


Figure 1. Comparison of 2 different genomic breeding programs: Type a) 2-step selection approach when considering bull dams; Type b) 1-step selection approach for the direct selection of male calves. GEBV = genomic EBV.

Correlations between indices and aggregate genotypes cannot fall below r_{mg} (Figure 2). For traits with heritabilities close to zero, values for r_{TI} and r_{mg} are identical. Hence, a cow's own phenotypic performance is competitive with a genomic approach only in situations with GEBV for traits that exhibit values for r_{mg}^2 lower than the realized (or observed) heritability.

The crucial parameter in improving breeding programs remains r_{mg} . However, even for low-heritability traits, a correlation of $r_{mg} = 0.5$, which has been used as the lowest value in our scenarios, is realizable when reviewing international results (Hayes et al., 2009) or when referring to simulation studies (Calus et al., 2008). Such crucial values for r_{mg} are also possible for low-heritability traits via the optimization of the number of animals to be genotyped. Formulas developed by Stricker and Fernando (2008) for Gaussian traits, or those derived by Daetwyler et al. (2008) for categorical traits have shown relationships between r_{mg} and the number of genotyped animals.

Analogous interpretations are valid when using RSR as the ultimate evaluation criterion (Figure 3) for breeding strategies of bull dams. For traits with

low heritability and high accuracy of GEBV, the RSR of conventional programs was only 10% of the genetic gain of genomic strategies. The importance of the cows' own phenotypic performance increases with increasing heritabilities and decreasing r_{mg} . For several German dairy cattle breeding organizations, one substantial goal when implementing central station tests for bull dams is the opportunity to improve the recording of functional traits. However, apart from the enhanced possibilities of improving the selection response for functional traits through a multitude of daughter records (König and Swalve, 2006) in a conventional breeding scheme, there is absolutely no requirement for testing bull dams on station in a genomic breeding scheme. Breeding organizations should invest in alternative breeding schemes that generate reliable data from daughter records as a source of reliable BLUP-EBV, which are necessary to estimate reliable SNP effects.

Scenario II for a Genotyped Sire and a Different Number of Daughters

The necessity for reliable data from daughter records can be obtained through the implementation of coop-

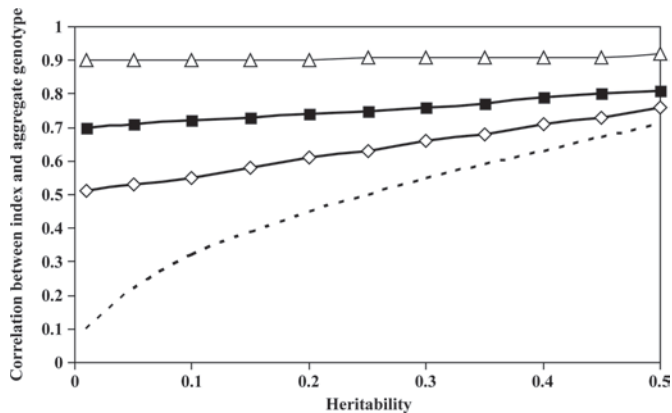


Figure 2. Correlation between index and aggregate genotype (r_{TI}) for scenario I by altering the heritability of the trait and the correlation between the true breeding value and genomic EBV (r_{mg}). Dotted line: index without genomic information; solid line with open diamonds: $r_{mg} = 0.5$; solid line with solid squares: $r_{mg} = 0.7$; solid line with open triangles: $r_{mg} = 0.9$.

erator herds for progeny testing. The use of cooperator herds has 2 advantages in a genomic breeding program. The first advantage is the generation of phenotypes for the estimation of SNP effects. The optimal number of daughters per sire and the resulting number of required herds for desired r_{mg} can be approximated by applying the formulae derived by Stricker and Fernando (2008) or Daetwyler et al. (2008).

