Animal (2012), **6:9**, pp 1369–1376 © The Animal Consortium 2012
doi:10.1017/S1751731112000626

animal

Integration of genomic information into sport horse breeding programs for optimization of accuracy of selection

A. M. Haberland[1†], U. König von Borstel[1], H. Simianer[1] and S. König[2]

[1]*Department of Animal Sciences, University of Goettingen, 37075 Goettingen, Germany;* [2]*Department of Animal Breeding, University of Kassel, Nordbahnhofstraße 1a, 37213 Witzenhausen, Germany*

(Received 2 March 2011; Accepted 6 January 2012; First published online 21 March 2012)

Reliable selection criteria are required for young riding horses to increase genetic gain by increasing accuracy of selection and decreasing generation intervals. In this study, selection strategies incorporating genomic breeding values (GEBVs) were evaluated. Relevant stages of selection in sport horse breeding programs were analyzed by applying selection index theory. Results in terms of accuracies of indices (r_{TI}) and relative selection response indicated that information on single nucleotide polymorphism (SNP) genotypes considerably increases the accuracy of breeding values estimated for young horses without own or progeny performance. In a first scenario, the correlation between the breeding value estimated from the SNP genotype and the true breeding value (= accuracy of GEBV) was fixed to a relatively low value of $r_{mg} = 0.5$. For a low heritability trait ($h^2 = 0.15$), and an index for a young horse based only on information from both parents, additional genomic information doubles r_{TI} from 0.27 to 0.54. Including the conventional information source 'own performance' into the before mentioned index, additional SNP information increases r_{TI} by 40%. Thus, particularly with regard to traits of low heritability, genomic information can provide a tool for well-founded selection decisions early in life. In a further approach, different sources of breeding values (e.g. GEBV and estimated breeding values (EBVs) from different countries) were combined into an overall index when altering accuracies of EBVs and correlations between traits. In summary, we showed that genomic selection strategies have the potential to contribute to a substantial reduction in generation intervals in horse breeding programs.

Keywords: accuracy of selection, breeding strategies, generation interval, genomic selection, sport horse

Implications

The availability of genomic information demands proper assessment of its impact on practical horse breeding programs. Accuracies of conventional breeding values do not increase significantly until a stallion is aged 8 to 12 years and his progeny enters competition. We showed that additional genomic information considerably increases the accuracy of breeding values estimated for foals, young horses without own performance, and horses without progeny performance. Therefore, genomic selection (GS) enables selection at an earlier stage, shortening generation intervals and opening room for increased genetic progress. Our results indicate that horse breeding organizations could likely benefit from the application of GS.

Introduction

Sport horse breeding programs are characterized by long generation intervals and suboptimal selection intensities (Philipsson *et al*., 1990; Niemann, 2009) because of the lack of efficient selection criteria early in life. Estimated breeding values (EBVs) including information on own performance and on progeny performance are generally not available until a horse is 8 to 12 years old (German Equestrian Federation, 2008).

Genomic selection (GS) has the potential to substantially improve existing breeding strategies. The notion of GS was formulated by Meuwissen *et al*. (2001) and is being implemented in dairy cattle breeding programs (Hayes *et al*., 2009). The benefit of GS to conventional breeding programs has been demonstrated for dairy cattle (Schaeffer, 2006; König *et al*., 2009) and for pigs (Simianer, 2009). A substantial increase in genetic gain was found for breeding programs characterized by long generation intervals, and those focusing on lowly heritable, functional traits (König *et al*., 2009). Both findings support the demand to evaluate the potential of GS for horse breeding programs.

The aims of our study were to: (i) evaluate the impact of genomic breeding values (GEBVs) on the accuracy of EBVs and on the relative selection response by applying selection

† E-mail: ahaberl@gwdg.de

index theory; (ii) develop a strategy to address the practical problem of how to combine different types of EBVs (e.g. GEBV and EBVs available from different countries) in an overall breeding goal.

Material and methods

The methodology of combining phenotypic observations (y) and the single nucleotide polymorphism (SNP) genotype as a marker trait (m) via selection index theory was developed by Dekkers (2007). Application of this methodology was put into practice by König and Swalve (2009) to evaluate genomic breeding strategies in dairy cattle breeding programs. The present study, this method was extended to specific scenarios relevant for selection decisions in horse breeding programs. Evaluation criteria were the correlation between aggregate genotype and selection index, referred to as accuracy of the EBV (r_{TI}), as well as the relative selection response (RSR), which was calculated applying the formula

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

with ΔG being the selection response per generation. In order to assess the impact of GS on practical situations, all scenarios were investigated for a lowly heritable, functional trait and a trait of high heritability. As an example for a lowly heritable trait, susceptibility to *osteochondrosis* (OC), with a heritability of 0.15 (Pieramati *et al.*, 2003; Schober, 2003), was chosen. OC can cause disorders of chondral growth. When diagnosed, it reduces the horse's sales value considerably (Van Hoogmoed *et al.*, 2003; Stock and Distl, 2007). The quality of trot ($h^2 = 0.52$; Jaitner and Reinhardt, 2008) is highly correlated to the other gaits and to the rideability (Schade, 1996; Thorén Hellsten *et al.*, 2006), and therefore represents an important high heritability trait not only for dressage horses.

