

Application of selection index theory comprising genomic information to breeding programs of sport horses and pigs

Dissertation
for the Doctoral Degree
at the Faculty of Agricultural Sciences,
Department of Animal Sciences,
Georg-August-University Goettingen

presented by
Anne Marie Haberland
born in Berlin

Goettingen, May 2013

D7

1. Referee: Prof. Dr. Henner Simianer
Animal Breeding and Genetics Group
Department of Animal Sciences
Georg-August-University Goettingen

2. Referee: Prof. Dr. Sven König
Department of Animal Breeding
Faculty of Organic Agricultural Sciences
University of Kassel

Date of disputation: 17th of May 2013

meinem Opa zum 100. Geburtstag

Zusammenfassung	1
Summary	3
1st chapter	General Introduction 5
	Preface 6
	The classical breeding tools..... 6
	The selection index..... 6
	Economic modeling 8
	The gene flow method 10
	Genomic selection 10
	ZPlan+ 12
	Sport horse breeding in Germany 13
	Pig breeding in Germany 14
	Scope of this thesis..... 15
	References 17
2nd chapter	Integration of genomic information into Sport Horse breeding programs for optimization of accuracy of selection 23
3rd chapter	Interplay between heritability, genetic correlation and economic weighting in a selection index with and without genomic information 43
4th chapter	Efficiency of different selection strategies against boar taint in pigs 69
5th chapter	General Discussion 93
	Prospects of genomic selection for horse breeding 94
	Prospects of genomic selection for pig breeding 97
	References 100
Acknowledgment	103
Curriculum vitae	104

Die Anwendung des Selektionsindex unter Verwendung genomischer Information auf Zuchtprogramme für Reitpferde und Schweine

In der vorliegenden Arbeit werden drei Modellrechnungen vorgestellt, die auf dem Selektionsindex basieren und mit denen der Nutzen genomischer Zusatzinformation auf aktuelle Selektionsstrategien für Reitpferde und Schweine ermittelt werden soll. In Kapitel eins werden sowohl die klassischen Instrumente der Tierzucht, der Selektionsindex, die ökonomische Modellierung und die Genfluss-Methode, als auch die genomische Selektion dargestellt. Zudem wird die Software ZPlan+, in welcher diese Instrumente miteinander kombiniert sind, kurz vorgestellt.

Im zweiten Kapitel der Arbeit wird das Potenzial der genomischen Selektion für die Reitpferdezucht dargestellt. Der Zugewinn an Genauigkeit eines Zuchtwerts durch Hinzunahme genomischer Information wurde für drei praxisnahe Selektionsschritte betrachtet, nämlich für Pferde ohne Eigenleistung, Pferde mit Eigenleistung und Pferde mit Eigen- und Nachkommenleistungen. Die Modellrechnung zeigte, dass sich durch Hinzunahme genomischer Information besonders die Genauigkeit von Zuchtwerten junger Pferde ohne Eigen- oder Nachkommenleistung beträchtlich steigern ließe. Sobald Nachkommenleistungen unterstellt wurden war der Zugewinn an Genauigkeit durch genomische Information jedoch zu vernachlässigen. Für die praktische Pferdezucht ist die genomische Selektion daher vor allem für die Absicherung vorläufiger Zuchtzulassungen für junge Hengste nach der Körung und für leistungsgeprüfte Hengste ohne Nachkommenleistungen empfehlenswert.

Im dritten Kapitel der Arbeit wurde untersucht, wie sich die Art der berücksichtigten Informationsquellen (konventionell oder genomisch) auf die Interaktion der Parameter genetische Korrelation, Heritabilität und ökonomische Gewichtung auswirkte. Als Vergleichsparameter wurde die Standardabweichung des jeweiligen Index herangezogen, welche sich direkt proportional zum Zuchtfortschritt verhält. Es wurden drei Indices mit zwei Zuchtzielmerkmalen verglichen, deren Informationsquellen in Anlehnung an die Schweinezucht gewählt wurden. Der erste Index wurde für ein Tier mit Eigenleistung aufgestellt, während im zweiten Index zusätzliche Vollgeschwisterleistungen angenommen wurden. Im dritten Index wurde die Eigenleistung mit genomischer Information für beide Zuchtzielmerkmale kombiniert. Die Genauigkeit der genomischen Zuchtwerte wurde durch Annahme unterschiedlich großer Referenzpopulationen variiert. Je mehr Information im Index berücksichtigt wurde, desto unabhängiger wurde die Standardabweichung des Index von den Parametern genetische Korrelation, Heritabilität und ökonomische Gewichtung. Dabei wurden für den Vollgeschwisterindex mit 7 Vollgeschwistern und den genomischen Indexes mit einer Referenzpo-

pulation von 1.000 Tieren vergleichbare Ergebnisse gefunden. Die Anzahl von 1.000 Tieren in der Referenzpopulation kann somit als Mindestmaß für die Schweinezucht angenommen werden. Für geringere Heritabilitäten zeigte sich eine deutliche Überlegenheit des genomischen Indexes über den Vollgeschwisterindex.

Im vierten Kapitel der Arbeit wurde ein Schweinezuchtprogramm optimiert. Dazu wurden verschiedene Zuchtziele und Selektionsstrategien gegen Ebergeruch verglichen. Das Zuchtziel war entweder die Selektion gegen den Gehalt der chemischen Leitkomponenten des Ebergeruchs, Androstenon, Skatol und Indol oder die Selektion gegen den, von Testpersonen bestimmten, Human Nose Score. Innerhalb der Selektion gegen die chemischen Komponenten wurden drei verschiedene Informationsquellen miteinander verglichen, nämlich eine Feldprüfung in Form einer Biopsie am lebenden Eber, genomische Selektion sowie die Kombination beider Informationen. Innerhalb der Selektion gegen den Human Nose Score wurden als Informationsquellen eine Stationsprüfung von Voll- und Halbgeschwistern des Selektionskandidaten sowie genomische Selektion miteinander verglichen. Das komplexe Zuchtprogramm wurde deterministisch mit der Software ZPlan+ modelliert. Bei der Selektion gegen die chemischen Komponenten des Ebergeruchs war der Zuchtfortschritt am höchsten, wenn als Informationsquelle die Eigenleistung in Form einer Biopsie der männlichen Selektionskandidaten genutzt wurde. Aufgrund der hohen Erblichkeit der Ebergeruchskomponenten war der Nutzen genomischer Zuchtwerte deutlich geringer und gleichzeitig teurer. Für die Selektion gegen den geringer erblichen Human Nose Score lieferte die genomische Selektion einen höheren Zuchtfortschritt als die Stationsprüfung der Geschwister des Selektionskandidaten. Auch wenn der Human Nose Score als Zielmerkmal angesehen wurde, erwies sich eine Selektion gegen die chemischen Komponenten als zielführend, da der (korrelierte) naturale Zuchtfortschritt des Human Nose Scores bei Durchführung einer Biopsie deutlich höher war, als bei direkter Selektion gegen den Human Nose Score.

Im fünften Kapitel wird das Potential der genomischen Selektion für die Reitpferde- und Schweinezucht diskutiert. Dabei wird besonders auf Möglichkeiten für die Vergrößerung der Referenzpopulation und der Kostenreduktion eingegangen. Für die Reitpferdezucht werden zusätzlich Möglichkeiten aufgezeigt, wie der Zuchtfortschritt mit konventionellen Methoden gesteigert werden könnte.

Application of selection index theory comprising genomic information to breeding programs of sport horses and pigs

This thesis presents three model calculations based on selection index theory to evaluate the benefit of genomic information for the optimization of current breeding strategies in sport horses and pigs. In chapter one, an introduction is given to the classical breeding tools selection index theory, economic modeling and the gene-flow method, as well as to genomic selection. Additionally, the software ZPlan+, which combines all of these tools, is introduced.

In chapter two of the thesis, the potential of genomic selection is evaluated for sport horse breeding. The increase in accuracy of selection due to inclusion of genomic information was assessed for three practical selection steps: horses without own performance, horses with own performance and horses with own and progeny performance. The model calculation showed that by including genomic information a considerable increase in accuracy of selection could be achieved for young horses without own and progeny performance. For progeny tested horses, the gain in accuracy when including genomic information was negligible. For practical sport horse breeding, genomic selection is recommended for increasing the accuracy of selection of young stallions at the time of stallion licensing and of already performance tested stallions without progeny records.

Chapter three of the thesis represents an investigation on how the kind of information (conventional vs. genomic) influences the interaction of the parameters genetic correlation, heritability and economic weighting. Three different indices were compared in regard to their standard deviation, which behaves directly proportional to response to selection. The indices had a two-trait breeding goal and information sources were chosen according to pig breeding. The first index only incorporated an own performance, while the second index incorporated additional performances of full sibs. Within the third index, the conventional own performance was combined with genomic information on both breeding goal traits. The accuracy of the genomic breeding values was varied by varying the size of the reference population. The standard deviation of an index became more independent of genetic correlation, heritability and economic weighting, the more information was considered. Standard deviations of the index with 7 full sibs and the genomic index with 1'000 animals in the reference population were found to be comparable. Based on these results, 1'000 animals can be considered a minimal size for reference populations in pig breeding. If the heritability of one breeding goal trait was reduced, the standard deviation of the genomic index was higher than the standard deviation of the full sib index.

The fourth chapter of the thesis represents an optimization of a practical pig breeding program. The aim was to compare different breeding goals and selection strategies against

boar taint in a sire line. Breeding goal traits were either the chemical compounds of boar taint (androstenone, skatole and indole) or the so-called human nose score of boar taint as measured by human panelists. Within the selection against chemical compounds, three different information sources were compared: a biopsy-based field test conducted in live boars, genomic selection, and the combination of both. Within the selection against the human nose score, station testing of full and half sibs was compared to genomic selection. The complex breeding program was modeled deterministically using ZPlan+. Due to high heritabilities, the annual genetic gain achieved in the chemical compounds was highest when conducting biopsy-based performance testing of male selection candidates. Genetic gain arising from genomic selection was considerably smaller but breeding costs were higher. For selection against the lowly heritable human nose score, the potential of genomic information was higher than the potential of station testing of the selection candidate's sibs. Even if the human nose score was assumed to be the target trait, the (correlated) response in units of the trait was highest when selecting against the chemical compounds by biopsy-based field testing of the selection candidate.

Chapter five presents a discussion of the prospects of genomic selection for sport horse and pig breeding programs in regard to possibilities of increasing the size of the reference population and possibilities of reducing variable breeding costs. Additionally, different options for increasing the response to selection with conventional methods are suggested for sport horse breeding.

1ST CHAPTER

General Introduction

Preface

Animal breeding started with the domestication about 6'000 to 20'000 years ago most probably in the late Paleolithic (Old Stone Age) or early in the Neolithic (New Stone Age) (Lush, 1945). However, most of the present livestock breeds were created during the 19th century (Simon and Buchenauer, 1993), providing a structure to the forthcoming selection procedures. Animal breeding primarily relied on the phenotype of the individuals (Lush, 1945). Animals were selected on several traits from the beginning, but selection was not always efficient. In the middle of the 20th century, animal breeding shifted from an appearance orientation to a performance orientation whereas the further evolution to an economic orientation is still incomplete (Harris and Newman, 1994). Although animal breeding often does not rely on sophisticated economic models, the productivity of livestock systems has considerably increased within the last decades (Amer, 2011). In the following, the classical breeding tools as well as new methodologies introduced to animal breeding will be described.

The classical breeding tools

The selection index

Selection originally relied on production (milk, meat) but one can reasonably assume that the temperament of the animals and their fertility were also criteria of choice. A first step to optimize selection was thus to combine several traits to better select on all of them simultaneously.

Hazel and Lush (1942) compared three methods of selection and found that selection on an index, that is a linear combination of the breeding values of each of the traits, was more efficient than selection for independent culling levels or tandem selection for one trait at a time. Hazel (1943) addressed the problem of maximizing the economic response for a multiple-trait selection. His objective was to increase the genetic gain of several traits differing in economic importance, heritability, variance and degree of genetic and environmental correlation by combining them in an overall breeding goal. For that purpose, the aggregate genotype was defined as the sum of the breeding values (i.e. the additive genetic values) of all considered traits weighted by their relative economic importance:

$$T = a_1G_1 + a_2G_2 + \dots + a_nG_n,$$

where T is the aggregate breeding value, G_i are the breeding values of the traits considered in the breeding goal and a_i are the relative economic weights of the traits. Because true breeding values cannot be determined directly, selection had to be based on a correlated variable, the selection index (I):

$$I = b_1X_1 + b_2X_2 + \dots + b_nX_n,$$

where X_i are the performances of an animal for each trait and b_i are the regression coefficients, which are obtained from a multiple regression procedure to maximize the correlation between the index and the aggregate breeding value (r_{TI}) while minimizing the residual variance. Required parameters are the phenotypic standard deviation of each trait, phenotypic and genetic correlations between each pair of traits, phenotypic correlations between the traits of relatives and the heritability of each trait.

Selection index theory in this form had two main drawbacks. On one hand the impact of environmental effects that could not be corrected for, and on the other hand the varying amount of information between selection candidates. To better evaluate the genetic potential of an individual, Hazel (1943) recommended integrating performances of additional relatives into the index for instance because selection for certain traits cannot be accomplished on the breeding animal itself (e.g. carcass traits) and / or can only be measured in animals of one sex (e.g. prolificacy or milk production). Furthermore, he suggested a better control of the environment and correcting for known environmental effects. The author concluded that among the parameters affecting genetic gain, increasing the accuracy of the index, r_{TI} , offers the best prospects of turning selection more efficient.

The problem of environmental effects, which often inhibited a correct comparison of animal data was addressed by Henderson (1949). He estimated breeding values combining least squares with selection index methods but later on discovered that they were biased by deficiencies of repeatability. For that purpose, a maximum likelihood method was applied to calculate annual correction factors for different dairy cattle herds. These correction factors were used to compute the genetic improvement of the herd and to estimate breeding values for cows. A further modification was separating the application of the selection index into two steps: (1) estimation of breeding values for each breeding goal trait and (2) application of the relative economic weights (Henderson, 1951 cited by Hazel *et al.*, 1994). With this adjustment economic weighting factors could be varied according to the selection objectives or production systems without having to recalculate breeding values at the same time. Moreover, this modification enabled the use of estimators for each index trait and to account for unequal amounts of information (Hazel *et al.*, 1994). It therefore provided a solution to both problems of the selection index. A further development was the introduction of the so-called mixed-model equations by Henderson *et al.* (1959). They permitted simultaneous estimation of fixed effects and prediction of random genetic effects. Henderson (1973) proved the best linear unbiased predictor (BLUP) for random effects to be identical with the selection index criterion when using generalized least squares solutions as means for fixed effects.

In practical breeding, selection indices were not used until the 1970's. In 1971, the United States Department of Agriculture introduced its first economic index comprising milk yield traits (Norman and Dickinson, 1971). Still, selection index theory remains an often used tool for the evaluation of breeding programs. The main advantages are its applicability to a wide range of problems in animal breeding and its small computing time compared to sophisticated empirical approaches.

Economic modeling

After the introduction of selection index theory by Hazel (1943), research had to focus on the definition of breeding goals, since the aggregate genotype basically represents the economic breeding objective on which breeders have to focus to achieve maximum profit. Willham (1979) proposed a system for sire evaluation programs where the breeders were allowed to choose their breeding direction according to their own ideas. However he suggested that the breeding organizations should provide the appropriate tools for that purpose by collecting and evaluating performance data.