The second, and even more important, question in this context addresses the optimal number of daughters for a genotyped sire to achieve a desired r_{TI} . The national German project to develop the methodology for estimating SNP or haplotype effects is a collaborative work that includes several scientific institutes and all Holstein dairy cattle breeding organizations (Thaller, 2007). However, the next step, the practical implementation of the genomic breeding program, may differ for the participating cattle breeding organizations. Details of centralized activities (e.g., related to the national genetic evaluations) and the competitive character of breeding organizations in Germany are outlined by König et al. (2007b). Hence, the primary intention of breeding organizations to establish cooperator herds within their region is to generate additional daughter records per sire to improve the accuracies of low-accuracy GEBV for functional traits.

Changes in r_{TI} depending on the heritability and r_{mg} for 50 daughter records and for 1 daughter record are shown in Figure 4 and Figure 5, respectively. When focusing on low-heritability functional traits, there is a substantial increase in r_{TI} when considering the sire's genotype as an index source. Differences in r_{TI} for scenarios with and without genomic information in the index increase with increasing r_{mg} and with a decrease-

ing number of daughters per sire (Figures 4 and 5). For moderate- and high-heritability traits, r_{TI} from 50 daughter records and without genomic information from the sire is as high as r_{TI} from genomic indices, supporting the conclusions of König and Swalve (2006) about the optimal design of cooperator herds in conventional breeding programs. They found that selection response in disease resistance per generation and accuracy of EBV of bulls for a trait with a heritability of 0.14 could be doubled when observations of 50 daughters were included as index traits.

The trends in RSR confirm these results (not shown). The RSR is 100% in a conventional breeding program for at least 50 daughters per sire and traits with high heritability (above 0.45), but RSR in conventional breeding programs cannot compete with genomic breeding programs for low-heritability functional traits.

The precise size of progeny groups for sires to achieve the desired values of r_{TI} were calculated by modifying heritabilities and r_{mg} (Table 1). Additional daughter records for genotyped sires are not necessary for optimal scenarios where $r_{mg} \geq r_{TI}$ (i.e., $r_{mg} = 0.9$ and $r_{TI} = 0.80$). Following the results from the most recent studies for estimating SNP effects (e.g., Hayes et al., 2009), an interpretation should focus on scenarios with $r_{mg} = 0.5$ and $r_{mg} = 0.7$. For these moderate correlations, there is still a necessity for additional daughter records to ensure a desired r_{TI} . Consider as an example a trait with a low heritability of 0.01 and an $r_{mg} = 0.5$. The critical value for breeder acceptance of genomic information is $r_{TI} = 0.80$ (König et al., 2009), and 581 daughters are needed to achieve this in this scenario, which is 129 less than in a conventional breeding pro-

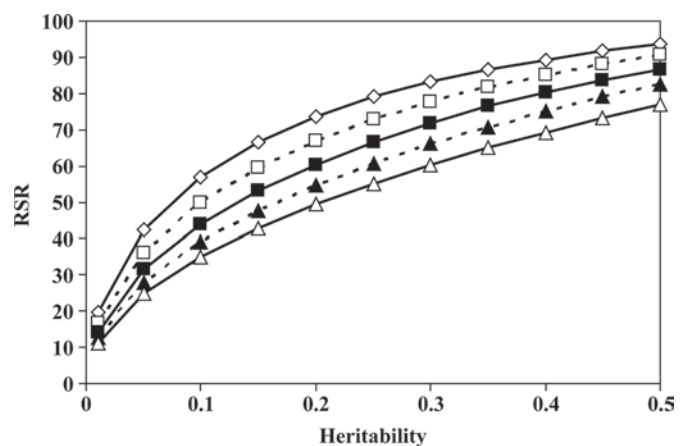


Figure 3. Relative selection response (RSR) for scenario I by altering the heritability of the trait, and the correlation between the true breeding value and genomic EBV (r_{mg}). Solid line with open diamonds: $r_{mg} = 0.5$; solid line with open squares: $r_{mg} = 0.6$; solid line with solid squares: $r_{mg} = 0.7$; solid line with solid triangles: $r_{mg} = 0.8$; solid line with open triangles: $r_{mg} = 0.9$.