Scenario I: Genotyped young horse without own performance

Scenario I was designed considering animals without own performance according to the first step of selection in the breeding scheme of the German Riding Horse (Figure 1). At the age of 6 months, foals are inspected by the breeding associations. Many breeders decide whether or not to castrate the young stallion based on these first results, even though the correlations between results of foal inspection and subsequent studbook inspections are low (Schorm, 1983). A high proportion of male foals is castrated at a very young age, resulting in low selection intensities in subsequent steps of selection (Philipsson *et al.*, 1990; von Lengerken and Schwark, 2002). Scenario I is also valid to achieve improvements in selection of young mares without own performance, for example, to select mares as potential donors for embryo transfer (ET). Hence, we constructed a scenario for the genomic era where the SNP genotype of the foal, as well as the performances of the dam and of the sire, was used as information sources in the index.

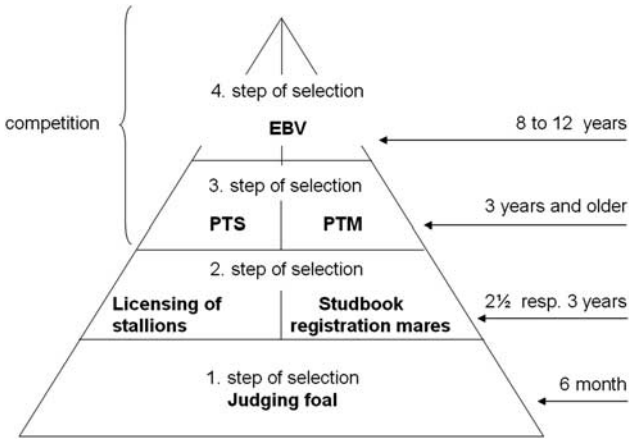

Figure 1 Pyramid structure of the breeding program of the German Riding Horse. EBV = estimated breeding value; PTS = performance test stallions; PTM = performance test mares.

The (co)variance matrix **P** of index sources was

$$\mathbf{P} = \begin{bmatrix} \sigma_m^2 & 0.5\sigma_{am} & 0.5\sigma_{am} \\ 0.5\sigma_{am} & \sigma_y^2 & 0 \\ 0.5\sigma_{am} & 0 & \sigma_y^2 \end{bmatrix}.$$

The first line of **P** refers to the marker genotype of the foal, the second line refers to the phenotypic performance of the sire and the third line corresponds to the phenotypic performance of the dam. According to Lynch and Walsh (1998), the heritability of the SNP genotype was fixed to a value of 1, which entails identical values for both phenotypic and genetic variance of the marker trait (σ_m^2). Variances for m were calculated using equation (1):

$$\sigma_m^2 = r_{mg}^2 \times \sigma_a^2, \quad (1)$$

where r_{mg} denotes the correlation between the breeding value estimated from the SNP genotype and the true breeding value (= accuracy of GEBV), and σ_a^2 is the additive-genetic variance of the trait. σ_y^2 is the phenotypic variance of the trait. The covariance σ_{am} between marker genotype m and phenotype y is described by the general equation (2):

$$\sigma_{am} = a_{ij} \times r_{mg}^2 \times \sigma_a^2, \quad (2)$$

with a_{ij} being the coefficient of relationship between animal i used in the index and animal j in the aggregate genotype. For this scenario, the coefficient of relationship between foal i and its dam and sire j was 0.5.

Covariances between traits in the index and traits in the breeding goals were included in matrix **G**, which was defined as

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

The first line of **G** refers to the marker genotype of the foal, the second line refers to phenotypic performance of the sire and the third line represents the phenotypic performance of the dam. The columns correspond to the GEBV estimated from the marker genotype *m* and to the conventional breeding value of the phenotypic trait *y*. Matrix **C** was the matrix for variances and covariances of breeding values, that is,

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

As the SNP genotype was considered as an auxiliary trait, the economic weight was put on the phenotypic performance, which resulted in vector **w** being

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

The variance of the aggregate genotype (**T**) was $\sigma_T^2 = w'\mathbf{C}w$, and the variance of the index (**I**) was $\sigma_I^2 = b'\mathbf{G}w$. These variances are essential to calculate $r_{TI} = \sigma_I/\sigma_T$.

Scenario II: Genotyped horse with own performance

In scenario II, the selection index was extended for a genotyped horse with own performance, but without progeny information. Such a scenario corresponds to step 3 of selection (Figure 1), that is, to young stallions or mares at the age of 4 to 7 years, which have accomplished performance testing. Information for the dam and for the sire was considered as done in scenario I. Matrix **P** of phenotypic (co)variances was

$$\mathbf{P} = \begin{bmatrix} \sigma_m^2 & \sigma_{am} & 0.5\sigma_{am} & 0.5\sigma_{am} \\ \sigma_{am} & \sigma_y^2 & 0.5\sigma_a^2 & 0.5\sigma_a^2 \\ 0.5\sigma_{am} & 0.5\sigma_a^2 & \sigma_y^2 & 0 \\ 0.5\sigma_{am} & 0.5\sigma_a^2 & 0 & \sigma_y^2 \end{bmatrix}$$

and matrix **G** was written as

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

Equations (1) and (2) were applied to calculate variance and covariance components for matrices **P** and **G**. The vector **w** of economic weights and the (co)variance matrix **C** for breeding values were identical to **w** and **C** in scenario I.