The first step towards an economic consideration of breeding programs was the determination of the relative economic importance of different traits in the aggregate breeding value. Costs and returns from a breeding program have to be investigated in order to determine the economic weights of single traits. Gjedrem (1972) recommended including all criteria which have a major impact on the efficiency of commercial production into the breeding objective. Costs to be considered are at least (1) feeding costs, (2) costs of labor and facility, (3) costs for the buildup and maintenance of breeding stock (Harris and Newman, 1994). Returns on the other hand are quantified through the value of products.

For the set-up of valuable economic weightings, Hazel (1943) proposed the marginal utility which is defined as the net increase in profit for one unit of genetic change in the selected trait independently of the effects from genetic changes in other traits. The economic values are hence considered as linear functions of the trait values, which is in practice often not true (Weller, 1994). Moav and Moav (1966) proposed the use of profit equations to compare the economic efficiency of lines or crosses, which provided a non-linear approach. Moav and Hill (1966) underlined the relationship between economics and genetics by computing economic values of traits as their partial differentials with respect to profit per unit of product. In order to overcome the issue of non-linearity, Harris (1970) discussed methods dealing with breeding objectives expressed by more complex functions. An approximation would be to divide a complex function into partial derivatives, which thus provides the rate of change of the profit function at the points of the population means. Melton *et al.* (1979) de-

veloped a method to estimate economic values by combining a profit function and a production function where every output is linked to the respective inputs.

Yet, the economic weights derived from profit equations depended on the perspective taken for the evaluation and hence varied if they were calculated regarding the consumer's interest, regarding the producer's interest or per unit of investment made (Moav, 1973; Brascamp *et al.*, 1985). The economic value of a trait thus differs depending on the perspective, e.g. between investors, farmers and consumers. Brascamp *et al.* (1985) showed that equal relative economic weights can be derived for all perspectives of evaluation if the derivation is based on the so-called 'normal profit'. This means considering profit as a cost of production and thus setting the profit equation to zero. Smith *et al.* (1986) discussed two conditions under which economic weights derived by different methods and considered from different perspectives are equivalent. The authors stated that profit from genetic change which is also achievable by changing the size of the production unit is not due to genetic improvement, and therefore should not be counted. Secondly, arguing that fixed costs depend on the level of output, the authors advised that 'fixed costs, like variable costs, should be expressed per unit of output'.

During the last decade, selection on functional traits like longevity and fertility moved into the focus of animal breeders. Fewson and Niebel (1986) concluded that the inclusion of functional traits into the breeding goal only marginally reduces genetic gain of production traits. Miesenberger *et al.* (1998) found that the overall profit of a breeding program can be even increased by including functional traits with their proper economic weighting. Assuming different market scenarios, Lind (2007) derived economic weights as well as optimal index weights for the aggregate breeding value for the selection of German and Austrian dairy- and dual purpose cattle. Wolfová *et al.* (2001) derived economic weights for crossbreeding systems in pigs. The position of the breed within the crossbreeding scheme was accounted for as a factor with an impact on economic weighting of the traits. The authors also found that traits should be weighted according to the kind of the breeding enterprise (multiplying of breeding animals versus fattening of end-products). Although numerous methods have been developed for designing economically precise breeding strategies, there are still shortcomings. Amer (2011) stated that a basic understanding of economic principles would prevent animal breeders from ignoring existing inefficiencies in making use of previous genetic improvements. To take full advantage of the potential of existing livestock breeding programs, the author proposed contemplating factors like vertical integration, value capture versus efficiency because of competition or intellectual property rights.

The gene flow method

The genetic progress resulting from selection in the breeding nuclei affects commercial production only after a time lag (Bichard, 1971). Costs and returns occur at different times and the realization of profit from one round of selection may be delayed according to the trait (Fewson, 1993). Especially in animals with long generation intervals, returns from a breeding program are realized long after the initial investment, which requires discounting of these parameters. The classical formula for annual improvement developed by Dickerson and Hazel (1944) and Rendel and Robertson (1950) only accounts for the fact that the same selection scheme was practiced for many generations. Hinks (1971) and Hill (1971) pointed out that genetic improvement resulting from one cycle of selection is not expressed constantly in successive years, but rather erratically approaches an equilibrium many years after a selection round.

Hill (1974) and Elsen and Mocquot (1974) concomitantly addressed the problem of how to predict rates of response to selection considering overlapping generations. The central point was how to specify the passage of genes between different age groups and sexes by using matrix notations.

The so-called P-matrix enables expressing the transmission of genes and the ageing of all selection groups. It is structured in four blocks which describe the four pathways of gene transmission being (1) males to breed males, (2) males to breed females, (3) females to breed males and (4) females to breed females (cp. Rendel and Robertson, 1950). Thus, the gene flow method offered a way to follow the genes deriving from a group of selected animals and to compute their contribution to subsequent generations. The genetic gain could be predicted as the summed products of the genetic selection differential and the proportions of genes deriving from sex-age groups. The gene-flow method permits to compute the time lag of improvement from nucleus to commercial stock and thus allows calculating discounted monetary returns from a breeding program.

Genomic selection

Since molecular technologies have been developed for determining the genotype of individuals at specific loci, attempts have been made to use this information for selection decisions. Fernando and Grossman (1989) proposed a method for combining information from a relevant locus with a polygenic term when predicting EBVs. This procedure was referred to as marker assisted selection (MAS) and the number of considered markers was limited to one or just a few, like e.g. microsatellites. With further development in genotyping technologies, large numbers of single nucleotide polymorphisms (SNPs) became commercially available.

The idea of using many markers at a time was first mentioned by Van Arendonk *et al.* (1989) and Lande and Thompson (1990) first proposed a method for combining traditional information sources with information on many markers. A simulation study conducted by Meuwissen *et al.* (2001) revealed the possibility of simultaneously using all markers for the prediction of breeding values, which was referred to as genomic selection (GS). The definition of GS incorporates the derivation of a prediction equation by means of a reference population of animals with accurately predicted conventional breeding values. This equation is then used for predicting the GBVs of the selection candidates. Further methodological developments were proposed, like computations using the genomic relationship matrix (VanRaden, 2009) or single step approaches combining phenotypic, genotypic and pedigree information at once (Misztal *et al.*, 2009). These were the basis for GS to become a routine procedure in many breeding programs, especially in dairy cattle.

A strategy for incorporating GS in a practical dairy cattle breeding scheme was first presented by Schaeffer (2006). In comparison with a conventional progeny testing scheme, the author found a considerable increase in accuracy of selection early in life as well as an enormous reduction of breeding costs. Although these first prognoses were very optimistic, the advantages of GS over traditional selection schemes were confirmed especially for the selection on traits with a low heritability (e.g. functional traits) or traits that cannot be measured on the selection candidate itself (Dekkers, 2004; König and Swalve, 2009, König *et al.*, 2009). Due to the great potential of shortening generation intervals, routine estimation of genomic breeding values (GBVs) was first implemented in cattle. The possibility to increase accuracy of selection early in life and thus shorten generation intervals was also found for sport horse breeding, as presented in chapter two of this thesis. In contrast, the main purpose of GS in pig breeding is enhancing the accuracy of selection. Model calculations showed the potential of GS with regard to fertility traits (Simianer, 2009; Cleveland *et al.*, 2010) and production traits (Haberland *et al.*, 2010).

The accuracy, with which a GBV can be estimated (r_{GBV}), is affected by different factors. Daetwyler *et al.* (2008; 2010) derived a formula

$$r_{GBV} = \sqrt{\frac{N_p r^2}{N_p r^2 + M_e}},$$

where N_p is the size of the reference population, r^2 is the reliability of the conventional EBVs of the animals in the reference population and M_e is the effective number of chromosome segments segregating in the population, which again is a function of the effective population size, the average length of a chromosome and the number of chromosome pairs.

In dairy cattle breeding, breeding organizations for Holstein Friesian have started to cooperate to increase the size of a common reference population and thus the accuracy of the estimated GBVs. For example, the reference population of the EuroGenomics project (Lund *et al.*, 2010) is composed of currently 25'000 progeny tested bulls. The situation is different in pig breeding, where breeding programs are organized more regionally and the genetic exchange between different populations, even within the same breed, is low. The reproduction rate of boars and therefore r^2 is much lower compared to bulls. The buildup of reference populations larger than several hundreds of animals thus requires either cooperation of different organizations similar to dairy cattle or including sows in the reference population.

Dekkers (2007) proposed a method for including marker information in selection index calculations. GBVs are considered as indicator traits with a heritability of 1, which are linked by genetic correlations to the breeding goal traits. This methodology was used in chapter three and four of this thesis and will be presented more in detail there.

ZPlan+, a software to optimize breeding programs in livestock

A preliminary concept for a software enabling the user to optimize breeding schemes with respect to investment parameters was developed by Niebel (1974). The program was used to perform model calculations on the optimization of purebreeding in dual purpose cattle (Niebel and Fewson, 1978) and of purebreeding in swine (Niebel and Fewson, 1979). A first version of ZPLAN combining the work of Niebel (1974) with the gene flow method (Hill, 1974, Elsen and Mocquot, 1974) was written by Karras (1984) and presented by Nitter and Graser (1994). The program is based on the classical tools selection index theory (Hazel, 1943) and gene flow method (Hill, 1974) and is a means to deterministically simulate breeding plans in any livestock species (Willam *et al.*, 2008). Additionally, it allows economic modeling of breeding programs and thus offers a good basis for the optimization of selection schemes in livestock breeding. The program was used by Wünsch *et al.* (1999; 2000) to evaluate the response to selection for a three-way crossbreeding system in pigs. Economic weights of production and reproduction traits were optimized according to sex and breed. Additionally, the optimal reproductive life time was determined using ZPLAN. Wolfová *et al.* (2001) investigated the impact of different crossbreeding systems on economic weights in purebred pig populations.

ZPLAN was originally written in FORTRAN and was rather demanding in its use. In the context of the project FUGATO+brain, the software was newly programmed incorporating a user-friendly interface. The main properties of the software remained in the new version, ZPlan+ (Täubert *et al.*, 2010). In a basic run, a breeding scheme is defined and evaluated.

Subsequently, parameters can be varied in order to compare different alternatives of the breeding program in terms of output parameters like genetic gain, breeding return, costs, profit, generation intervals, and accuracies of selection. With its deterministic approach, the advantage of the program compared to stochastic simulations is the feasibility of multi-trait modeling and the requirement of much less computing time (Willam *et al.*, 2008).

In addition to the basic applications of the previous version, ZPlan+ incorporates the possibility to straightforwardly include genomic information following the approach proposed by Dekkers (2007) with slight modifications. GBVs can be added as genomic indicator traits with a heritability of 1. The accuracy of the GBVs can be determined by defining the parameters N_p , r^2 and M_e , according to the formula proposed by Daetwyler *et al.* (2010). ZPlan+ thus allows for comparing conventional and genomic breeding strategies.

Sport horse breeding in Germany

In 1732, Frederick William I of Prussia founded the Prussian Stud Administration (Preußische Gestütsverwaltung) for providing the cavalry and agriculture with horses of good quality. Next to the foundation of state studs, a system of performance testing and selection of stallions was implemented (Graf, 2006). At the same time, facilities were constructed in East Prussia with the focus on improving the quality of the Trakehner horse, a light horse used by the cavalry. The availability of these institutions and genetic resources was later an important factor for Germany to become a leading horse breeding nation for the classical disciplines dressage, show jumping and 3-day eventing.

In the 1950's, the importance of horses as means of transport and agricultural production decreased due to the proceeding motorization and as a consequence, the number of horses in Germany declined dramatically. Horse breeding officials, e.g. Gustav Rau, made great efforts to promote the foundation of riding clubs and the organization of sport horse competitions (German Equestrian Federation, 2005). A major task for breeders at this time was changing the type of horses from the ancient working type to the now required modern type suitable for sport purposes. One advantage over other European countries was the already existing infrastructure of state controlled horse breeding institutions dating back to the 18th century. Additionally, the Trakehner horse proved to be an important genetic resource for transforming the old working type into a modern sport horse. The importance of breeding sport horses as partners for leisure activities and riding competitions grew in the beginning 1970's and its organization by breeding organizations and the German Equestrian Federation (FN) proceeded.

Nowadays, sport horse breeding programs are conducted by 25 German breeding organizations which are members of the FN. About 7'800 stallions and 89'000 mares are registered as breeding animals. The number of matings has been declining within the last years and about 39'200 foals were born in 2012 (German Equestrian Federation, 2013). Stallions and a small percentage of mares are performance tested in eleven test stations. Within the last decade, German sport horse breeding is in increasing competition with other horse breeding countries like the Netherlands, Belgium, Denmark or Sweden. Due to an increasing genetic exchange between breeding organizations within Germany but also between countries (Koenen *et al.*, 2004; Thorén Hellsten *et al.*, 2009), what matters is no longer the availability of good genetics, but the quality of the breeding program. In this regard, the diversity of breeding organizations in Germany is sometimes challenging. Because the breeding goal is largely the same for most of the breeding organizations, there is competition for breeders and the development of common strategies together with the FN is often time consuming.

In contrast to other livestock species, the breeding goal for the German sport horse includes important characteristics for different breeding directions, because several breeding directions even exist within one breed (e.g. dressage, show jumping, driving). There exists no precise determination of economic values for breeding goal traits (Koenen *et al.*, 2004; Haberland and Simianer, 2009). An important development for approaching this problem was to gradually separate the breeding stock according to the main breeding directions dressage and show jumping within the last 20 years (Schade, 1996; von Lengerken and Schwark, 2002, Niemann, 2009).

Pig breeding in Germany

Germany is the biggest producer of pork meat after China and the USA and the degree of self-sufficiency is around 100% (FAO, 2013). Within the last 20 years, the number of slaughtered pigs increased from 43'700 to 59'700 and an annual consumption of about 55 kg pork meat per person illustrates the significance of pig breeding in Germany (BMELV, 2013). The German pig breeding industry is structured into regionally operating breeding organizations, which conduct their own breeding programs. A classical crossbreeding program incorporates a three-way cross of F₁ sows (Landrace x Large White) with a sire line, mostly Piétrain. Other sire lines are e.g. Duroc or Hampshire (Willam and Simianer, 2011). In Switzerland, a Large White line especially selected for meat quality is used as sire line (chapter 4 of this thesis). Boars and gilts are usually selected according to their own-performance in a field test in combination with the performance data of their full and half sibs, as proposed by Niebel and Fewson (1979). The development of a 60K SNP array for *Sus scrofa* (Ramos *et al.*, 2009) and investigations concerning linkage disequilibrium carried out by Uimari and Tapio (2011)

and Badke *et al.* (2012) permit to assess genomic information as an additional information source. Cleveland *et al.* (2010) found the accuracies of GBVs for female fertility traits to be higher than accuracies of conventional EBVs that are normally available at the time of selection.

A currently much debated issue in pig breeding is finding possibilities for the reduction of boar taint, an unpleasant odor expressed by uncastrated males. Due to an increasing disapproval of consumers, the practice of surgical castration without using anesthesia, as performed in the past, is no longer an option (von Borell *et al.*, 2009). Next to immunocastration (Prunier *et al.*, 2006; Fàbrega *et al.*, 2010; Rydhmer *et al.*, 2010), which raises concerns of the retailers, fattening of intact boars is proposed as an option. However, reducing the occurrence of boar taint in the meat of the end-product is a prerequisite. Selection against the main components of boar taint seems promising due to high heritabilities (Sellier *et al.*, 2000; Windig *et al.*, 2012) and the availability of methods to assess their amount in carcasses or live boars (e.g. Baes *et al.*, 2012).