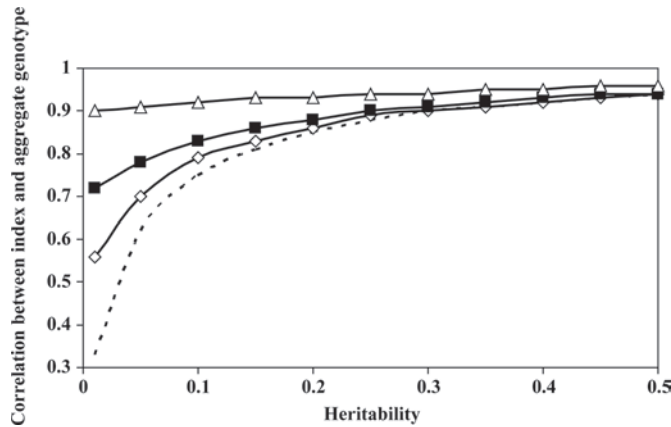


Figure 4. Correlation between index and aggregate genotype (r_{TI}) for scenario II and 50 daughter records per sire by altering the heritability of the trait and the correlation between the true breeding value and genomic EBV (r_{mg}). Dotted line: index without genomic information; solid line with open diamonds: $r_{mg} = 0.5$; solid line with solid squares: $r_{mg} = 0.7$; solid line with open triangles: $r_{mg} = 0.9$.

gram (Table 1). Apart from the scenarios with $r_{mg} = 0.9$ that were already discussed, the inclusion of genomic information is most beneficial for $r_{mg} = 0.7$. Here, the number of daughters necessary to achieve $r_{TI} = 0.80$ is halved. The general framework of a stringent genomic breeding program, as discussed by Schaeffer (2006) or König et al. (2009), does not include risk calculations that account for the loss of acceptance of genotyped sires by the dairy farmers when r_{TI} is decreasing, given that progeny tests should be abolished. However, the definition of a threshold for r_{TI} , with or without additional daughter records, may be a crucial parameter to encourage dairy cattle farmers to select genotyped sires for AI.

Considering genotyped daughters in the index would be another alternative for evaluation, but has limited practical relevance. Dairy farmers will not pay for genotyping “ordinary” cows, apart from potential bull dams. This is in agreement with Schaeffer (2006), who stated that the “dam of cow” pathway would be unaffected because these cows would not be genotyped.

Scenario III for 2 Traits in a Combined Breeding Goal

Application of scenario III allows the evaluation of a multitude of breeding strategies by altering genetic parameters, accuracies of GEBV, economic weights, and selection intensities. Figure 6 shows the selection response in genetic standard deviations for the 2 single traits 305MY and CTFS for 2 scenarios of different economic weights and a variation of replacement rates by keeping genetic parameters constant. For this analysis, 2 traits with moderate ($h^2 = 0.30$; 305MY) and low

heritability ($h^2 = 0.05$; CTFS) with an antagonistic relationship between them ($r_g = -0.20$; $r_p = -0.20$) were considered in the breeding goal with equal economic weights per genetic standard deviation in the first run. In the second run, the economic weight for CTFS was decreased by 50%. In first 2 runs, r_{mg} was 0.8 for 305MY and 0.5 for CTFS, and was 0.8 for both traits in the third scenario. High selection intensities or low replacement rates increase the absolute difference in selection response between these 2 antagonistic traits, but the ratio in selection response is independent of the replacement rate. For example, consider the scenario with $r_{mg} = 0.8$ for both traits and equal economic weights per genetic standard deviation. For a replacement rate of 5%, the selection response for 305MY and CTFS is 0.81 and 0.30 genetic standard deviations, respectively. For a replacement rate of 50%, the selection response for 305MY is 0.22 and is 0.08 for CTFS. Hence, the ratio of selection responses is 2.68 for both replacement rates.