Scenario III: Genotyped horse with own performance and performance of progeny

The fourth step of selection in a conventional breeding program is the estimation of breeding values (Figure 1). In order to estimate highly reliable EBVs, progeny records are needed. Therefore, this scenario corresponds to stallions at the age of 8 to 12 years. As a consequence, five index sources were considered in this scenario: records of a varying number of progeny (5, 50 or 100, respectively), the SNP genotype of the stallion, the own performance of the stallion and again the performance of the dam and the sire. Hence, matrix **P** was

$$\mathbf{P} = \begin{bmatrix} \frac{(1+(n-1)0.25h^2)}{n}\sigma_y^2 & 0.5\sigma_{am} & 0.5\sigma_a^2 & 0.25\sigma_a^2 & 0.25\sigma_a^2 \\ 0.5\sigma_{am} & \sigma_m^2 & \sigma_{am} & 0.5\sigma_{am} & 0.5\sigma_{am} \\ 0.5\sigma_a^2 & \sigma_{am} & \sigma_y^2 & 0.5\sigma_a^2 & 0.5\sigma_a^2 \\ 0.25\sigma_a^2 & 0.5\sigma_{am} & 0.5\sigma_a^2 & \sigma_y^2 & 0 \\ 0.25\sigma_a^2 & 0.5\sigma_{am} & 0.5\sigma_a^2 & 0 & \sigma_y^2 \end{bmatrix}$$

and matrix **G** was

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

Again, the vector **w** of economic weights and the (co)variance matrix **C** for breeding values were identical to **w** and **C** in scenario I.

Scenario IV: Combination of breeding values into a combined index

Scenarios I to III were constructed to evaluate the potential of additional SNP information in terms of r_{TI} or relative selection response. The following approach addresses the question how to combine different EBVs with different accuracies and different correlations among each other into a combined index, which is constructed to match an overall breeding goal. This would be of use, for example, in case EBVs from different countries were available for a stallion. Another application would include the merging of conventional EBVs with GEBVs or joining EBVs measured at different stages of the horse's life, as was assumed in this scenario. Methodology is also based on selection index calculations, but using EBVs rather than phenotypic observations.

In this scenario, a combined index (*T*), illustrating an overall breeding goal, is composed of three different EBVs (TBVs), which are considered as traits of *T*. Hence, the overall index of an animal *k* was

$$T_k = \sum_{i=1}^{i=n} b_i TBV_i.$$

According to selection index theory, *b*-values were calculated as

$$b = \mathbf{P}^{-1}\mathbf{G}\mathrm{w}$$

with matrices **P** and **G** as explained below.

For each TBV, a separate type of EBV is available as information source: (1) an integrated breeding value (IEBV)

incorporating all available information on relatives of a stallion, as well as progeny information, which commonly has a high accuracy (r_{TI}) and in this example is arbitrarily set to 0.85 for all runs. Because progeny information is considered, IEBVs are only available later in life; (2) an *EBV* including the result of the stallion's performance test as well as the results of its performance tested male relatives (stallion estimated breeding value, SEBV). Due to less information and estimation earlier in life, the accuracy of *SEBV*s is generally lower than for *IEBV*s, and therefore we have chosen the values of 0.5 and 0.8, respectively; and (3) a *GEBV* with accuracy varying from 0.1 to 0.9. The correlation between *IEBV* and *SEBV* was set to 0.95, and for all *TBV*s equal economic weights per genetic standard deviation were assumed. In a second run, the correlation between *IEBV* and *SEBV* was reduced to 0.5. Notations for matrices were chosen in analogy to index calculations used above. The standard deviation of the *EBV* for a trait *i* was

$$\sigma_{EBV_i} = r_{EBV_i:TBV_i} \times \sigma_{TBV_i},$$

where σ_{TBV_i} denotes the standard deviation of the breeding value TBV_i, which was used in the overall index and $r_{EBV_i:TBV_i}$ is the correlation between EBV_i and TBV_i, or in other words the accuracy of the EBV_i (r_{TI_i}). The correlation between an EBV_i and an EBV_j is

$$r_{EBV_i:EBV_j} = r_{EBV_j:TBV_j} \times r_{EBV_i:TBV_i} \times r_{TBV_i:TBV_j}.$$

This information is needed to compute matrix **P**, the variance–covariance matrix for *n* EBVs:

$$\mathbf{P} = \begin{bmatrix} \sigma^2_{EBV_1} & \sigma_{EBV_1:EBV_2} & \cdots & \sigma_{EBV_1:EBV_n} \\ & \sigma^2_{EBV_2} & \cdots & \sigma_{EBV_2:EBV_n} \\ & & \cdots & \vdots \\ sym. & & & \sigma^2_{EBV_n} \end{bmatrix}.$$

Covariances between an EBV_i and an EBV_j were calculated using the following formula:

$$\sigma_{EBV_i:EBV_j} = r_{EBV_i:EBV_j} \times \sigma_{EBV_i} \times \sigma_{EBV_j}.$$

Matrix **C**, the quadratic variance–covariance matrix for *m* TBVs in the overall index, was

$$\mathbf{C} = \begin{bmatrix} \sigma^2_{TBV_1} & \sigma_{TBV_1:TBV_2} & \cdots & \sigma_{TBV_1:TBV_m} \\ & \sigma^2_{TBV_2} & \cdots & \sigma_{TBV_2TBV_m} \\ & & \cdots & \vdots \\ sym. & & & \sigma^2_{TBV_m} \end{bmatrix}.$$