Scope of this thesis

This thesis aims at assessing the prospects of GS for different applications in horse and pig breeding programs. Selection index theory is applied to compare currently practiced breeding schemes to breeding schemes incorporating genomic information, similar to the investigations carried out by König and Swalve (2009) for dairy cattle. Chapter one gives an introduction to classical and new breeding tools, which were used for the analyses.

In chapter two, the additional gain in accuracy of selection when including genomic information is investigated for different selection steps of sport horse breeding. The focus is especially on possibilities to increasing the accuracy of selection for young horses with no riding performance available, as e.g. young stallions in the time of castrating decisions or stallions licensing.

In chapter three, the interplay between different factors, which have an impact on the response to selection, is evaluated. Those are the heritability, the genetic correlation and the relative economic weighting of traits. The aim was to assess, whether there are differences in the interplay of these factors, which depend on the kind of information (conventional vs. genomic). For this purpose, conventional and genomic indices with information sources chosen according to pig breeding were compared in regard of the standard deviation of the index. The standard deviation of the index is the product of accuracy and the standard deviation of the aggregate breeding goal, and therefore directly proportional to response to selection. Additionally, the required size of a reference population for GS in pig breeding is assessed in this context.

In chapter four, the prospects of GS in pig breeding are evaluated in regard to the selection against boar taint. A complex breeding program of a Swiss sire line was modeled in ZPlan+, which is evaluated as a means to optimize breeding schemes with respect to the potential benefit of genomic selection. The efficiency of different information sources is compared for the selection against the chemical compounds of boar taint, androstenone, skatole and indole, as well as for the selection against the human nose score of boar taint, as subjectively assessed by test persons. The different breeding strategies are compared in terms of genetic gain and variable breeding costs per selection candidate.

A general discussion on the potential of genomic selection for sport horse and pig breeding programs is presented in chapter five. The main focus is on possibilities to increase the size of the reference population and to reduce breeding costs, as well as on options to optimize response to selection by conventional selection tools, especially in sport horse breeding.

References

- Amer P 2011. The economics of successes and failures in animal breeding. *J Anim Breed Genet* 128, 327-328.
- Badke YM, Bates RO, Ernst CW, Schwab C and Steibel JP 2012. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* 13, 24.
- Baes C, Mattei S, Luther H, Ampuero Kragten S, Sidler X, Bee G, Spring P and Hofer A 2012. A performance test for boar taint compounds in live boars. Published online in *Animal* (doi:10.1017/S1751731112002273).
- Bichard M 1971. Dissemination of genetic improvement through a livestock industry. *Anim Prod* 13, 401-411.
- BMELV (Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz) 2013. <http://berichte.bmelv-statistik.de/SJT-4010500-0000.pdf> (accessed on 5th of March 2013).
- Brascamp EW, Smith C and Guy DR 1985. Derivation of economic weights from profit equations. *Anim Prod* 40, 175-179.
- Cleveland MA, Forni S, Garrick DJ and Deeb N 2010. Prediction of Genomic Breeding Values in a Commercial Pig Population. In: *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, Germany: August 2010 47.
- Dekkers JCM 2004. Commercial application of marker- and gene- assisted selection in livestock: Strategies and lessons. *J Anim Sci* 82, E313-328.
- 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.
- Dickerson GE and Hazel LN 1944. The effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J Agric Res* 69, 459-476.
- Elsen JM and Mocquot JC 1974. Méthode de prévision de l'évolution du niveau génétique d'une population soumise à une opération de sélection et dont les générations se chevauchent. *Bull. tech. Dépt. Génét. Anim.* 17, INRA, 30-54.
- Fàbrega E, Velarde A, Cros J, Gispert M, Suárez P, Tibau J and Soler J 2010. Effect of vaccination against gonadotrophin-releasing hormone, using Improvac[®], on growth performance, body composition, behaviour and acute phase proteins. *Livest Sci* 132, 53-59.
- FAO (Food and Agriculture Organization of the United Nations) 2013. Statistics <http://www.fao.org/corp/statistics/en/> (accessed on 6th of march 2013).
- Fernando RL and Grossman M 1989. Marker assisted selection using best linear unbiased prediction. *Genet Sel Evol* 21, 467-477.

- Fewson D 1993. Definition of breeding objective. In: Design of livestock breeding programs. The University of New England, Armidale, New South Wales, Australia.
- Fewson D and Niebel E 1986. Berücksichtigung indirekter Merkmale in Zuchtplänen für Zweinutzungsrinder. Züchtungskunde 58, 4-20.
- German Equestrian Federation 2005. 100 Jahre Pferdezucht und Pferdesport in Deutschland. FN-Verlag, Warendorf, Germany.
- German Equestrian Federation 2013. Jahresbericht 2012. FN-Verlag, Warendorf, Germany.
- Gjedrem T 1972. A study on the definition of the aggregate genotype in a selection index. Acta Agr Scand. 22, 11-16.
- Graf MW 2006. Die Königlich-Preußische Gestütsverwaltung. Thüringer Druckhaus, Eisenach/Mühlhausen, Germany.
- Haberland AM and Simianer H 2009. Ökonomische Gewichte in der Pferdezucht. In: Göttinger Pferdetage 2009, pp 33-40, FN-Verlag, Warendorf, Germany.
- Haberland AM, Ytournel F, Luther H and Simianer H 2010. Evaluation of selection strategies including genomic breeding values in pigs. In: Proc. of 61th Annu. Mtg. Eur. Assoc. Anim. Prod., Heraklion, Greece, August 23-27, 2010. Wageningen Academic Publishers, The Netherlands.
- Harris DL 1970. Breeding for efficiency in livestock production: defining the economic objectives. J Anim Sci 30, 860-865.
- Harris DL and Newman S 1994. Breeding for profit: synergism between genetic improvement and livestock production (a review). J Anim Sci 72, 2178-2200.
- Hazel LN and Lush JL 1942. The efficiency of three methods of selection. J Hered 33, 393-399.
- Hazel LN 1943. The genetic basis for constructing selection indexes. Genetics 28, 476-490.
- Hazel LN, Dickerson GE and Freeman AE 1994. The selection index – then, now, and in the future. Symp. Selection index theory. J Dairy Sci 77, 3236-3251.
- Henderson CR 1949. Estimation of changes in herd environment. J Dairy Sci 32, 706.
- Henderson CR 1952. Specific and general combining ability. Heterosis. J. W. Gowen, ed. Iowa State College Press, Ames.
- Henderson CR 1973. Sire evaluation and genetic trends. Pages 10-41 in Proc. Anim. Breeding Genet. Symp. in honor of Dr. J. L. Lush. Am. Soc. Anim. Sci., Am. Dairy Sci. Assoc., Champaign, IL.
- Henderson CR, Kempthorne O, Searle SR and von Krosigk CM 1959. The estimation of environmental and genetic trends from records subject to culling. Biometrics 15, 192.
- Hill WG 1971. Investment appraisal for national breeding programmes. Anim Prod 13, 37-50.
- Hill WG 1974. Prediction and evaluation of response to selection with overlapping generations. Anim Prod 18, 117-139.

- Hinks CJM 1971. The genetic and financial consequences of selection amongst dairy bulls in artificial insemination. *Anim Prod* 13, 209-218.
- Karras K 1984. ZPLAN. EDV-Programm zur Optimierung der Zuchtplanung bei landwirtschaftlichen Nutztieren. Institut für Tierhaltung und Tierzucht, Universität Hohenheim.
- Koenen EPC, Aldridge LI and Philipsson J 2004. An overview of breeding objectives for Warmblood sport horses. *Livest Prod Sci* 88, 77-84.
- König S, Simianer H and Willam A 2009. Economic evaluation of genomic breeding programs. *Journal of Dairy Science* 92, 382-391.
- 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.
- Lande R and Thompson R 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743-756.
- Lind B 2007. Ableitung der Wirtschaftlichkeitskoeffizienten und optimalen Indexgewichte des Gesamtzuchtwertes für die deutschen Milch- und Zweinutzungsrasen unter Berücksichtigung aktueller und erwarteter zukünftiger Rahmenbedingungen. PhD thesis, Georg-August University Goettingen, Germany.
- Lund MS, de Roos APW, de Vries AG, Druet T, Ducrocq V, Fritz S, Guillaume F, Guldbandsen B, Liu Z, Reents R, Schrooten C, Seefried M 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 2010, Leipzig, Germany.
- Lush JL 1945. *Animal breeding plans*. Iowa State College Press. Ames, Iowa, U.S.A.
- Melton BE, Heady EO and Willham RL 1979. Estimation of economic values for selection indices. *Anim Prod* 28, 279-286.
- Meuwissen THE, Hayes BJ and Goddard ME 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Miesenberger J, Sölkner J and Essl A 1998. Economic weights for fertility and reproduction traits relative to other traits and effects of including functional traits into a total merit index. *Proc. Intern. Workshop on Genetic Improvement of Functional Traits in Cattle*, Grub, Germany, Nov. 1997. *Interbull-Bulletin* 18, 78-84.
- Misztal I, Legarra A and Aguilar I 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J Dairy Sci* 92, 4648-4655.
- Moav R 1973. Economic evaluation of genetic differences. Pages 319-352 in *Agricultural Genetics, Selected Topics*. R. Moav (Ed.) John Wiley and Sons. New York, NY.

- Moav R and Hill WG 1966. Specialised sire and dam lines. IV. Selection within lines. *Anim Prod* 8, 375-390.
- Moav R and Moav J 1966. Profit in a broiler enterprise as a function of egg production of parent stocks and growth rate of their progeny. *Br Poult Sci* 7, 5-15.
- Niebel E 1974. Methodik der Zuchtplanung für die Reinzucht beim Rind bei Optimierung nach Zuchtfortschritt und Züchtungsgewinn. PhD-Thesis, University of Hohenheim, Germany.
- Niebel E and Fewson D 1978. Untersuchungen zur Zuchtplanung für die Reinzucht beim Zweinutzungsgrind. *Zeitschr. f. Züchtungskunde* 50, 245-345.
- Niebel E and Fewson D 1979. Untersuchungen zur Optimierung der Zuchtplanung für die Reinzucht beim Schwein. *Zeitschr. f. Züchtungskunde* 51, 1-32.
- Niemann B 2009. Untersuchungen zu Veränderungen im Zuchtgeschehen und deren Auswirkungen auf die Hannoveraner Pferdezeit. PhD, Goettingen Univ.
- Nitter G and Graser H-U 1994. ZPLAN - a PC program to optimize livestock selection programs. *Proceedings 5th World Congress on Genetics Applied to Livestock Production* 22, 77-78.
- Norman HD and Dickinson FN 1971. An economic index for determining the relative value of milk and fat in predicted differences of bulls and cow index values of cows. *ARS-44-223. DHI. Lett.* 47, 1-34.
- Prunier A, Bonneau M, von Borell EH, Cinotti S, Gunn M, Fredriksen B, Giersing M, Morton DB, Tuytens FAM and Velarde A 2006. A review of the welfare consequences of surgical castration in piglets and the evaluation of non-surgical methods. *Animal Welfare* 15, 277-289.
- Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu Z-L, Kerstens HH, Law AS, Megens H-J, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TPL, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB and Groenen MAM 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE* 4, e6524. doi:10.1371/journal.pone.0006524.
- Rendel JM and Robertson A 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *Genetics* 50, 1-8.
- Rydhmer L, Lundström K and Andersson K 2010. Immunocastration reduces aggressive and sexual behavior in male pigs. *animal* 4, 965-972.
- Schade W 1996. Entwicklung eines Besamungszuchtprogramms für die Hannoversche Warmblutzucht. PhD thesis, Georg-August University Goettingen, Germany.

- Schaeffer LR 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics* 123, 218-223.
- Sellier P, Le Roy P, Fouilloux MN, Gruand J and Bonneau M 2000. Responses to restricted index selection and genetic parameters for fat androstenone level and sexual maturity status of young boars. *Livest Prod Sci* 63, 265-274.
- Simianer H 2009. The potential of genomic selection to improve litter size in pig breeding programs. In: Proc. of 60th Annu. Mtg. Eur. Assoc. Anim. Prod., Barcelona, Spain, August 24-27, 2009. Wageningen Academic Publishers, The Netherlands.
- Simon DL and Buchenauer D 1993. Genetic diversity of European livestock breeds. EAAP Publication No. 66, Wageningen Press.
- Sitzenstock F 2012. Genetic progress and inbreeding rate in complex breeding programmes - applications to sport horses and laying hens. PhD thesis, Georg-August University, Goettingen.
- Smith C, James JW and Brascamp EW 1986. On the derivation of economic weights in livestock improvement. *Anim Prod* 43, 545-551.
- Täubert H, Reinhardt F and Simianer H 2010. ZPLAN+, a new software to evaluate and optimize animal breeding programs. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production. 1-6 August 2010, Leipzig, Germany.
- Thorén Hellsten E, Näsholm A, Jorjani H, Strandberg E and Philipsson J 2009. Influence of foreign stallions on the Swedish Warmblood breed and its genetic evaluation. *Livest Sci* 121, 207-214.
- Uimari P and Tapio M 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *J Anim Sci* 89, 609-614.
- van Arendonk JAM, Smith C, Kennedy BW 1989. Method to estimate genotype probabilities at individual loci in farm livestock. *Theor Appl Genet* 78, 735-740.
- 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.
- von Borell E, Baumgartner J, Giersing M, Jäggin N, Prunier A, Tuyttens FAM and Edwards SA 2009. Animal welfare implications of surgical castration and its alternatives in pigs. *animal* 3, 1488-1496.
- von Lengerken G and Schwark H-J 2002. Exterieur und Leistungen in der Pferdezucht – Alleskönner oder Spezialisten. *Arch Tierz Dummerstorf* 45, 68-79.
- Weller JI 1994. Economic aspects of animal breeding. Chapman & Hall, London, Great Britain.
- Willham RL 1979. Evaluation and direction of beef sire evaluation programs. *J Anim Sci* 49, 592-599.

- Willam A and Simianer H 2011. Tierzucht. Eugen Ulmer Verlag, Stuttgart, Germany.
- Willam A, Nitter G, Bartenschlager H, Karras K, Niebel E and Graser H-U 2008. ZPLAN- Manual for a PC-program to optimize livestock selection schemes.
- Windig JJ, Mulder HA, ten Napel J, Knol EF, Mathur PK and Crump RE 2012. Genetic parameters for androstenone, skatole, indole, and human nose scores as measures of boar taint and their relationship with finishing traits. *J Anim Sci* 90, 2120-2129.
- Wolfová M, Nitter G, Wolf J and Fiedler J 2001. Impact of crossing system on relative economic weights of traits in purebred pig populations. *J Anim Breed Genet* 118, 389-402.
- Wünsch U, Nitter G and Schüler L 1999. Genetic and economic evaluation of genetic improvement schemes in pigs. *Arch Tierz Dummerstorf* 42, 571-582.
- Wünsch U, Nitter G, Bergfeld U and Schüler L 2000. Genetic and economic evaluation of genetic improvement schemes in pigs. *Arch Tierz Dummerstorf* 43, 139-149.