Of more practical relevance is the impact of economic weights on the selection response for both traits. The differences increase with increasing economic weights for the more heritable trait (305MY), and decrease with increasing r_{mg} (results are not shown) for the low-heritability functional trait (CTFS). König et al. (2009) used a different approach for comparing selection response and annual monetary genetic gain in a genomic breeding program for a low- and a high-heritability trait. For a moderate accuracy of a genomic index of 0.70, annual monetary genetic gain was mostly due to the annual genetic gain in the high-heritability trait, whereas for extremely high accuracies of 0.90 or 0.99, annual genetic gain for both traits became

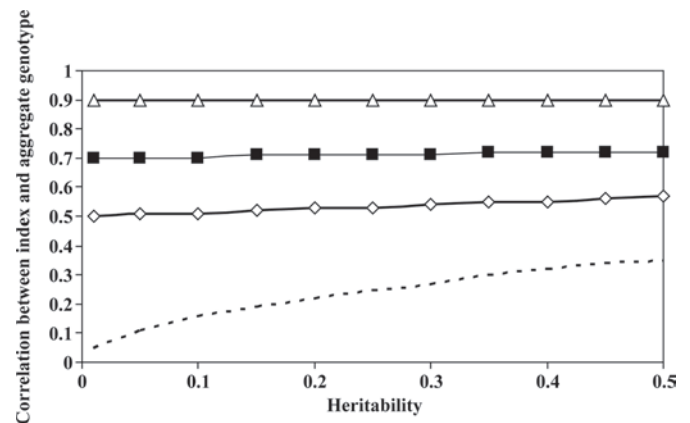


Figure 5. Correlation between index and aggregate genotype (r_{TI}) for scenario II and 1 daughter record per sire by altering the heritability of the trait and the correlation between the true breeding value and genomic EBV (r_{mg}). Dotted line: index without genomic information; solid line with open diamonds: $r_{mg} = 0.5$; solid line with solid squares: $r_{mg} = 0.7$; solid line with open triangles: $r_{mg} = 0.9$.

Table 1. Required number of daughters per genotyped sire to achieve predefined correlations between index and aggregate genotype (r_{TI}) for different heritabilities and different correlations between genomic EBV and true EBV (r_{mg})

h ²	Correlation between index and aggregate genotype (r_{TI})						Without genomic information	
	$r_{mg} = 0.5$		$r_{mg} = 0.7$		$r_{mg} = 0.9$			
	$r_{TI} = 0.8$	$r_{TI} = 0.95$	$r_{TI} = 0.8$	$r_{TI} = 0.95$	$r_{TI} = 0.8$	$r_{TI} = 0.95$	$r_{TI} = 0.8$	$r_{TI} = 0.95$
0.01	581	3,561	330	3,310	0	1,993	710	3,694
0.05	115	705	65	656	0	395	141	732
0.10	57	348	32	324	0	195	70	361
0.15	38	230	21	213	0	129	46	238
0.20	28	170	16	158	0	95	34	176
0.25	22	134	13	125	0	75	27	139
0.30	18	111	11	103	0	62	22	115
0.35	16	94	9	87	0	53	19	97
0.40	13	81	8	75	0	45	16	84
0.45	12	71	7	66	0	40	15	74
0.50	11	63	6	59	0	35	13	65

similar. Lower accuracies for GEBV for low-heritability functional traits compared with accuracies of GEBV for production traits increase the risk for an increasing gap in selection response between production and functionality. A decrease in this difference is possible by increasing economic weights for functional traits in the combined breeding goal. Another possibility to think about in this context is the application of desired gain indices (e.g., Brascamp, 1984), in which the rate of genetic gain can be predetermined. However, Gibson and Kennedy (1990) have shown that desired gain indices can cause substantial losses in genetic gain when the ultimate goal is to improve economic merit. This is the common goal in dairy cattle breeding and may be a mixture using selection indices based on economic values, and selection indices with constraints, as suggested by Brascamp (1984), can solve some problems. On the other hand, from a practical point of view, dairy cattle breeding programs will shift substantially from production toward functionality in the combined breeding goal (Miglior et al., 2005). Hence, an increase in economic weights for functional traits can be expected, which will make the gap in genetic gain between production and functionality smaller, even for a genomic breeding program.