Covariances between TBV_i and TBV_j were calculated by using the following formula:

$$\sigma_{TBV_i:TBV_j} = r_{TBV_i:TBV_j} \times \sigma_{TBV_i} \times \sigma_{TBV_j}.$$

Matrix **G** of dimension $n \times m$ is the covariance matrix between the *n* EBVs used in the index and the *m* TBVs used in the aggregate genotype (= breeding goal):

$$\mathbf{G} = \begin{bmatrix} \sigma_{EBV_1:TBV_1} & \sigma_{EBV_1:TBV_2} & \cdots & \sigma_{EBV_1:TBV_m} \\ \vdots & \sigma_{EBV_2:TBV_2} & \cdots & \sigma_{EBV_2:TBV_m} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{EBV_n:TBV_1} & \vdots & \cdots & \sigma_{EBV_nTBV_m} \end{bmatrix}.$$

Those covariances in **G** were calculated by using the formula:

$$\sigma_{EBV_j:TBV_i} = r_{EBV_j:TBV_j} \times r_{TBV_i:TBV_j} \times \sigma_{EBV_j} \times \sigma_{TBV_i}.$$

Hence, the individual weighting factor *b* of an EBV_i for an animal in the overall breeding goal depends on the accuracy of this EBV. Correlations among TBVs in the breeding goal, as well as economic weights *w*, are equal for all groups of animals.

The accuracy of T_k was

$$r_{T\hat{T}} = \sqrt{\frac{b'\mathbf{P}b}{w'\mathbf{C}w}}.$$

Results and discussion

Estimation of SNP effects within a calibration group, and transferring those estimates to animals in the whole population, is the key feature of GS methodology. The availability of 50-K SNP chip technology, as well as the recent release of the horse genome sequence (Wade *et al.*, 2009), provides the framework to estimate highly reliable GEBVs analogous to dairy cattle (e.g. VanRaden *et al.*, 2009). The extent of linkage disequilibrium (LD) was analyzed by Corbin *et al.* (2010) for the Thoroughbred horse. Owing to the high LD, the authors concluded that GS could be applied in the observed population. However, strength of LD as a function of the effective population size (N_e) may be lower in European sport horse breeds compared with the Thoroughbreds ($N_e = 100$). For example, in the Hanoverian Warmblood, N_e was estimated to be 372 (Hamann and Distl, 2008). Wade *et al.* (2009) found unusually high LD in Thoroughbreds compared with other horse breeds. Nevertheless, GS should be feasible, because strength of LD across several horse breeds (Wade *et al.*, 2009) is comparable to LD in Holstein cattle (Qanbari *et al.*, 2010), where GS was implemented successfully. First practical investigations in terms of estimation of GEBVs in horses are carried out for Franches-Montagnes horses in Switzerland (Hasler *et al.*, 2011).

Scenario I: Genotyped young horse without own performance

Owing to insufficient sources of information at this stage of selection, accuracies of breeding values estimated for young horses are particularly low, especially for low heritability traits. Selection of foals at this early point in time reduces

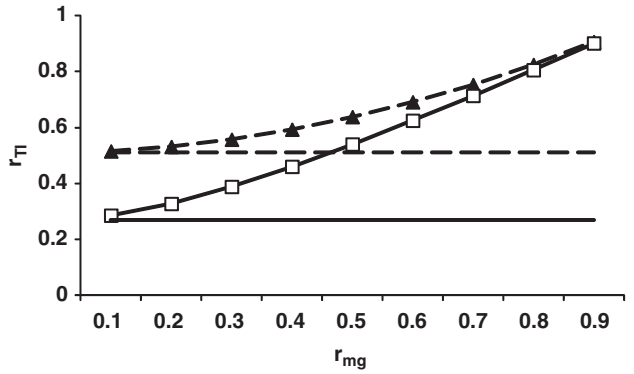

Figure 2 Correlation between index and aggregate genotype (r_{TI}) in dependency of accuracy of GEBV (r_{mg}) for a genotyped horse without own performance (Scenario I). Dashed line with black triangles: trot ($h^2 = 0.52$); solid line with white squares: OC ($h^2 = 0.15$). Parallel lines to *x*-axis: conventional accuracies not including GEBVs in the index, that is, dashed line for trot and solid line for OC. GEBV = genomic estimated breeding value; OC = *osteochondrosis*.

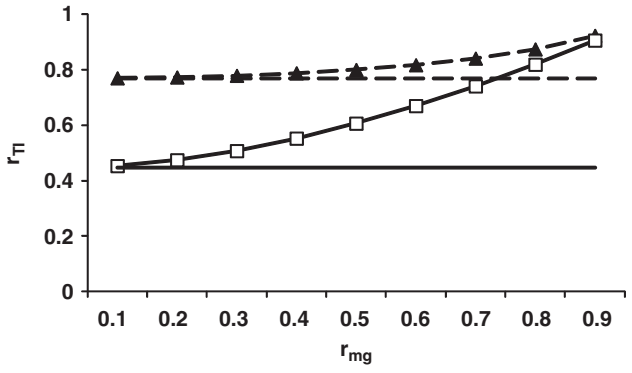

Figure 3 Correlation between index and aggregate genotype (r_{TI}) in dependency of accuracy of GEBV (r_{mg}) for a genotyped horse with own performance and performance of parents (Scenario II). Dashed line with black triangles: trot ($h^2 = 0.52$); solid line with white squares: OC ($h^2 = 0.15$). Parallel lines to *x*-axis: conventional accuracies not including GEBVs in the index, that is, dashed line for trot and solid line for OC. GEBV = genomic estimated breeding value; OC = *osteochondrosis*.

generation intervals, but is associated with a higher risk for practical breeders. Model calculations by Schade (1996) showed that genetic gain is reduced by 70% if stallions are used for matings before being performance tested. This is probably due to the fact that the phenotype itself, that is, riding quality, cannot be tested at this early point in time. Particularly with regard to castrating, there is a high risk of unfortunate selection decisions as long as there is no information on own performance available.