2ND CHAPTER**Integration of genomic information into Sport Horse breeding programs for optimization of accuracy of selection**

A.M. Haberland¹, U. König von Borstel¹, H. Simianer¹ and S. König²

¹Department of Animal Sciences, Georg-August-University Goettingen, Goettingen, Germany

²Department of Animal Breeding, University of Kassel, Witzenhausen, Germany

Published in *Animal*

Volume 6, Issue 9, pages 1369–1376, March 2012

© 2012 The Animal Consortium

doi: 10.1017/S1751731112000626

Abstract

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 genomic breeding value) 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 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) due to 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 yrs 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 index theory; ii) and 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 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. In 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 *osteocondrosis* (**OC**) with a heritability of 0.15 (Schober, 2003; Pieramati *et al.*, 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 (Fig. 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, e.g. to select mares as potential donors for embryo transfer. 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 were used as information sources in the index.

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 \mathbf{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 genomic breeding value), 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 \mathbf{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 \mathbf{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 genomic breeding value estimated from the marker genotype m and to the conventional breeding value of the phenotypic trait y . Matrix \mathbf{C} was the matrix for variances and covariances of breeding values, i.e.

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

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

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

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

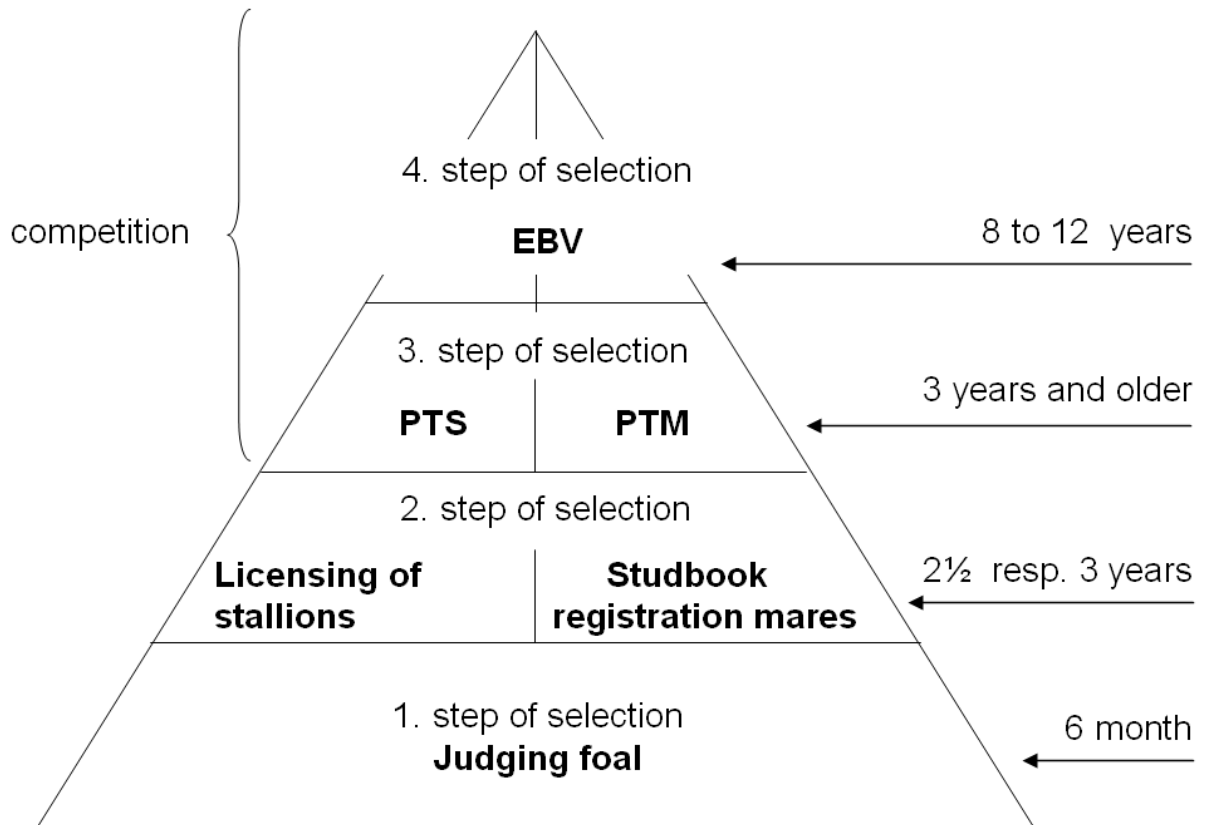


Figure 1. Pyramid structure of the breeding program of the German riding horse (EBV = Estimated Breeding Value, PTS = Performance Test Stallions, PTM = Performance Test Mares).

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 (Fig. 1), i.e. 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 \mathbf{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 \mathbf{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 \mathbf{P} and \mathbf{G} . The vector \mathbf{w} of economic weights and the (co)variance matrix \mathbf{C} for breeding values were identical to \mathbf{w} and \mathbf{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 (Fig. 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 \mathbf{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 \mathbf{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 \mathbf{w} of economic weights and the (co)variance matrix \mathbf{C} for breeding values were identical to \mathbf{w} and \mathbf{C} in scenario I.

Scenario IV: Combination of breeding values into a combined index.

Scenarios I, II, and 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 $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{w}$ with matrices \mathbf{P} and \mathbf{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 (SEBV). Due to less information and estimation earlier in life, the accuracy of SEBVs are generally lower than for IEBVs 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 TBVs, equal economic weights per genetic SD 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}). 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 \mathbf{P} , the variance-covariance matrix for n EBVs:

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

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 \mathbf{C} , the quadratic variance-covariance matrix for m TBVs in the overall index was:

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

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 \mathbf{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):

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

Those covariances in \mathbf{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:

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

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 50K SNP chip technology, as well as the recent release of the horse genome sequence (Wade *et al.*, 2009), provide the framework to estimate highly reliable genomic breeding values 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. Due 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 to 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 to 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 genomic breeding values in horses are carried out for Franches-Montagnes horses in Switzerland (Hasler *et al.*, 2011).

Scenario I: Genotyped young horse without own performance.

Due 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 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, i.e. 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 (Fig. 2). As known from selection index theory, additional information from further close relatives of the foal would increase 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 (Fig. 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,

e.g. dairy cattle, accuracies of GEBVs are substantially higher for production traits compared to fertility, somatic cell score, or longevity (VanRaden *et al.*, 2009). Nevertheless, based on 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 (Fig. 2) compared to the accuracy of the conventional index.

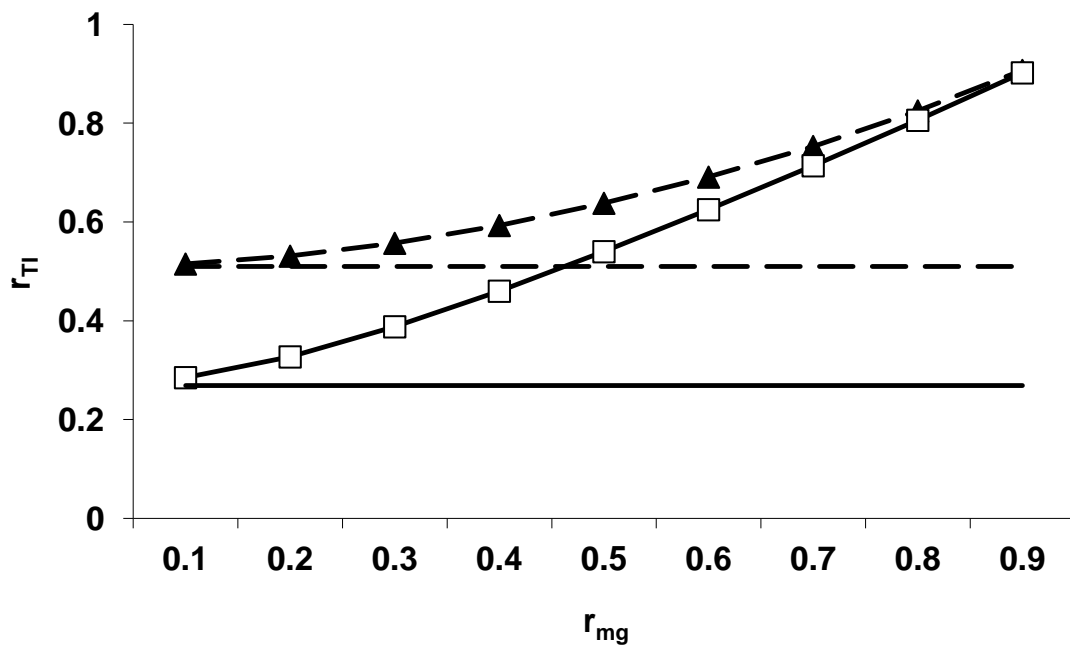


Figure 2. Correlation between index and aggregate genotype (r_{TI}) in dependency of accuracy of GEBV (r_{mg}) for a genotyped horse without own performance (Scenario I). Dashed line with black triangles: Trot ($h^2 = 0.52$); solid line with white squares: OC ($h^2 = 0.15$). Parallel lines to x-axis: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

When additionally considering economic 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 two and a half years and allow a limited number of matings until the stallion is performance tested at the age of three or four 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 embryo transfer (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 embryo transfer 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 Fig. 2 and Fig. 3). The conventional index sources own performance, performance of sire and performance of dam result in $r_{TI} = 0.45$ for $h^2 = 0.15$ (Fig. 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.

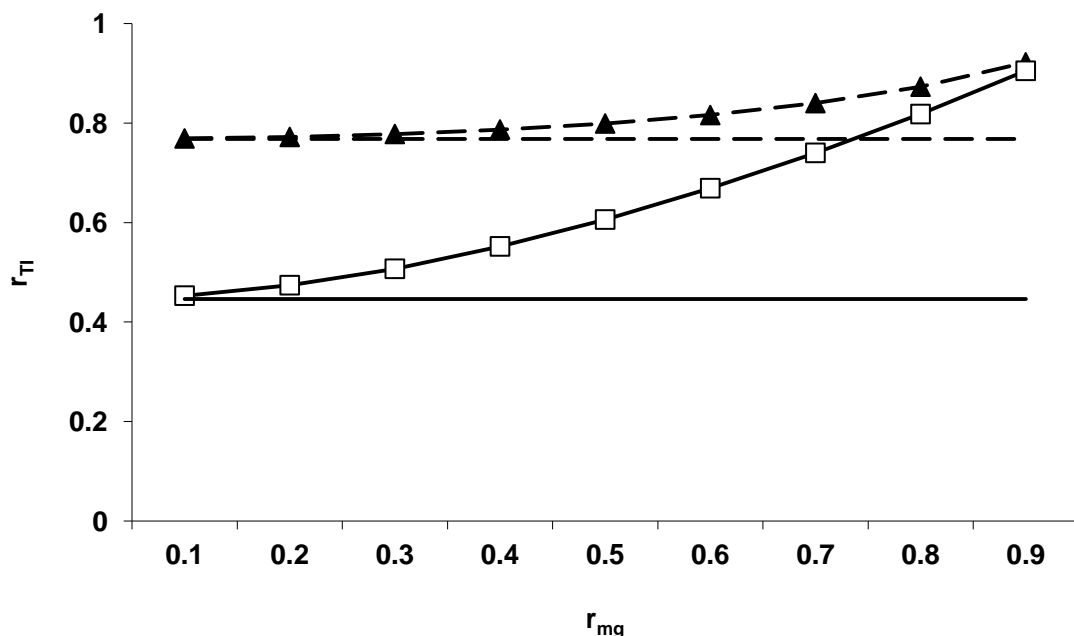


Figure 3. Correlation between index and aggregate genotype (r_{TI}) in dependency of accuracy of GEBV (r_{mg}) for a genotyped horse with own performance and performance of parents (Scenario II). Dashed line with black triangles: Trot ($h^2 = 0.52$); solid line with white squares: OC ($h^2 = 0.15$).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

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 (Thorén Hellsten *et al.*, 2006, Gerber Olsson *et al.*, 2000). 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 QTL caused by recombination events (Sonesson and Meuwissen, 2009). Also, 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 yrs 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 one. 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 5 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 5 offspring, leading to a RSR of 1.65.

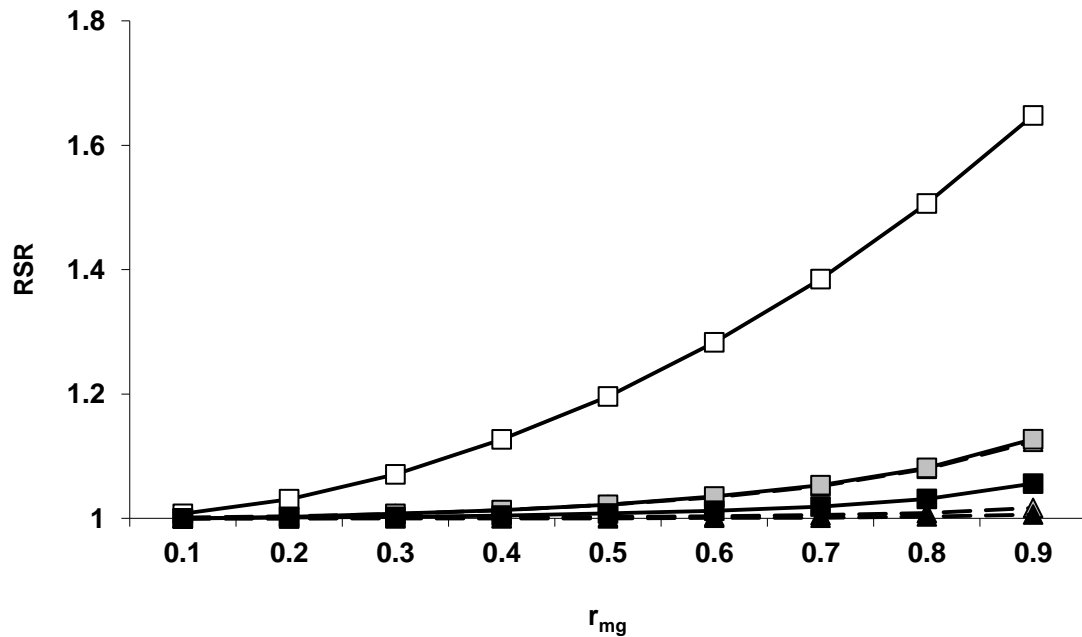


Figure 4. Relative selection response (RSR) for a genotyped horse with own performance, performance of parents, and a different number of progeny in dependency of accuracy of GEBV (r_{mg}). Dashed lines with triangles: Trot ($h^2 = 0.52$); solid lines with squares: OC ($h^2 = 0.15$). White triangles or squares: 5 progeny; grey triangles or squares: 50 progeny; black triangles or squares: 100 progeny.

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.

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, e.g. 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, e.g. 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, II, and 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 set-up 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 e.g. 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 $r_{\overline{T}}$ and weighting factors (b-values) for the combination of GEBV, IEBV, and SEBV by altering r_{mg} are depicted in Figure 5 and Figure 6, respectively. Independently of the accuracy for SEBV, i.e. 0.8 versus 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} (Fig. 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.

However, in reality accuracies of GEBVs will be not 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 which is 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 genomic selection and performance testing. Considerably earlier selection of male and female animals for breeding is possible due to the estimation of genomic breeding values for animals without phenotypic data.