2-Pathway Selection Strategy

Figure 7 shows the selection response in genetic standard deviations per year for a 4-pathway genomic breeding program and for a 2-pathway genomic breeding program in relation to replacement rates. The term “replacement rate” denotes the fraction of preselected potential bull dams in relation to all registered cows in the 4-pathway genomic breeding program, and also denotes the fraction of male calves directly selected

according to GEBV when applying the 2-pathway approach. These replacement rates were derived from selection intensities and vice versa, and were altered for preselection of bull dams and for the final selection of male calves in a range from 0 to 100% (Figure 7). The chosen scenario for the 4-pathway strategy (Table 2) was adopted from Schaeffer (2006), but by considering preselection of bull dams according to conventional BLUP-EBV. These bull dams need to be genotyped, and final selection is based on GEBV. Accuracies of GEBV for all genotyped groups (i.e., bull dams, bull sires, and cow sires) were assumed to be 0.75. An identical accuracy of 0.75 was used to select genotyped male calves in the 2-pathway breeding program. The idea for such a simplified scheme was taken from the fact that, in Germany, breeding organizations are often also involved in trading of commercial and slaughter cattle. This also pertains to trading of male calves at the age of 2 wk. These conditions in turn would thus allow the genotyping of these calves before reselling them as feeder cattle because, at that moment, they are owned by the breeding organizations. Genotyping a large number of male calves would imply high selection intensities and a high possibility of finding the top bulls for AI by using a direct approach. Because of low accuracies or even because of the unavailability of BLUP-EBV of functional traits for cows, practical preselection of potential bull dams in the 4-pathway program is mainly focused on production. This implies relatively high selection intensities for production but extremely low selection intensities for functionality at selection step 1. Hence, direct selection on genotyped male calves would result in distinct gains in the selection response for such scenarios, for which selection intensities and accuracies of BLUP-EBV of cows are quite low.

Table 2. Parameters for calculating selection response in a genomic breeding program for four pathways of selection

Item	Replacement rate (p , in %)	Selection intensity (i)	Accuracy (r_{TI})	Generation interval (L)
Sire of bulls	5	2.06	0.75	2.1
Sire of cows	10	1.40	0.75	2.1
Dams of cows	80	0.35	0.50	5.5
Dams of bulls (step 1)	Varied	Varied	0.50	2.0
Dams of bulls (step 2)	50	According to Niebel and Fewson (1976)	0.75	2.0

Results from Figure 7 clearly depict the advantage of a genomic breeding program according to Schaeffer (2006) for all pair-wise comparisons in replacement rates. For identical replacement rates of bull dams and male calves, the advantage of the 4-pathway model in selection response increases with decreasing selection intensities. However, the comparison should focus on the practical situation. For example, for low selection

intensities for bull dams in the step of preselection for a functional trait (replacement rates higher than 0.8), selection response using the 4-pathway program with a replacement rate of 0.8 for bull dams is 0.31 genetic standard deviations per year. The same selection response of 0.31 genetic standard deviations per year can be achieved in the 2-pathway model if the replacement rate for male calves using GEBV is lower than 0.01. A