Only including phenotypic records from the sire and the dam of the foal in the index results in $r_{TI} = 0.27$ for OC (Figure 2). As known from selection index theory, additional information from further close relatives of the foal would increases r_{TI} only marginally. In contrast, a distinct gain of accuracies can be achieved when including the SNP genotype in addition to the sire's and dam's performance, even for low accuracies of GEBVs (r_{mg}) in combination with a low heritability (Figure 2). For $r_{mg} = 0.3$, the additional information of the SNP genotype increases r_{TI} to 0.39. Extremely high r_{mg} of 0.8, or even higher, enable similar r_{TI} for the low and the high heritability trait ($r_{TI} = 0.81$ to 0.91). However, when referring to other species, for example, dairy cattle, accuracies of GEBVs are substantially higher for production traits compared with fertility, somatic cell score or longevity (VanRaden *et al.*, 2009). Nevertheless, on the basis of results from simulation studies or deterministic predictions (Calus *et al.*, 2008; Daetwyler *et al.*, 2010), a correlation of $r_{mg} = 0.5$ should be feasible also for GS for functional traits in horses. Such a crucial value doubles r_{TI} at this very early point of selection (Figure 2) compared with the accuracy of the conventional index. When additionally considering economical aspects, even r_{mg} lower than 0.5 enable additional gain in terms of return of investment for pig breeding programs (Simianer, 2009), or in terms of breeding profit for dairy cattle breeding programs (König *et al.*, 2009). Thus, pre-selection of genotyped foals can be used for the identification of promising selection candidates very early and therefore helps to avoid improper castrating decisions. Until further testing, the issue of temporary breeding permissions may contribute to shorten generation intervals and is already practised by several organizations. Those breeding permissions are valid from the stallion's licensing carried out at the age of $2\frac{1}{2}$ years and allow a limited number of matings until the stallion is performance tested at the age of 3 or 4 years. However, the accuracy and effectiveness of this practice could be improved by considering GEBVs as additional information source.

According to Schaeffer (2006), more accurate breeding values on the dam side of selection can be achieved when genotyping females. This can be of economic importance when choosing young females without own performance as donors for ET, which still is an expensive biotechnology. New commercial reproductive technologies such as ET have been adopted by some horse breeding associations. Long *et al.* (2003) focused on two examples, the American Quarter Horse Association and the United States Polo Association. Advantages of ET will increase with decreasing generation intervals, provided that sufficiently reliable EBVs of young mares are available. For this specific case in horse breeding, the combination of both reproduction technologies and molecular genetic tools is a powerful approach to further increase selection response (e.g. Spelman and Garrick, 1998).

Scenario II: Genotyped horse with own performance

The higher the basic result for r_{TI} without considering genomic information, the lower the gain in r_{TI} when including additional SNP information in the index (comparison of Figures 2 and 3).

The conventional index sources own performance, performance of sire and performance of dam result in $r_{TI} = 0.45$ for $h^2 = 0.15$ (Figure 3). Including additionally a GEBV with $r_{mg} = 0.5$ in the index, r_{TI} increased by 40%. For $r_{mg} = 0.9$, relative selection response (not shown) is doubled to a value of 2.03. Hence, in scenario II, the benefit of GS in terms of gains in r_{TI} is substantial, in particular for the lowly heritable, functional trait.

Generally, r_{TI} cannot drop below r_{mg} (König and Swalve, 2009). As a practical consequence, provided that r_{mg} is 0.7 or higher, r_{TI} reaches at least the same level as can be obtained by running performance testing. For example, for $r_{mg} = 0.8$ and $h^2 = 0.15$, r_{TI} for the combination of the SNP genotype and own performance is 0.82. Formulas developed by Stricker and Fernando (2008) or by Daetwyler *et al.* (2008 and 2010) can be used to derive r_{mg} dependent on the number of genotyped animals.

However, reliable phenotypes are an essential pre-requisite to derive reliable SNP effects. The most effective source of data for genetic evaluation of young stallions and their parents is phenotyping carried out in the form of performance tests on station (Gerber Olsson *et al.*, 2000; Thorén Hellsten *et al.*, 2006). Continuous phenotyping in this format and frequent re-estimation of SNP effects within a calibration group are required because the accuracy of GEBVs is declining over generations as shown in simulation studies by Habier *et al.* (2007). This is due to a decreasing relationship between calibration group and selection candidates, as well as due to decay in LD between SNP-markers and surrounding quantitative trait loci caused by recombination events (Sonesson and Meuwissen, 2009). In addition, as priorities in breeding goals change or new assessment techniques become available, from time to time new phenotypes (e.g. König von Borstel *et al.*, 2011) may be introduced into the breeding program, requiring estimation and calibration of SNP effects for these new traits.

In order to keep generation intervals as short as possible, breeding organizations should encourage the use of young stallions. The accuracy of EBVs of performance tested stallions can still be enhanced by including genomic information.