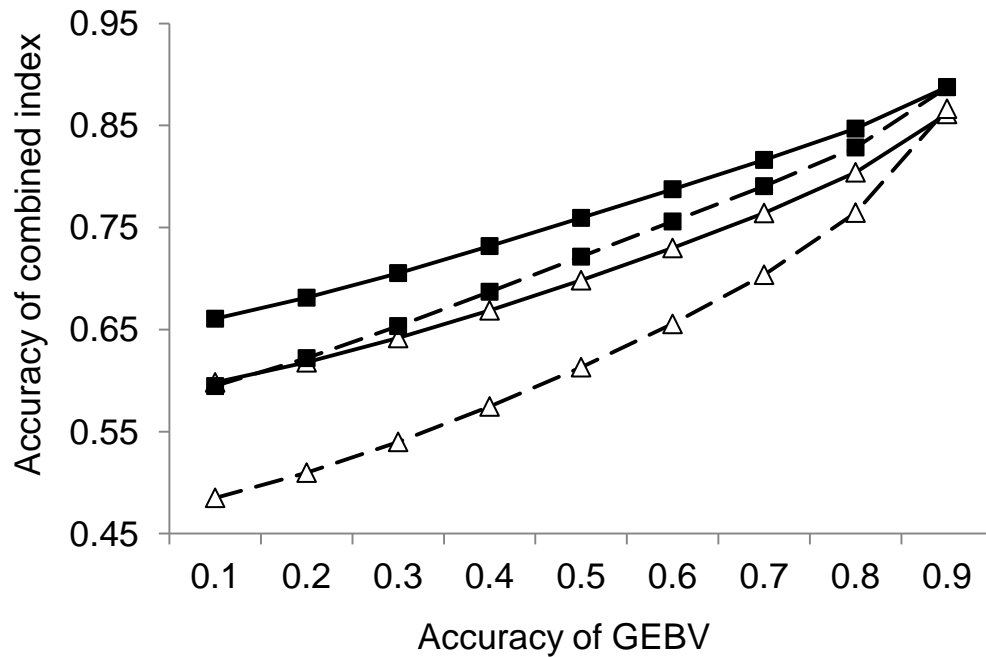


Figure 5. Accuracy of the combined index including GEBV, IEBV, and SEBV in dependency of accuracies of GEBV (r_{mg}) for equal economic weights per genetic SD (solid lines: correlation between SEBV and IEBV = 0.95; dashed lines: correlation between SEBV and IEBV = 0.50; black squares: accuracy of SEBV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

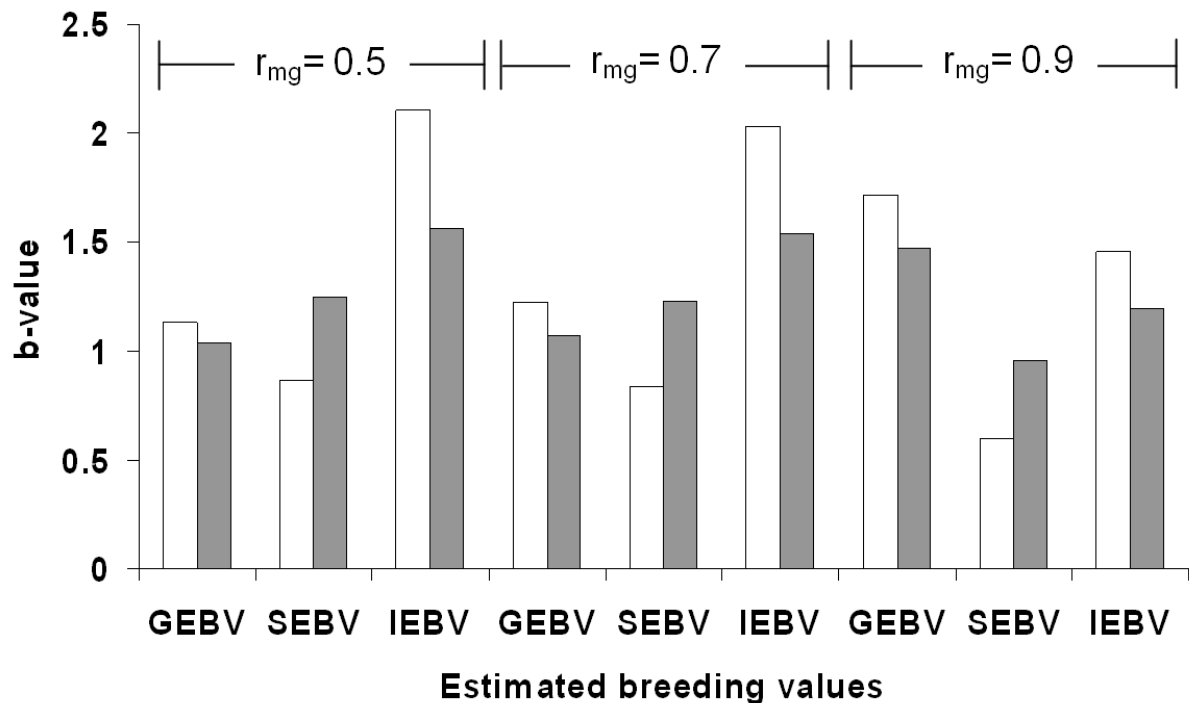


Figure 6. Weighting factors (b-values) for the combination of GEBV, IEBV, and SEBV in a combined index in dependency of accuracies of GEBV (r_{mg}) for equal economic weights per genetic SD (white bars: accuracy of SEBV = 0.5; black bars: accuracy of SEBV = 0.8; accuracy of IEBV = 0.85 for all scenarios).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

Conclusions

Based on our results, application of genomic selection 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. *J Anim Breed Genet* 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. *Livest Sci* 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. *Livest Prod Sci* 65, 81-89.
- German Equestrian Federation (FN) 2008. Breeding program of the German Riding Horse. In *Jahrbuch Sport und Zucht* 2008, pp. 84. 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. *J Anim Sci* 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. *J Anim Breed Genet* 128, 394-406.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci* 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. *Livest Prod Sci* 88, 77-84.
- König S and Swalve HH 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. *J Dairy Sci* 92, 5292-5303.
- König S, Simianer H and Willam A 2009. Economic evaluation of genomic breeding programs. *J Dairy Sci* 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, Guldbbrandtsen B, Liu Z, Reents R, Schrooten C, Seefried FR and Su G 2010. Improving genomic prediction by EuroGenomics collaboration. In *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 Assoc. 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 thesis, Goettingen Univ.
- Philipsson J, Árnason Th and Bergsten K 1990. Alternative selection strategies for Performance of the Swedish Warmblood horse. *Livest Prod Sci* 24, 273-285.
- Pieramati C, Pepe M, Silvestrelli M and Bolla A 2003. Heritability estimation of *osteocondrosis dissecans* in Maremmano horses. *Livest Prod Sci* 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 thesis, Goettingen Univ.
- Schaeffer LR 2006. Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet* 123, 218-223.
- Schober M 2003. Schätzung von genetischen Effekten beim Auftreten von *Osteochondrosis dissecans* beim Warmblutpferd. PhD thesis, Goettingen Univ.
- Schorm G 1983. Analyse der phänotypischen Entwicklung des Warmblutpferdes von der Geburt bis zum 3jährigen Pferd und Einflüsse von genetischen und umweltbedingten Faktoren. PhD thesis, Leipzig Univ.
- 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 An-*

- imal 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. *Gen Sel Evol* 41, 37.
- Spelman RJ and Garrick DJ 1998. Genetic and economic responses for within-family marker-assisted selection in dairy cattle breeding schemes. *J Dairy Sci* 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. *J Anim Sci* 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 in 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. *Livest Sci* 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. *J Dairy Sci* 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 – Alteskönner oder Spezialisten. *Archiv für Tierzucht, Dummerstorf* 45, 68-79.
- Wade CM, Giolotto 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, Zody MC, Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander ES and Lindblad-Toh K 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* 326, 865-867.

3RD CHAPTER**Interplay between heritability, genetic correlation and economic weighting in a selection index with and without genomic information**

A.M. Haberland¹, E.C.G. Pimentel², F. Ytournal¹, M. Erbe¹ and H. Simianer¹

¹Department of Animal Sciences, Georg-August-University Goettingen, Goettingen, Germany

²Department of Animal Breeding, University of Kassel, Witzenhausen, Germany

Published online in Journal of Animal Breeding and Genetics
on 29th of August 2013

© 2013 Blackwell Verlag GmbH

doi:10.1111/jbg.12051

Summary

The availability of genomic information demands proper evaluation on how the kind (phenotypic versus genomic) and the amount of information influences the interplay of heritability (h^2), genetic correlation ($r_{G_i G_j}$) and economic weighting of traits with regards to the standard deviation of the index (σ_I). As σ_I is directly proportional to response to selection it was the chosen parameter for comparing the indices. Three selection indices incorporating conventional and genomic information for a two trait (i and j) breeding goal were compared. Information sources were chosen corresponding to pig breeding applications. Index I incorporating an own performance in trait j served as reference scenario. In index II, additional information in both traits was contributed by a varying number of full sibs (2, 7, 50). In index III the conventional own performance in trait j was combined with genomic information (GBVs) for both traits. The number of animals in the reference population ($N_p = 1'000, 5'000, 10'000$) and thus the accuracy of GBVs were varied. With more information included in the index, σ_I became more independent of $r_{G_i G_j}$, h_j^2 and relative economic weighting. This applied for index II (more full sibs) and for index III (more accurate GBVs). Standard deviations of index II with 7 full sibs and index III with $N_p = 1'000$ were similar when both traits had the same heritability. If the heritability of trait j was reduced ($h_j^2 = 0.1$), σ_I of index III with $N_p = 1'000$ was clearly higher than for index II with 7 full sibs. When enhancing the relative economic weight of trait j , the decrease in σ_I of the conventional full sib index was much stronger than for index III. Our results imply that $N_p = 1'000$ can be considered a minimum size for a reference population in pig breeding. These conclusions also hold for comparing the accuracies of the indices.

Keywords: Genetic gain; economic weighting; genomic selection; pig breeding; selection index.

Introduction

Accounting for the interrelationships between heritability, genetic correlation and economic weighting is essential for optimizing the response to selection in all traits of a multi-trait breeding goal. Accuracy of the selection index (r_{TI}) is strongly influenced by the combination of these parameters. However, the standard deviation of an index (σ_I) is more appropriate for comparison of indices with varying economic weights since σ_I shows more proportionality to expected response to selection than accuracy. Smith (1983) investigated the influence of changing economic weights on efficiency of indices comprising two or more traits. He found the balance of traits considered in an index (measured as the product of economic weighting and heritability) to significantly influence its efficiency. Accordingly, an unbalanced index in which one or more traits dominate the index is more sensitive to changes in economic weighting within these dominant traits. For example, genetic changes for functional traits in bull dam selection typically lean towards the undesirable direction when using current economic weighting which favors production traits (Hansen Axelsson et al., 2011). In order to meet the currently required balance between production and functional traits, the authors propose a readjustment of the economic weighting of breeding goal traits. Assessing the efficiency of an index by the balance of its traits, however, does not account for different information sources.

The availability of genomic information sources raises the question whether the kind of information (phenotypic versus genomic) has an impact on the interaction of heritability, genetic correlation and economic weighting of traits concerning σ_I as an indicator for genetic gain. After this concept was formulated by Meuwissen et al. (2001) genomic selection (GS) has been demonstrated to be a valuable tool for increasing the accuracy of selection. Benefits of GS are considered highest for dairy cattle due to the long generation interval, which can be shortened dramatically by increasing accuracy of breeding values early in life (Schaeffer, 2006). Moreover, genetic gain for functional traits has been found to be higher when selecting on the basis of genomic breeding values (GBVs) if the training population sizes are assumed to be equal for production and functional traits (König et al., 2009; König and Swalve, 2009; Buch et al., 2012). In this case GS enables more sustainable breeding strategies.

The potential of shortening generation intervals due to high accuracy of selection at a young age as well as more efficient selection on functional traits was also found for sport horse breeding (Haberland et al., 2012). Studies considering the use of molecular data in beef cattle selection (Van Eenennaam et al., 2011; Pimentel and König, 2012) have con-

firmed the potential for increasing the accuracy of the index, discounted return and genetic gain for different production systems and breeding goal traits, such as meat quality. Considering GBVs as an additional information source can also be beneficial in pig breeding. Simianer (2009) and Haberland et al. (2010) found increased accuracy and response to selection which enables more intensive selection.

Dekkers (2007) proposed an approach for combining conventional and genomic information using selection index theory. The studies mentioned above have applied this or similar approaches to a particular production system, assuming a specific configuration of genetic parameters. The aim of our study was a more general assessment of the impact of heritability, genetic correlation and relative economic weighting among breeding goal traits on σ_I in the presence of genomic information. Understanding this multidimensional interaction is challenging, as it is additionally affected by the composition of the respective index, i.e., of the available information sources.

We report results of a case study with a two trait breeding goal and information sources chosen according to a pig breeding scheme. We used the selection index approach proposed by Dekkers (2007) in order to account for genomic information. His formula for the correlation between GBVs of different traits was slightly modified so that the proportion of genetic variance captured by the markers can be assumed to be diverse for different traits. Specifically we were interested in the following questions: (i) How does the composition of information sources affect the impact of heritability, genetic correlation and relative economic weighting on the standard deviation of an index? (ii) How many animals are required for the setup of a reference population to achieve a meaningfully higher σ_I for the genomic index than for the conventional index? (iii) What is the influence of heritability on the relative performance of a genomic and a conventional index? (iv) How can the correlation of prediction errors of GBVs and the expected genetic progress be affected by an overlap of the training sets used for different traits?

Materials and Methods

In order to assess the interrelationship between the parameters heritability (h^2), genetic correlation ($r_{G_i G_j}$) and economic weighting, three different indices were investigated according to selection index theory (Hazel, 1943), which is briefly summarized as follows:

In general we consider a true breeding goal (\mathbf{T}) which is a linear combination of n traits, i.e. $\mathbf{T} = \mathbf{w}'\mathbf{u}$ where \mathbf{w} is a column vector of length n containing the economic weights, and \mathbf{u} is a column vector of length n containing the true breeding values for trait 1, ..., n . \mathbf{T} is

estimated by an index (I) which is a linear combination of m observations, $I = \mathbf{b}'\mathbf{y}$, where \mathbf{b} is a column vector of length m containing the unknown index weights and \mathbf{y} is a column vector of length m containing the available phenotypic observations, which are pre-corrected for non-genetic effects.

Index weights are estimated from the equation $\mathbf{Pb} = \mathbf{Gw}$ as $\mathbf{b} = \mathbf{P}^{-1}\mathbf{Gw}$ which maximizes the accuracy of the index as a predictor of the breeding goal which is reflected by the correlation between T and I :

$$r_{TI} = \sqrt{\frac{\sigma_I^2}{\sigma_T^2}}$$

where $\sigma_I^2 = \mathbf{b}'\mathbf{Pb} = \mathbf{b}'\mathbf{Gw}$ and $\sigma_T^2 = \mathbf{w}'\mathbf{Cw}$. Matrix \mathbf{P} contains the phenotypic (co)variances of the tested traits. Matrix \mathbf{G} includes the genetic covariances among m information sources (rows) and n breeding goal traits (columns). Matrix \mathbf{C} comprises the covariances among true breeding values for breeding goal traits.