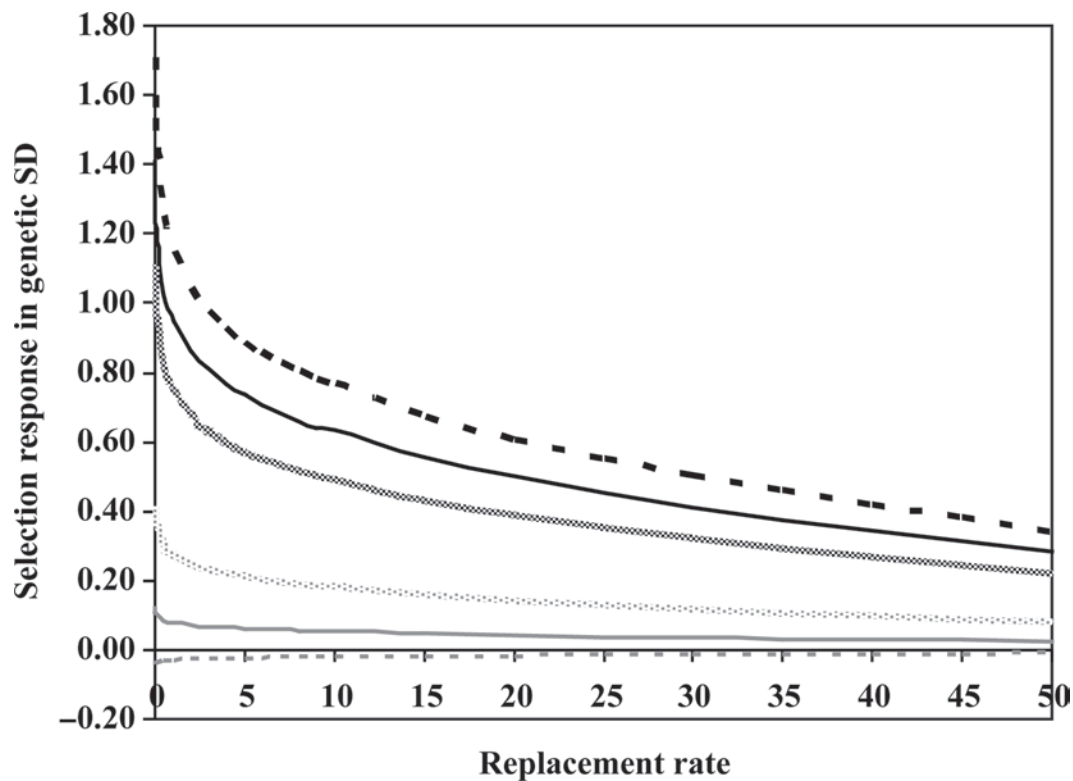


Figure 6. Selection response (ΔG) per generation in genetic standard deviations for 305MY and CTFS by altering replacement rates, economic weights, and correlations between the true breeding value and genomic EBV (r_{mg}). Solid black line and solid gray line: ΔG for 305-d milk yield (305MY) and ΔG for calving to first service (CTFS), respectively, for the scenario equal economic weights per genetic standard deviation, $r_{mg_{305MY}} = 0.8$, $r_{mg_{CTFS}} = 0.5$; dotted black line and dotted gray line: ΔG for 305MY and ΔG for CTFS, respectively, for the scenario economic weight per genetic standard deviation for CTFS = 50% of the economic weight per genetic standard deviation for 305MY, $r_{mg_{305MY}} = 0.8$, $r_{mg_{CTFS}} = 0.5$; dotted black band and dotted gray band: ΔG for 305MY and ΔG for CTFS, respectively, for the scenario equal economic weights per genetic standard deviation, $r_{mg_{305MY}} = 0.8$, $r_{mg_{CTFS}} = 0.8$.

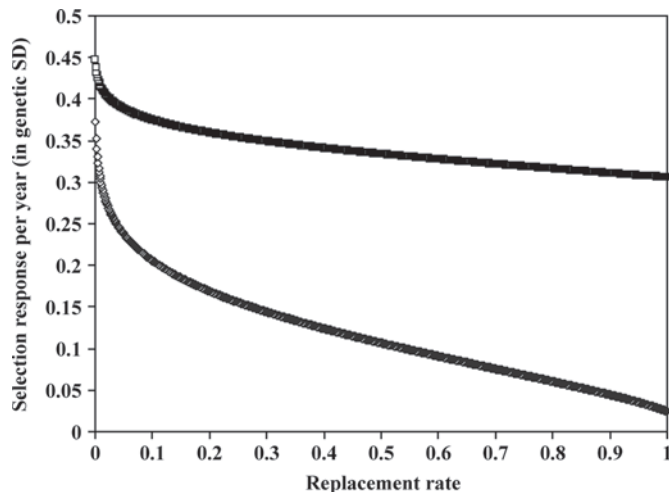


Figure 7. Selection response per year (in genetic standard deviations) for a genomic breeding program considering 4 pathways of selection (squares), and a genomic breeding program considering 2 pathways of selection (diamonds) by altering replacement rates for bull dams in selection step 1 and for male calves, respectively.

replacement rate of 1% out of a population of 50,000 male calves would generate 500 potential sires for AI. From a theoretical point of view, even replacement rates lower than 0.0001 could be envisaged. For a regionally operating breeding organization normally encountered in Germany, this might imply the genotyping of all male calves to find the top 5 sires. However, the 2-pathway selection strategy has potential in the future under the assumption of decreasing costs for genotyping and an increasing importance of functional traits. A decrease in the cost of genotyping would also be possible when using low-density SNP chips.