Scenario III: Genotyped horse with own performance and performance of progeny

Generally, for the estimation of highly reliable EBVs, the availability of progeny records is of major importance. This implies that stallions are at the age of 8 years or even older once their EBVs reach high accuracies of 0.92 to 0.99 (German Equestrian Federation, 2008). Dubois and Ricard (2007) focused on the problems of long generation intervals due to extended progeny testing systems, and they encouraged breeders to use younger stallions with a reduced number of progeny as a compromise.

Additional gain in r_{TI} from GEBV is relatively low when performance of parents, own performance and progeny records are available as index sources. This finding is illustrated by the relative selection response (Figure 4). For the highly heritable trait and 50 or 100 progeny records, the value of RSR is, independent from r_{mg}, close to 1. This implies negligible gain when considering the GEBV as additional index information. For the lowly heritable trait and 50 or 100 progeny records, and for the highly heritable trait and five progeny records, *RSR* ranged from 1.06 to 1.13 for $r_{mg} = 0.9$. Substantial gain in r_{TI} in scenario III was found only for the lowly heritable trait and five offspring, leading to an *RSR* of 1.65.

Methodology developed for scenario III can additionally be used to derive the optimal number of progeny records to achieve a pre-defined r_{TI}. Additional progeny records contribute to realize a high r_{TI} for an index considering lowly heritable traits and genomic information with moderate r_{mg} in the range from 0.4 to 0.7.

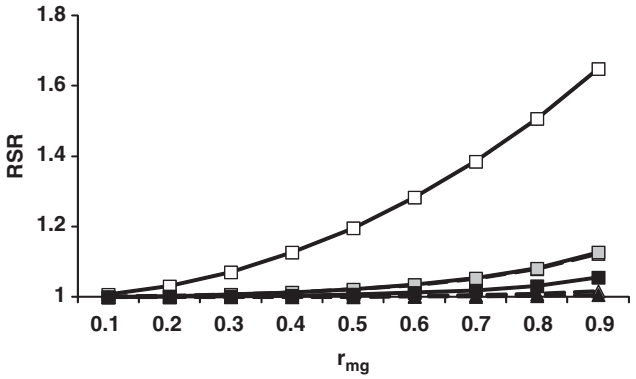

Figure 4 Relative selection response for a genotyped horse with own performance, performance of parents, and a different number of progeny in dependency of accuracy of GEBV (r_{mg}). Dashed lines with triangles: trot ($h^2 = 0.52$); solid lines with squares: OC ($h^2 = 0.15$). White triangles or squares: five progeny; gray triangles or squares: 50 progeny; black triangles or squares: 100 progeny. GEBV = genomic estimated breeding value; OC = *osteochondrosis*.

A crucial point for the practical implementation and ultimate success of GS will be the acceptance of GEBV by practical breeders, as well as the additional cost component for genotyping sport horses. Applied selection strategies in horse breeding programs traditionally have a strong focus on phenotypic performances rather than EBVs (Koenen *et al.*, 2004). Beyond dressage and show jumping, breeders have the opportunity to use GEBVs of health traits, for example, OC, for selection decisions. Van Hoogmoed *et al.* (2003) showed that radiographic findings of OC severely reduce the sales value of a horse. Some further studies, for example, Stock and Distl (2007) investigated the correlations between radiographic findings and performance traits in warmblood riding horses. They concluded that riding horse performance will likely benefit from the reduction of prevalence of radiographic findings. Hence, all available tools should be applied to reduce incidence of disorders.

Scenario IV: Combination of breeding values into a combined index

Application of scenarios I to III is appropriate for such situations where detailed information for selection index calculations are available, for example number of animals within selection groups, phenotypic parameters and genetic parameters. In general, availability of those parameters is guaranteed within an own breeding program or on the national scale. A major requirement for the practical implementation of GS in sport horse breeding programs is the setup of a calibration group for the estimation of SNP effects. For achieving adequate accuracies of the GEBV, it will be beneficial for breeding organizations to cooperate in this matter, like exemplified by the European breeding organizations of Holstein–Friesian cattle (Lund *et al.*, 2010).

A calibration group composed of animals being registered in different breeding organization would be justified by the extensive genetic exchange between breeding populations within Germany, as well as between European countries (Koenen *et al.*, 2004). Scenario IV could be applied for the combination of single breeding values from different countries in an index constructed to match an overall breeding goal. There is also the possibility to use stallions from, for example, Sweden or the Netherlands in German breeding programs, and those stallions have different sources of EBVs. Hence, addressing the question of an optimal combination of EBVs is important.

Results in terms of $r_{T\hat{T}}$ and weighting factors (*b*-values) for the combination of GEBV, IEBV and SEBV by altering r_{mg} are depicted in Figures 5 and 6, respectively. Independently of the accuracy for SEBV, that is, 0.8 *v.* 0.5, accuracy of the combined index substantially increases with increasing r_{mg}. For different accuracies of SEBV, the gap in accuracies for the combined index decreases with increasing r_{mg} (Figure 5). This is due to the impact of highly accurate GEBVs explaining most of the genetic variance of the aggregate genotype. Hence, for high accuracies of GEBVs, further correlated information sources only marginally improve the accuracy of the aggregate breeding value.