Since we considered a two-trait situation, $m = 2$ and the relative economic weighting of the traits was

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

The weighting factor of trait i (w_i) was varied from 0 to 1. Relative economic weighting of trait j equals $1 - w_i$.

Index I (own performance): As a reference scenario with minimal complexity a selection index was set up for two breeding goal traits i and j . The only information source was an own performance in trait j . Hence, matrices \mathbf{P} , \mathbf{G} and \mathbf{C} were:

$$\mathbf{P} = \begin{bmatrix} \sigma_{P_j}^2 \end{bmatrix},$$

$$\mathbf{G} = \begin{bmatrix} r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & \sigma_{G_j}^2 \end{bmatrix}$$

and

$$\mathbf{C} = \begin{bmatrix} \sigma_{G_i}^2 & r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} \\ r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & \sigma_{G_j}^2 \end{bmatrix}$$

where σ_{G_i} and σ_{G_j} are the additive-genetic standard deviations of both traits.

Index II (own and full sib performance): Conventional performance of a varying number of full sibs (2; 7; 50) in both traits i and j was added to index I described above. Matrices \mathbf{P} and \mathbf{G} were set up as follows:

$$\mathbf{P} = \begin{bmatrix} \sigma_{P_i}^2 & ar_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & a\sigma_{G_j}^2 \\ \cdot & \left(\frac{1+(n-1)ah_i^2}{n} \right) \sigma_{P_i}^2 & \frac{1}{n} \left(\sigma_{P_i P_j} \sigma_{P_i} \sigma_{P_j} + (n-1) ar_{G_i G_j} \sigma_{G_i} \sigma_{G_j} \right) \\ \text{sym.} & \cdot & \left(\frac{1+(n-1)ah_j^2}{n} \right) \sigma_{P_j}^2 \end{bmatrix}$$

and

$$\mathbf{G} = \begin{bmatrix} r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & \sigma_{G_j}^2 \\ a\sigma_{G_i}^2 & ar_{G_i G_j} \sigma_{G_i} \sigma_{G_j} \\ ar_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & a\sigma_{G_j}^2 \end{bmatrix}$$

where n denotes the number of full sibs and $a = 0.5$ is the relationship coefficient between full sibs.

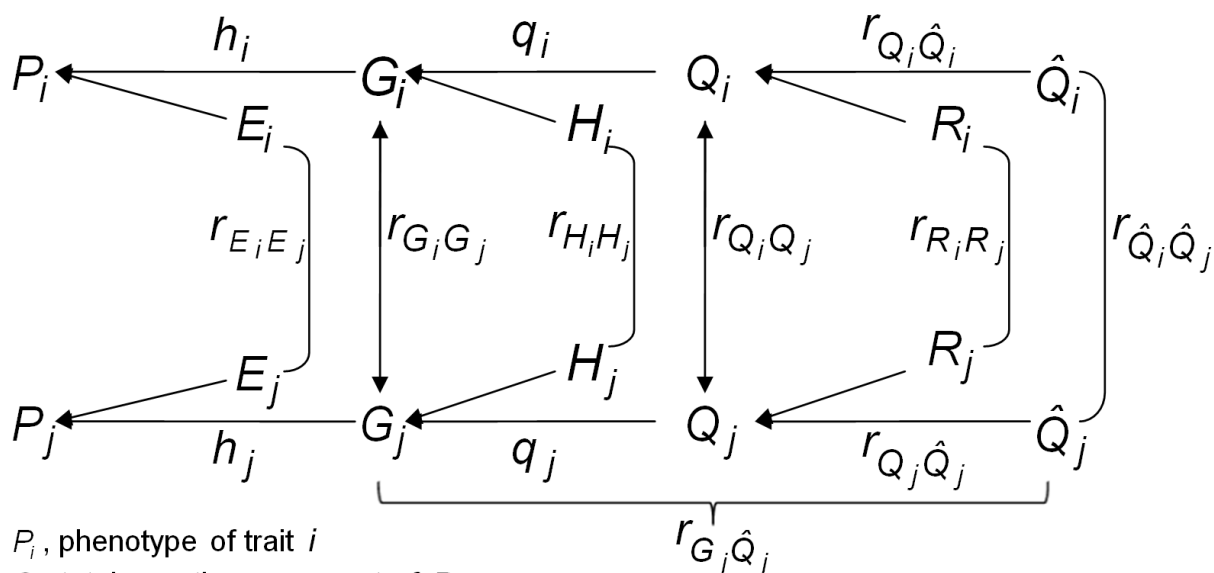
The first line of \mathbf{P} and \mathbf{G} refers to the conventional own performance in trait j , the second and third line correspond to the performance of full sibs in traits i and j , respectively.

Matrix \mathbf{C} is the same as in index I.

Index III (conventional and genomic own performance): In addition to the conventional own performance in trait j , it was assumed that the selection candidate was genotyped and thus GBVs for both traits i and j were incorporated into the index considering them as auxiliary traits with a heritability of 1 (Dekkers, 2007). Genomic information on sibs, therefore, provides no additional information for the candidate as the genomic prediction accurately estimates one part of the genetic variance of the trait and no amount of correlated observations can improve its accuracy. Components of matrix \mathbf{P} and \mathbf{G} of index III were for the most part calculated following the selection index approach by Dekkers (2007). For this reason, we will restate the main points of his considerations.

The path coefficient diagram depicted in Figure 1 shows the interrelationships between two traits i and j , their corresponding GBVs and further components of the model. In brief, the phenotype P is dissected into an additive genetic component G and a residual E . G is considered to be the sum of a true genomic effect Q that is associated with the SNP array used,

and a remaining component H that is not associated with the SNP array. Q is predicted by \hat{Q} and the corresponding residual is termed R . This scheme is set up for both traits, so that all corresponding components are linked by a correlation.



P_i , phenotype of trait i

G_i , total genetic component of P_i

E_i , environmental component of P_i

Q_i , component of G_i being associated with markers

H_i , component of G_i being independent of markers

\hat{Q}_i , EBV for Q_i based on marker data

R_i , prediction error of \hat{Q}_i

h_i^2 , heritability of trait i

q_i^2 , proportion of genetic variance associated with markers for trait i

$r_{Q_i \hat{Q}_i}$, accuracy of \hat{Q}_i as a predictor for Q_i

$r_{G_i G_j}$, genetic correlation between traits i and j

$r_{Q_i Q_j}$, correlation between Q_i and Q_j

$r_{\hat{Q}_i \hat{Q}_j}$, correlation between GBVs \hat{Q}_i and \hat{Q}_j

$r_{E_i E_j}$, correlation between environmental components E_i and E_j

$r_{H_i H_j}$, correlation between components H_i and H_j

$r_{R_i R_j}$, correlation between prediction errors R_i and R_j of \hat{Q}_i and \hat{Q}_j

Figure 1 Path coefficient diagram adapted from Dekkers (2007) showing the relationships between two traits and their genetic components.

Seven parameters are needed to set up the required matrices: the phenotypic standard deviation σ_{P_j} , the additive genetic standard deviations σ_G of both traits, the correlation $r_{G\hat{Q}}$ between the true breeding value G and the corresponding GBV \hat{Q} for both traits, the phenotypic correlation $r_{P_i P_j}$, and the genetic correlation $r_{G_i G_j}$. In the following, the calculations for the required parameters will be explained.

As the heritability of the genomic traits \hat{Q}_i and \hat{Q}_j was assumed to be 1, the phenotypic variance of the genomic traits corresponds to the genetic variance which was calculated following Dekkers (2007) as

$$\sigma_{\hat{Q}}^2 = r_{G\hat{Q}}^2 \sigma_G^2.$$

For trait i , the genetic correlation between the true breeding value G_i and the corresponding GBV \hat{Q}_i is

$$r_{G_i \hat{Q}_i} = q_i r_{Q_i \hat{Q}_i}.$$

The parameter q_i^2 describes the proportion of genetic variance captured by markers for trait i ; and $r_{Q_i \hat{Q}_i}$ is the accuracy of \hat{Q}_i as a predictor of Q_i . This accuracy is assumed to be < 1 because q^2 depends on marker density and on the extent and pattern of LD (Dekkers, 2007). Erbe et al. (2011) used cross-validated data to empirically determine q for genotyped Holstein Friesian bulls and found it to be in the area of $q \approx 0.9$ for the traits milk yield and somatic cell score. Due to the lack of such empirical data in pigs, we adopted this value of $q = 0.9$ for our calculations of $r_{G\hat{Q}}$ for both GBVs.

The correlation between the phenotype P_i and the GBV of trait i is:

$$r_{P_i \hat{Q}_i} = h_i r_{G_i \hat{Q}_i}$$

For the calculation of the accuracy of \hat{Q}_i as a predictor of Q_i we used a formula derived by Daetwyler *et al.* (2010):

$$r_{Q_i \hat{Q}_i} = \sqrt{\frac{N_p h_i^2}{N_p h_i^2 + M_e}}$$

where N_p is the number of individuals in the reference population which was varied within our genomic index ($N_p = 1'000; 5'000; 10'000$). The heritability h_i^2 was replaced with the reliability r_i^2 of the quasi-phenotypes, i.e. of the conventional EBVs of the animals in the ref-

erence population. We assumed $r^2 = 0.49$ for both traits i and j . M_e is the effective number of chromosome segments segregating in the population and can be approximated with $2N_e Lk / \log(N_e L)$ as proposed by Goddard et al. (2011). N_e denotes the effective population size, L is the average length of a chromosome in Morgans and k is the number of chromosome pairs. Assuming $N_e = 100$, $k = 19$ and $L = 1.2$ Morgans (the length of the porcine genome being ~ 23 Morgans, Rohrer et al., 1996) $M_e = 952.57$ which we rounded to $M_e = 1'000$.

The genetic correlation between the true breeding value of trait i and the GBV of trait j is

$$r_{G_i \hat{Q}_j} = r_{G_i G_j} r_{G_j \hat{Q}_j}$$

and the corresponding phenotypic correlation is

$$r_{P_i \hat{Q}_j} = h_i r_{G_i G_j} r_{G_j \hat{Q}_j}.$$

Our formula to determine the genetic correlation $r_{\hat{Q}_i \hat{Q}_j}$ between GBVs of two different traits differs from Dekkers (2007) in that it was derived assuming q_i and q_j not being necessarily identical (see Appendix). Dekkers (2007) assumed the proportion of genetic variance captured by the markers, q_i^2 and q_j^2 , to be the same in both traits. This might be the case with high density SNP-arrays (and ultimately with the whole genome sequence) where the proportion of genetic variance captured by the markers should approach 1 for any single trait. In practice, SNP-arrays might not be dense enough to guarantee that the same amount of genetic variance is captured for all traits. Our formula

$$r_{\hat{Q}_i \hat{Q}_j} = r_{G_i G_j} q_i q_j r_{Q_i \hat{Q}_i} r_{Q_j \hat{Q}_j} = r_{G_i G_j} r_{G_i \hat{Q}_i} r_{G_j \hat{Q}_j}$$

allows for choosing different proportions q^2 for different traits, which might be useful when empirically determined values for q^2 are available from studies like Erbe et al. (2011).

Matrices \mathbf{P} and \mathbf{G} were set up as follows:

$$\mathbf{P} = \begin{bmatrix} \sigma_{P_i}^2 & r_{G_i \hat{Q}_i} \sigma_{G_i} \sigma_{\hat{Q}_i} & r_{G_i \hat{Q}_j} \sigma_{G_i} \sigma_{\hat{Q}_j} \\ \cdot & \sigma_{\hat{Q}_i}^2 & r_{\hat{Q}_i \hat{Q}_j} \sigma_{\hat{Q}_i} \sigma_{\hat{Q}_j} \\ \text{sym.} & \cdot & \sigma_{\hat{Q}_j}^2 \end{bmatrix}$$

and

$$\mathbf{G} = \begin{bmatrix} r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} & \sigma_{G_j}^2 \\ r_{G_i \hat{Q}_i} \sigma_{G_i} \sigma_{\hat{Q}_i} & r_{G_j \hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_i} \\ r_{G_i \hat{Q}_j} \sigma_{G_i} \sigma_{\hat{Q}_j} & r_{G_j \hat{Q}_j} \sigma_{G_j} \sigma_{\hat{Q}_j} \end{bmatrix}$$

Again, the first line of both matrices corresponds to the own performance in trait j . The second and third line refer to the genomic own performance in traits i and j , respectively. Matrix \mathbf{C} was the same as in index I.

Index IV (commonly available conventional performance): To relate σ_I of index III to a more practical situation we set up an index comprising information sources commonly used in pig breeding. This index included an own performance in trait j as well as parents' performance, performance of 5 full sibs and 80 half sibs in both traits (matrices not shown).

Using this basic model, the following cases were studied:

1) Variation of relative economic weighting

Indices I, II and III were compared in terms of σ_I . The impact of genetic correlation and relative economic weighting among traits i and j on the standard deviation of an index was analyzed by varying $r_{G_i G_j}$ from -1 to 1 and w_{-i} from 0 to 1, both in steps of 0.01, respectively. The heritability and the phenotypic standard deviation of both traits were assumed to be equal ($h^2 = 0.3$; $\sigma_P = 1$) and the phenotypic correlation between the traits was assumed to be $r_{P_i P_j} = 0$.

2) Variation of heritability of one trait

The heritability of trait j was reduced to 0.1 while the heritability of trait i remained 0.3. The genetic correlation and relative economic weighting for traits i and j were varied as in the first scenario, and the phenotypic correlation remained 0. Results were assessed in terms of σ_I . Besides assuming equal reliabilities ($r_i^2 = r_j^2 = 0.49$) of the quasi-phenotypes in the reference population for both traits we also considered a scenario with different reliabilities of the conventional breeding values used as quasi-phenotypes ($r_i^2 = 0.49$ and $r_j^2 = 0.20$). This

would e.g. reflect a situation where different numbers of offspring per boar are tested for the different traits. The effect of the heritability on σ_I was assessed in more detail by varying h_j^2 from 0.1 to 1 in steps of 0.01 while keeping $w_{\cdot i}$ constant at 0.5 and varying $r_{G_i G_j}$ as described above.

A crucial assumption in the genomic scenarios described so far is, that GBVs of the two traits are only correlated to an extent which results automatically from the genetic correlation between the two traits and the different accuracies, as described in the Appendix. In many cases, though, GBVs for the two traits are derived from the same training set, and this may cause the errors of the GBVs to be more (or less) correlated than expected in the closed setting. Therefore, we also analyzed the case in which the correlation of the errors of the GBV differs from the result in eq. 4 in the Appendix by adding a constant

Results and Discussion

1) Variation of relative economic weighting

To give a first impression on how the parameter's relative economic weighting and genetic correlation of traits i and j interact, the standard deviation of index I has been displayed in Figure 2 using a three-dimensional graph. On the x-axis, $r_{G_i G_j}$ varies from 1 to -1. The y-axis shows the variation of relative economic weighting, i.e. $w_{\cdot i} = 1$ implies that economic weighting lies completely on trait i and $w_{\cdot i} = 0$ implies that economic weighting lies completely on trait j . The standard deviation of the index is shown on the z-axis. With $\sigma_P = 1$ and $h^2 = 0.3$ for both traits, the maximum achievable σ_I for all indices is 0.548 ($=\sqrt{0.3}$).