The 2-pathway selection strategy cannot compete with breeding goals focusing on production traits. The BLUP-EBV for production traits are available to preselect potential bull dams, and the possibility of recruiting the genetically valuable cows for production in step 1 is considerable. Genomic EBV for all traits are used for the final selection of bull dams. However, without the selection pressure on functional traits in step 1, compensation via intensive selection in step 2 is not possible. Theoretically, there would also be the possibility of genotyping all cows, resulting in high selection intensities for functional traits. However, for several reasons, the advantages of genotyping and selecting male calves are as follows: 1) male calves are owned by breeding organizations, which are also involved in cattle trading, but bull dams belong to the farmers; 2) genotyping bull dams and bull sires is a way to create elite matings, but genotyping male calves is a direct selection strategy focused on the final product; and 3) several interesting bull dams are not available for the German market, but

imported embryos or male offspring can be used via the 2-pathway approach.

Depending on the costs for genotyping, the effectiveness of bull dam selection, the accuracies of GEBV, the defined breeding goal, and, of course, the importance of livestock trade for breeding organizations, the 2-pathway approach should be taken into account and should be considered as an alternative breeding strategy for distinct circumstances. Furthermore, given the fact that the dams of cows pathway does not contribute significantly to genetic progress, neglecting this pathway completely would essentially give the same answers in the model calculations undertaken in the work presented here. Thus, under specific circumstances and cost situations, the era of GS that has now begun for dairy cattle breeding programs could bring a return of the 1-path model when assessing the possible response to selection.

CONCLUSIONS

In general, it can be concluded from this study that selection index calculations can be used to evaluate a broad variety of breeding strategies when genomic information is available. Distinct answers were found to assess the optimal size of progeny groups for genotyped sires with respect to cooperator herds and to the future role of central station tests for potential bull dams. Even for relatively low accuracies of GEBV, there is no longer a need to test genotyped potential bull dams on station to improve the accuracies of an index that includes genomic information. Phenotypic daughter records can improve those accuracies in an index for genotyped bulls for relatively low accuracies of GEBV below 0.80. Hence, a shift from station tests of bull dams toward the establishment of cooperator herds for progeny is recommended to increase the selection response for low-heritability functional traits. However, the availability of highly accurate GEBV for production traits and low-accuracy GEBV for functional traits also increases the risk of widening the gap between selection responses in production and functionality. Counteractions are possible, such as via higher economic weights for the low-heritability traits. For insufficiencies in bull dam selection strategies, an alternative genomic selection strategy considering only the 2 pathways for male calves and dams of cows was suggested. This strategy requires a large pool of genotyped male calves owned by breeding organizations and extremely high selection intensities when selecting according to GEBV.

REFERENCES

- ADR. 2007. Annual Statistics. German Cattle Breeders Federation, Bonn, Germany.