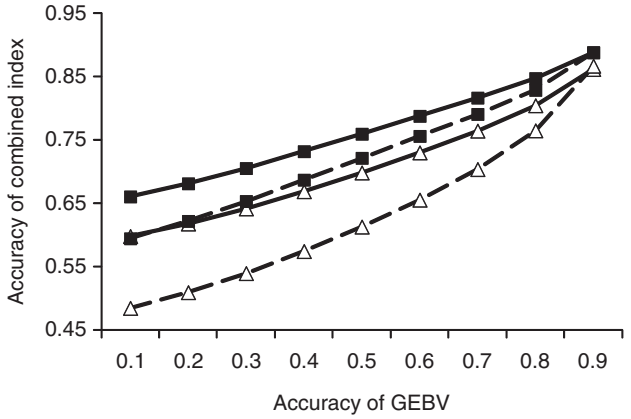

Figure 5 Accuracy of the combined index including GEBV, IEBV and SEBV in dependency of accuracies of GEBV (r_{mg}) for equal economic weights per genetic s.d. (solid lines: correlation between *SEBV* and *IEBV* = 0.95; dashed lines: correlation between *SEBV* and *IEBV* = 0.50; black squares: accuracy of *SEBV* = 0.8; white triangles: accuracy of *SEBV* = 0.5; accuracy of *IEBV* = 0.85 for all scenarios). GEBV = genomic estimated breeding value; IEBV = integrated estimated breeding value; SEBV = stallion estimated breeding value.

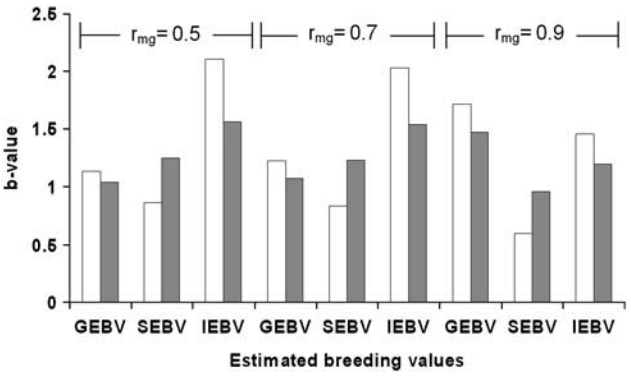

Figure 6 Weighting factors (*b*-values) for the combination of GEBV, IEBV and SEBV in a combined index in dependency of accuracies of GEBV (r_{mg}) for equal economic weights per genetic s.d. (white bars: accuracy of *SEBV* = 0.5; black bars: accuracy of *SEBV* = 0.8; accuracy of *IEBV* = 0.85 for all scenarios). The correlation between *IEBV* and *SEBV* was 0.95. GEBV = genomic estimated breeding value; IEBV = integrated estimated breeding value; SEBV = stallion estimated breeding value.

However, in reality accuracies of GEBVs will not be high enough to justify the complete abolishment of own-performance testing within horse breeding programs, such as the performance test for stallions. Moreover, performance testing provides phenotypic data that are of major importance for the re-estimation of marker effects in genomic breeding programs. For those reasons, the most likely breeding strategy remains a combination of both GS and performance testing. Considerable, earlier selection of male and female animals for breeding is possible because of the estimation of GEBVs for animals without phenotypic data.

Conclusions

On the basis of our results, application of GS can contribute to well-founded selection decisions within several selection stages of equine breeding programs particularly with regard to lowly heritable (e.g. functional) traits. For animals with a large number of progeny records available, additional gain in accuracy from GEBV is small. Accurate selection of genotyped young horses without own or progeny performance leads to a considerable reduction in generation intervals, and thereby increases the genetic response.

In order to apply GS in practice, breeding organizations will have to convince horse breeders that GS can be a valuable tool to increase selection response.

Acknowledgments

We would like to thank the German Ministry of Education and Research for the financial support of the program FUGATO + brain, which is also supported by the five horse breeding associations for Hanoverian, Trakehner, Oldenburger, Holsteiner and Westfalian horses.

References

Calus MP, Meuwissen TH, de Roos AP and Veerkamp RF 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178, 553–561.

Corbin LJ, Blott SC, Swinburne JE, Vaudin M, Bishop SC and Woolliams JA 2010. Linkage disequilibrium and historical effective population size in the Thoroughbred horse. Animal Genetics 41, 8–15.

Daetwyler HD, Villanueva B and Woolliams JA 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS One 3, e3395.

Daetwyler HD, Pong-Wong R, Villanueva B and Woolliams JA 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185, 1021–1031.

Dekkers JCM 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. Journal of Animal Breeding and Genetics 124, 331–341.

Dubois C and Ricard A 2007. Efficiency of past selection of the French Sport Horse: Selle Français breed and suggestions for the future. Livestock Science 112, 161–171.

Gerber Olsson E, Árnason Th, Näsholm A and Philipsson J 2000. Genetic parameters for traits at performance test of stallions and correlations with traits at progeny tests in Swedish Warmblood horses. Livestock Production Science 65, 81–89.

German Equestrian Federation (FN) 2008. Breeding program of the German Riding Horse. In Jahrbuch Sport und Zucht 2008, Deutsche Reiterliche Vereinigung, 84pp. FN-Verlag, Warendorf, Germany.

Habier D, Fernando RL and Dekkers JCM 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177, 2389–2397.

Hamann H and Distl O 2008. Genetic variability in Hanoverian Warmblood horses using pedigree analysis. Journal of Animal Science 86, 1503–1513.