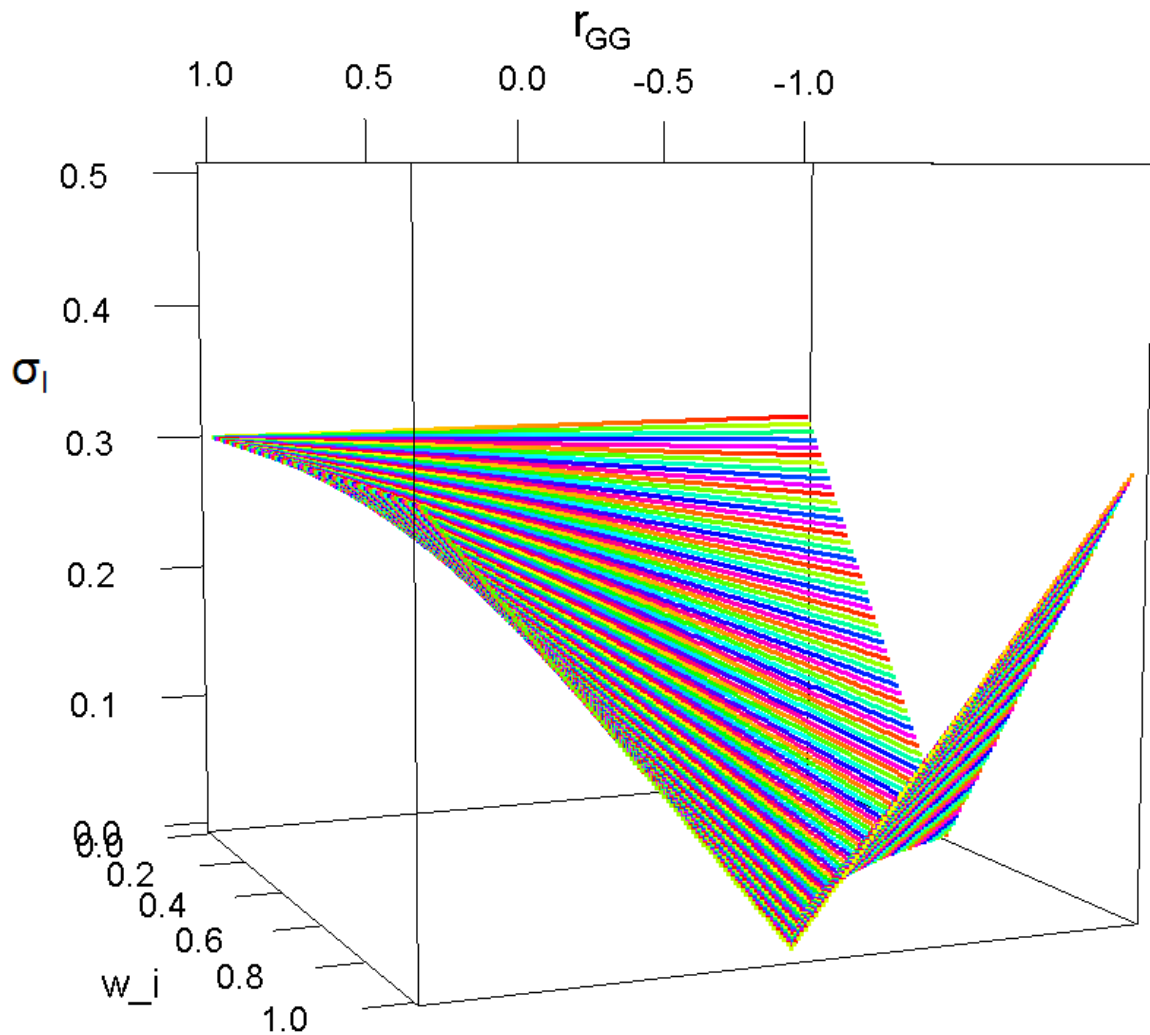


Figure 2 The standard deviation (σ_I) of index I as a function of the genetic correlation ($r_{G_iG_j}$) between traits i and j , and of relative economic weighting of trait i (w_i).

Some general trends can be observed in all three indices. Depending on $r_{G_iG_j}$, information on one trait can contribute to the other trait. Thus, information on trait j increases σ_I even if economic weighting lies completely on trait i . This is true for conventional information sources as well as for genomic information, but the extent to which σ_I can be increased by the correlated trait differed between indices I, II and III. The maximal value of σ_I of all indices occurred at $r_{G_iG_j} = 1$ for all w_i as well as at $r_{G_iG_j} = -1$ for w_i being either 0 or 1. For $r_{G_iG_j} = -1$, σ_I is a direct function of w_i . In this case, σ_I and therefore expected genetic gain is maximal with w_i being 0 or one, and σ_I is close to zero when w_i approaches 0.5. For exactly equal economic weights of both traits ($w_i = 0.5$) and $r_{G_iG_j} = -1$, σ_I was zero in all scenarios considered, because the same weight is assigned to traits with an opposite economic value.

A minimum of σ_I occurs for uncorrelated traits and $w_{-i} = 1$, because trait i contributes less information to the index than trait j (an own performance is only available for trait j). We also compared indices I, II and III at $r_{G_iG_j} = 0$ and $w_{-i} = 0.75$, as described below.

Index I: The standard deviation of index I was maximal ($\sigma_I = 0.3$) for $r_{G_iG_j} = 1$ irrespective of w_{-i} (Figure 2). This value was also reached for $r_{G_iG_j} = -1$ and w_{-i} being either 1 or 0. Because no phenotypic information on trait i is contributing to index I, σ_I was also maximal if economic weighting was assigned completely to trait j irrespective of $r_{G_iG_j}$.

At $w_{-i} = 1$ and $r_{G_iG_j} = 0$ no information on the only information source (own performance in trait j) was contributed to the index resulting in $\sigma_I = 0$. A minimum close to zero persisted if the economic weight of trait i was lowered towards an equal economic weighting ($w_{-i} = 0.5$) while simultaneously shifting $r_{G_iG_j}$ between traits i and j from zero to -1. This interaction between genetic correlation and relative economic weighting of trait i at minimum σ_I for $w_{-i} > 0.5$, can be expressed by

$$r_{G_iG_j} = \frac{w_{-i} - 1}{w_{-i}}.$$

For $w_{-i} = 0.25, 0.5$ and 0.75 , σ_I was 0.225, 0.150 and 0.075, respectively.

Index II: The availability of full sib performance in both traits increased σ_I in all cases. Adding performance of 2 (7; 50) full sibs in both traits to the own performance in trait j resulted in a maximum $\sigma_I = 0.356$ (0.390; 0.415) which was an increase of 19% (30%; 38%) compared to index I.

Figure 3A depicts σ_I of index I and II as a function of the economic weighting of trait i for uncorrelated traits. Figure 3B depicts σ_I of index I and index III as a function of the genetic correlation ($r_{G_iG_j}$) between traits i and j for $w_{-i} = 1$. The difference between minimum and maximum values of σ_I decreased as the number of full sibs increased. Hence, σ_I became more independent of w_{-i} and $r_{G_iG_j}$ the more full sibs were included in the index. As in index I, σ_I decreased substantially for the combination of equal economic weighting and strongly negative $r_{G_iG_j}$.

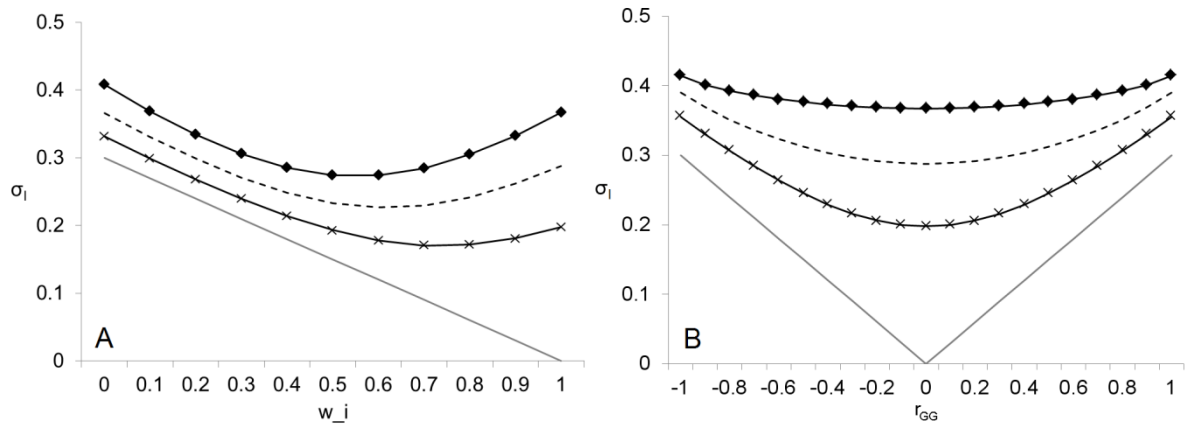


Figure 3 (A, B) Standard deviations (σ_I) of indices I (grey) and II (black) for 2 full sibs (crosses); 7 full sibs (dashed); and 50 full sibs (diamonds) as a function of the economic weighting of trait i (w_i) for uncorrelated traits (A) and as a function of the genetic correlation ($r_{G_iG_j}$) between traits i and j for $w_i = 1$ (B).

Index III: Figure 4A depicts σ_I of index I and index III as a function of the economic weighting of trait i for uncorrelated traits. Figure 4B depicts σ_I of index I and index III as a function of the genetic correlation ($r_{G_iG_j}$) between traits i and j for $w_i = 1$. The maximum σ_I of an index which includes an own performance in trait j as well as information on GBVs for both traits was 0.401 for $N_p = 1'000$. This value was higher than that of index II with 7 full sibs, but somewhat lower than σ_I of index II with 50 full sibs. It would nearly correspond to the standard deviation of index II when assuming 12 full sibs. However, the availability of 12 or 50 full sib performances in pig breeding is unlikely, while a reference population of 1'000 animals seems far more realistic. With $N_p = 5'000$, the maximum σ_I was 0.477, representing an increase of 22% compared to index II (7 full sibs). This value could not be reached by index II regardless of how many full sibs were included as information sources (asymptotic value ~ 0.420). For $N_p = 10'000$ σ_I was further increased by 27% ($\sigma_I = 0.496$).

The standard deviation of index III at $r_{G_iG_j} = 0$ and $w_i = 0.75$ was 0.231 ($N_p = 1'000$), which was slightly lower than σ_I of index II incorporating performance of 7 full sibs. For high genetic correlations, σ_I of index III ($N_p = 1'000$) was higher than σ_I of index II with 7 full sibs, whilst for a low $r_{G_iG_j}$ the standard deviation of index II with 7 full sibs was larger than σ_I of the genomic index. This superiority of index II grew slightly with increasing w_i . As for larger numbers of animals in the reference population, σ_I at $w_i = 0.75$ reached 0.330 ($N_p = 5'000$) and 0.356 ($N_p = 10'000$), respectively. Thus, the standard deviation of index III became increasingly independent of $r_{G_iG_j}$ and w_i if more animals were included in the refer-

ence population. This pattern is visualized by the flattening curves displaying σ_I in Figure 4A and 4B. Again, σ_I and therefore genetic gain was lowest when assigning similar economic weighting to traits with strongly negative $r_{G_i G_j}$.

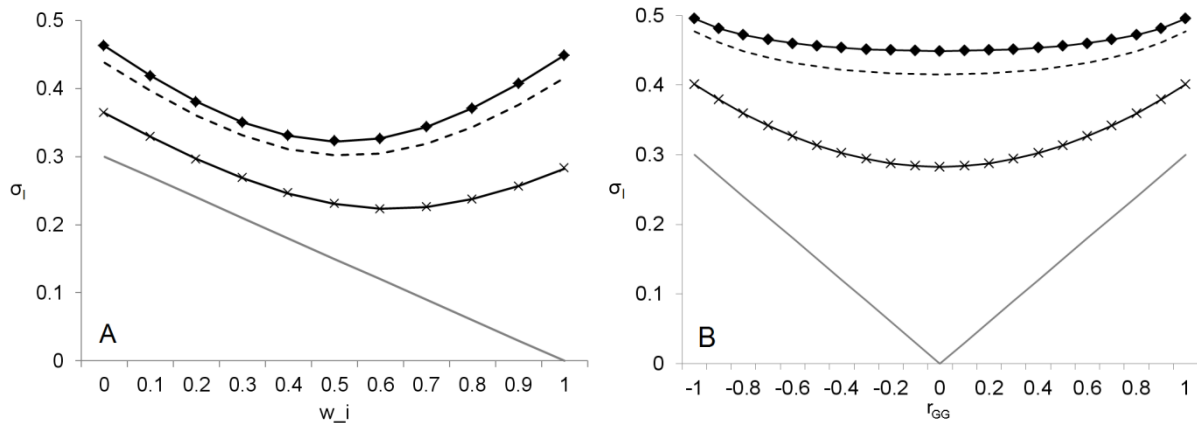


Figure 4 (A, B) Standard deviations (σ_I) of indices I (grey) and III (black) for $N_p = 1'000$ (crosses); $N_p = 5'000$ (dashed); and $N_p = 10'000$ (diamonds) as a function of the economic weighting of trait i (w_{-i}) for uncorrelated traits (A) and as a function of the genetic correlation ($r_{G_i G_j}$) between traits i and j for $w_{-i} = 1$ (B).

Index IV: The standard deviation of index III was also compared to a more practical index comprising an own performance in trait j as well as the performances of the parents, 5 full sibs and 80 half sibs in both traits, respectively. This conventional index resulted in maximum $\sigma_I = 0.402$. For $r_{G_i G_j} = 0$ and $w_{-i} = 0.75$, σ_I was 0.261. Thus, it slightly outperformed σ_I of index III with $N_p = 1'000$ which therefore can be considered a minimum size for the reference population, if GS is to be beneficial for pig breeding programs. This threshold based on rather theoretical assumptions was also indicated by our model calculations carried out for a practical breeding scheme of a Swiss sire-line comprising ten breeding goal traits (data not shown). In this investigation, the consideration of information on GBVs for all breeding goal traits and $N_p = 1'000$ resulted in higher accuracy than conventional information on own performance and performance of parents, full and half sibs. In contrast, GBVs based on $N_p = 500$ did not outperform an index accounting for conventional information only.

The results showed a tradeoff between economic weighting and genetic correlation of the traits. Standard deviations of the indices were highest when assigning economic weighting mostly to trait j for which an own performance was available. The more weight was assigned to trait i implying less performance information in the index, the more information had to be contributed via a strong genetic correlation with trait j in order to achieve high σ_I . In the extreme case in which economic weighting was assigned to trait i only while only infor-

mation on trait j was available in index I, trait j became an auxiliary trait for i . Hence, information could only be contributed via the genetic correlation and σ_I increased from 0 ($r_{G_i G_j} = 0$) to the maximum value which was achieved for $r_{G_i G_j} = -1$ or 1 (Figure 2).

In practical breeding programs this tradeoff was shown for instance by Pimentel and König (2012) who set up an index incorporating meat quality traits for beef cattle. When increasing economic weighting of a breeding goal trait for which no phenotypic information was available, accuracy of the index and genetic gain decreased. Hansen Axelsson et al. (2011) found that undesirable genetic changes in functional traits were caused by economic weights favoring highly heritable production traits in bull dam selection.