- Brascamp, E. W. 1984. Selection indices with constraints. *Anim. Breed. Abstr.* 52:645–654.
- Calus, M. P., T. H. Meuwissen, A. P. de Roos, and R. F. Veerkamp. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178:553–561.
- Daetwyler, H.D., B. Villanueva, and J. A. Wooliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE* 3(10):e3395.
- Dekkers, J. C. M. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. *J. Anim. Breed. Genet.* 124:331–341.
- Falconer, D. S., and T. C. F. Mackay. 1996. *Introduction to Quantitative Genetics*. Longman, Harlow, UK.
- Garrick, D. J., and L. D. van Vleck. 1987. Aspects of selection for performance in several environments with heterogeneous variances. *J. Anim. Sci.* 65:409–421.
- Gernand, E., R. Waßmuth, U. U. v. Borstel, and S. König. 2007. Estimation of variance components for production traits in large-scale dairy farms. *Livest. Sci.* 112:78–89.
- Gibson, J. P., and B. W. Kennedy. 1990. The use of constrained selection indexes in breeding for economic merit. *Theor. Appl. Genet.* 80:801–805.
- Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92:433–443.
- Henderson, C. R. 1964. Selecting the young sire to sample in artificial insemination. *J. Dairy Sci.* 47:439–441.
- Junge, W., and E. Stamer. 2006. Stand der Bullenmutterprüfung auf Station in Deutschland. *Rinder-Workshop Uelzen* 4:148–153.
- Kandzi, A. 1988. Untersuchungen zu einem offenen Nukleuszuchtprogramm beim Milchrind. PhD Diss. Göttingen University, Göttingen, Germany.
- König, S., F. Bosselmann, U. U. v. Borstel, and H. Simianer. 2007a. Genetic analysis of traits affecting the success of embryo transfer in dairy cattle. *J. Dairy Sci.* 90:3945–3954.
- König, S., S. Lessner, and H. Simianer. 2007b. Application of controlling instruments for improvements in cow sire selection. *J. Dairy Sci.* 90:1967–1980.
- König, S., H. Simianer, and A. Willam. 2009. Economic evaluation of genomic breeding programs. *J. Dairy Sci.* 92:382–391.
- König, S., and H. H. Swalve. 2003. Optimum design of an on station test for potential bull dams. Page 74 in *Proc. 54th Annu. Mtg. Eur. Assoc. Anim. Prod.*, Rome, Italy. Wageningen Academic Publishers, Wageningen, the Netherlands.
- König, S., and H. H. Swalve. 2006. Modellkalkulationen zu züchterischen Möglichkeiten auf Klauengesundheit beim Milchrind. *Züchtungskunde* 78:345–356.
- Kuhn, M. T., P. J. Boettcher, and A. E. Freeman. 1994. Potential bias in transmitting abilities of females from preferential treatment. *J. Dairy Sci.* 77:2428–2437.
- Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124:743–753.
- Lynch, M., and B. Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Assoc. Inc., Sunderland, MA.
- Meuwissen, T. H. E., B. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829.
- Miglior, F., B. L. Muir, and B. J. Van Dormal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255–1263.
- Niebel, E., and D. Fewson. 1976. Untersuchungen zur Zuchtplanung für die Reinzucht beim Zweinutzungsrand. II. Zuchtwahl in zwei Selektionsstufen. *Z. Tierz. Züchtungsbiol.* 93:169–177.
- Rendel, J. M., and A. Robertson. 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *J. Genet.* 50:1–8.
- Robertson, A., and J. M. Rendel. 1950. The use of progeny testing with artificial insemination in dairy cattle. *J. Genet.* 50:21–31.
- Ronningen, K. 1969. Studies on selection in animal breeding. 1. The efficiency of two-stage selection compared with single-stage selection with respect to progeny testing in animal breeding. *Acta Agric. Scand.* 19:149–174.
- Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223.
- Skjervold, H., and H. J. Langholz. 1964. Factors affecting the optimum structure of AI breeding in cattle. *Z. Tierzucht. Züchtungsbiol.* 80:25–40.
- Stricker, C., and R. L. Fernando. 2008. Genomewide genetic evaluation: How many individuals to genotype? *Int. Postgraduate Course and Workshop on Whole Genome Association and Genomic Selection*, Salzburg, Austria.
- Swalve, H. H. 2008. Recording of functional traits in contract herds for progeny testing of bulls in dairy cattle breeding programs. Page 210 in *Proc. 59th Annu. Mtg. Eur. Assoc. Anim. Prod.*, Vilnius, Lithuania. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Swalve, H. H., and S. König. 2007. Testherden in zuchtprogrammen für Milchrinder. 1. Mitteilung: Allgemeine Überlegungen. *Züchtungskunde* 79:249–263.
- Thaller, G. 2007. *GenoTrack: Ein Forschungsprojekt zur genomischen Selektion und mehr*. *Milchrind* 4:4.
- VanRaden, P. M., C. P. Van Tassel, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16–24.