Hasler H, Flury C, Menet S, Haase B, Leeb T, Simianer H, Poncet PA and Rieder S 2011. Genetic diversity in an indigenous horse breed – implications for mating strategies and the control of future inbreeding. Journal of Animal Breeding and Genetics 128, 394–406.

Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 92, 433–443.

Jaitner J and Reinhardt F 2008. Beschreibung Integrierte Zuchtwertschätzung Pferd. Retrieved February 24, 2011, from http://www.vit.de/index.php?id=zws-pferd.

Koenen EPC, Aldridge LI and Philipsson J 2004. An overview of breeding objectives for warmblood sport horses. Livestock Production Science 88, 77–84.

König S and Swalve HH 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. Journal of Dairy Science 92, 5292–5303.

König S, Simianer H and Willam A 2009. Economic evaluation of genomic breeding programs. Journal of Dairy Science 92, 382–391.

König von Borstel U, Euent S, Graf P, König S and Gauly M 2011. Equine behaviour and heart rate in temperament tests with or without rider or handler. Physiology and Behavior 104, 454–463.

Long CR, Walker SC, Wang RT and Westhusin ME 2003. New commercial opportunities for advanced reproductive technologies in horses, wildlife, and companion animals. Theriogenology 59, 139–149.

Lund MS, de Roos APW, de Vries AG, Druet T, Ducrocq V, Fritz S, Guillaume F, Guldbrandtsen B, Liu Z, Reents R, Schrooten C, Seefried FR and Su G 2010. Improving genomic prediction by EuroGenomics collaboration. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, 1–6 August, Leipzig, Germany, ISBN 978-3-00-031608-1.

Lynch M and Walsh B 1998. In genetics and analysis of quantitative traits. Sinauer Associates Inc., Sunderland, MA.

Meuwissen THE, Hayes BJ and Goddard ME 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

Niemann B 2009. Untersuchungen zu Veränderungen im Zuchtgeschehen und deren Auswirkungen auf die Hannoveraner Pferdezucht. PhD, Goettingen University.

Philipsson J, Árnason Th and Bergsten K 1990. Alternative selection strategies for performance of the Swedish Warmblood horse. Livestock Production Science 24, 273–285.

Pieramati C, Pepe M, Silvestrelli M and Bolla A 2003. Heritability estimation of *osteochondrosis dissecans* in Maremmano horses. Livestock Production Science 79, 249–255.

Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi AR and Simianer H 2010. The pattern of linkage disequilibrium in German Holstein cattle. Animal Genetics 41, 346–356.

Schade W 1996. Entwicklung eines Besamungszuchtprogrammes für die hannoversche Warmblutzucht. PhD, Goettingen University.

Schaeffer LR 2006. Strategy for applying genome-wide selection in dairy cattle. Journal of Animal Breeding and Genetics 123, 218–223.

Schober M 2003. Schätzung von genetischen Effekten beim Auftreten von *Osteochondrosis dissecans* beim Warmblutpferd. PhD, Goettingen University.

Schorm G 1983. Analyse der phänotypischen Entwicklung des Warmblutpferdes von der Geburt bis zum 3jährigen Pferd und Einflüsse von genetischen und umweltbedingten Faktoren. PhD, Leipzig University.

Simianer H 2009. The potential of genomic selection to improve litter size in pig breeding programs. In Proceedings of 60th Annual Meeting of the European Association for Animal Production, Barcelona, Spain, August 24–27, 2009. Wageningen Academic Publishers, the Netherlands.

Sonesson AK and Meuwissen THE 2009. Testing strategies for genomic selection in aquaculture breeding programs. Genetics Selection Evolution 41, 37.

Spelman RJ and Garrick DJ 1998. Genetic and economic responses for within-family marker-assisted selection in dairy cattle breeding schemes. Journal of Dairy Science 81, 2942–2950.

Stock KF and Distl O 2007. Genetic correlations between performance traits and radiographic findings in the limbs of German Warmblood riding horses. Journal of Animal Science 85, 31–41.

Stricker C and Fernando RL 2008. Genomewide genetic evaluation: how many individuals to genotype? International postgraduate course and workshop 'Whole Genome Association and Genomic Selection', September 1–8, Salzburg, Austria.

Thorén Hellsten E, Viklund Å, Koenen EPC, Ricard A, Bruns E and Philipsson J 2006. Review of genetic parameters estimated at stallion and young horse performance tests and their correlations with later results in dressage and show-jumping competition. Livestock Science 103, 1–12.

VanRaden PM, Van Tassel CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF and Schenkel FS 2009. Invited Review: reliability of genomic predictions for North American Holstein bulls. Journal of Dairy Science 92, 16–24.

Van Hoogmoed LM, Snyder JR, Thomas HL and Harmon FA 2003. Retrospective evaluation of equine prepurchase examinations performed. Equine Veterinary Journal 35, 375–381.

von Lengerken G and Schwark H-J 2002. Exterieur und Leistungen in der Pferdezucht – Alleskönner oder Spezialisten. Archiv für Tierzucht, Dummerstorf 45, 68–79.

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, Blöcker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MCT, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guérin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Røed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvänen AC, Tozaki T, Valberg SJ, Vaudin M, White JR and Zody MC 2009, Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander ES and Lindblad-Toh KGenome sequence, comparative analysis, and population genetics of the domestic horse. Science 326, 865–867.