As for the impact of an auxiliary trait on genetic gain, Buch et al. (2012) concluded that phenotypic as well as genomic information on an auxiliary trait further increases genetic gain. Investigating different selection indices in dairy cattle this finding was confirmed even in cases where both phenotypic and genomic information on the breeding goal trait was available. The authors assumed an accuracy of the GBV of 0.71, which in our investigations would correspond to $N_p \approx 3'400$. Comparing σ_I of index III and $w_{-i} = 1$ for uncorrelated traits with σ_I for strongly correlated traits ($r_{G_i G_j} = -1$ or 1) showed the gain in σ_I contributed by the auxiliary trait j . Our results are in agreement with Buch et al. (2012) and indicate that additional genetic gain contributed by an auxiliary trait decreases with increasing accuracy of the GBV, i.e. increasing N_p . This finding is also in agreement with the results of Pimentel and König (2012) who evaluated the additional genetic gain contributed by an indicator trait for meat quality in beef cattle. In their study, additional gain was small assuming a GBV for the breeding goal trait with more than 2'500 animals in the reference set.

2) Variation of heritability

For equal heritabilities we found a superiority of index II with 7 full sibs over index III with $N_p = 1'000$ if traits were uncorrelated. However, when reducing h_j^2 from 0.3 to 0.1, we observed a clear advantage of index III. Although σ_I decreased for both indices, the superiority of the genomic index over the full sib index, which was observed for strong correlations, further increased. Maximum σ_I was 0.331 for index II (7 full sibs) and 0.370 for index III ($N_p = 1'000$). In Figure 5A, σ_I of index II and III is shown for uncorrelated traits and equal heritabilities of both traits, as well as for the case where h_j^2 was reduced. For an economic weight of $w_{-i} = 0$, the reduction of h_j^2 resulted in a decrease of σ_I by 60% for index II but only by

51% for index III. Up to $w_i = 0.65$ the genomic index ($N_p = 1'000$) resulted in higher σ_I than the full sib index. The advantage of index III with $N_p = 1'000$ decreased, the more economic weighting was shifted towards the trait with the higher heritability. To further analyze this matter, we also compared the accuracies of indices II (7 fullsibs) and III ($N_p = 1'000$) for uncorrelated traits, which is shown in Figure 5B. As the accuracy of an index cannot drop below the accuracy of the GBV used as information source (König and Swalve, 2009), $r_{G\hat{Q}}$ of 0.516 was a minimum value for r_{TI} of index III. In contrast, the impact of low h_j^2 on r_{TI} was more direct when using the conventional full sib index, which led to a higher sensitivity to the reduction of h_j^2 . Increasing the relative economic weighting of the low heritability trait in the conventional index with 7 full sibs led to substantial losses in r_{TI} . In contrast, r_{TI} even rose when relative economic weighting of the lowly heritable trait was increased in index III with $N_p = 1'000$. However, as σ_I better accounts for the economic variation in the aggregate breeding goal than r_{TI} , comparison of σ_I shows more proportionality to expected response to selection. We also compared response to selection for the trait with low heritability, j , when σ_I of index II and III were the same assuming selection intensity of 1.4 and σ_G of 0.32. At this point of equal σ_I and w_i for both indices, the response to selection for the trait with low heritability was greater (0.081) using index III ($N_p = 1'000$) than using index II with 7 full sibs (0.053). Similar findings were also reported by Hayes et al. (2009), König and Swalve (2009) and Buch et al. (2012). Both findings are due to the fact that lower h_j^2 caused a decrease in σ_I , but this effect was less pronounced in index III.

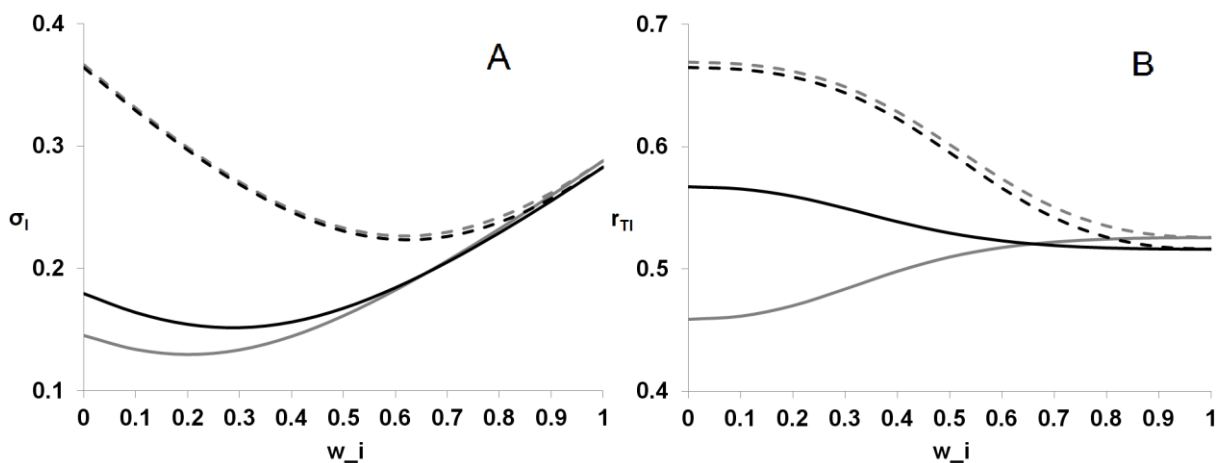


Figure 5 (A, B) Standard deviations (σ_I ; A) and accuracies (r_{TI} ; B) of index II with 7 full sibs (grey) and of index III with $N_p = 1'000$ (black) as a function of relative economic weighting of trait i (w_i) for uncorrelated traits i and j . Dashed lines: equal heritabilities of 0.3; solid lines: heritability of trait j reduced to 0.1.

Reducing r_j^2 from 0.49 to 0.20 resulted in comparable σ_I of index II (7 full sibs) and III ($N_p = 1'000$) for $r_{G_i G_j} = 0$ (results not shown). Maximum σ_I of index III remained higher than σ_I of index II. Consequently, the standard deviation of the genomic index with a reference population of 1'000 animals remained in a comparable range with σ_I of the full sib index (7 full sibs) even if accuracy of the GBV for trait j was low.

The effect of the heritability on σ_I was also assessed by varying h_j^2 from 0.1 to 1. If more information was incorporated into the index, σ_I became increasingly independent of the heritability of trait j . Figure 6 shows σ_I of indices I, II (7 full sibs) and III ($N_p = 1'000$) for an equal distribution of economic weights ($w_i = 0.5$). Only for a strongly negative genetic correlation ($r_{G_i G_j} = -1$) we observed almost identical σ_I of indices I, II and III (Figure 6C).

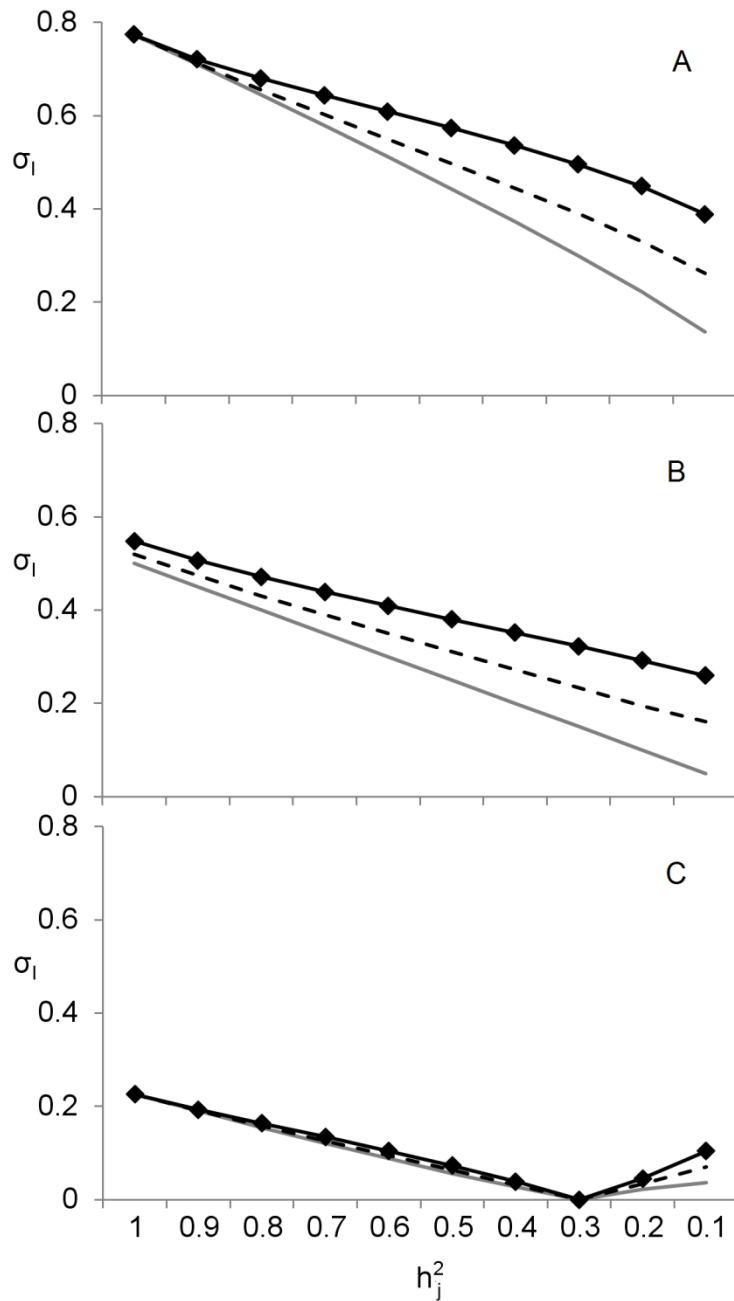


Figure 6 (A, B, C) Standard deviations (σ_1) of index I (grey), index II with 7 full sibs (dashed), and index III with $N_p = 10'000$ (diamonds) for a genetic correlation ($r_{G_iG_j}$) between traits i and j of 1 (A), 0 (B) and -1 (C) as a function of heritability of trait j for equal economic weighting.

Finally we considered the case that the correlation of the prediction errors of the genomic component of the breeding value is inflated by adding a constant Δ , so that eq. (4) of the Appendix becomes

$$r_{R_i R_j}^{\Delta} = \sqrt{1 - r_{Q_i \hat{Q}_i}^2} q_i r_{G_i G_j} q_j \sqrt{1 - r_{Q_j \hat{Q}_j}^2} + \Delta.$$

Solving eq. (6) of the Appendix under this assumption, the correlation between the GBVs of traits i and j becomes

$$r_{\hat{Q}_i \hat{Q}_j}^{\Delta} = r_{G_i G_j} q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j} + \frac{\sqrt{1 - r_{Q_i \hat{Q}_i}^2}}{r_{Q_i \hat{Q}_i}} \frac{\sqrt{1 - r_{Q_j \hat{Q}_j}^2}}{r_{Q_j \hat{Q}_j}} \Delta.$$

The first part of this equation is identical with our original result, and since the factor for Δ is by definition non-negative, $r_{\hat{Q}_i \hat{Q}_j}^{\Delta}$ is inflated (deflated) relative to $r_{\hat{Q}_i \hat{Q}_j}$ if Δ is positive (negative).

To assess the effect of including Δ on σ_i and hence the expected genetic progress, we consider the case where the accuracy of the quasi-phenotypes, $r_{Q \hat{Q}}$, is 0.573 for both traits with $N_p = 1'000$ and $q = 0.9$ for both traits. In this case

$$r_{\hat{Q}_i \hat{Q}_j}^{\Delta} = r_{G_i G_j} \times 0.266 + 1.297 \times \Delta.$$

If we further assume that $r_{G_i G_j} = 0.5$ and $\Delta = 0.2$, the correlation of the prediction errors when including Δ is inflated ($r_{R_i R_j}^{\Delta} = 0.373$ vs. $r_{R_i R_j} = 0.173$), and the same is true for the correlation between the GBVs ($r_{\hat{Q}_i \hat{Q}_j}^{\Delta} = 0.393$ vs. $r_{\hat{Q}_i \hat{Q}_j} = 0.133$). For an equal relative economic weight of 0.5 for both traits, the standard deviation of the index, σ_i^{Δ} , is now smaller (0.304) than without inclusion of Δ ($\sigma_i = 0.320$), which means that an inflated correlation of the prediction errors due to external factors will decrease the expected genetic progress. This also means that not accounting for the impact of correlated residuals due to external effects will lead to an overestimation of σ_i in the discussed scenario.

The direction of the effect of including a biased correlation of prediction errors is depending on three factors, namely the signs of Δ , $r_{G_i G_j}$, and w_i (when assuming $w_j > 0$). This results in eight different combinations which were assessed with the parameters of the assumed scenario, and results are given in Table 1.

Table 1 Correlations of prediction errors and standard deviations of index III if not accounting for Δ ($r_{R_i R_j}; \sigma_I$) and if accounting for Δ ($r_{R_i R_j}^\Delta; \sigma_I^\Delta$) for different combinations of Δ , genetic correlation between the traits ($r_{G_i G_j}$) and relative economic weighting of trait i (w_{-i})

	Combination of $\Delta; r_{G_i G_j}; w_{-i}$	$r_{R_i R_j}$	$r_{R_i R_j}^\Delta$	σ_I	σ_I^Δ
(1)	0.2; 0.5; 0.5	0.173	0.373	0.320	0.304
(2)	-0.2; 0.5; 0.5	0.173	-0.027	0.320	0.346
(3)	0.2; -0.5; 0.5	-0.173	0.027	0.126	0.117
(4)	-0.2; -0.5; 0.5	-0.173	-0.373	0.126	0.142
(5)	0.2; 0.5; -0.5	0.173	0.373	0.126	0.142
(6)	-0.2; 0.5; -0.5	0.173	-0.027	0.126	0.117
(7)	0.2; -0.5; -0.5	-0.173	0.027	0.320	0.346
(8)	-0.2; -0.5; -0.5	-0.173	-0.373	0.320	0.304

Combination (1) and (2) reflect the case where both $r_{G_i G_j}$ and w_{-i} are positive. Combination (1) with a positive Δ is described above. In combination (2), the negative Δ implies a weaker correlation of prediction errors, $r_{R_i R_j}^\Delta$, than if Δ is not accounted for ($r_{R_i R_j}$). This leads to an underestimation of σ_I . In combination (7) and (8) the signs of both $r_{G_i G_j}$ and w_{-i} are negative. Again, as in combination (1), a stronger $r_{R_i R_j}^\Delta$ compared to $r_{R_i R_j}$ implies an overestimation of σ_I . If $r_{R_i R_j}^\Delta$ is weaker than when neglecting Δ , as in combination (2), σ_I is underestimated.

In the cases where the signs of $r_{G_i G_j}$ and w_{-i} are antagonistic (combinations (3) to (6)), a weaker $r_{R_i R_j}^\Delta$ implies an overestimation of σ_I and a stronger $r_{R_i R_j}^\Delta$ implies an underestimation of σ_I .

Conclusions

As the amount of information increases, the standard deviation of an index, determining the expected genetic gain, is increasingly independent of genetic correlation, heritability and relative economic weighting among traits. This applies for both conventional and genomic information, e.g., more full sibs or GBVs with higher accuracies.

The standard deviation of a conventional full sib index is more sensitive to a reduction of heritability than a genomic index. For a two-trait scenario where the economic importance is the same for both traits, the genetic gain of the trait with low heritability can therefore be

higher than genetic gain of the trait with a higher heritability when applying the genomic index rather than the full sib index.

Based on our results, a reference population of 1'000 animals is a minimum requirement for GS to be competitive with a conventional selection scheme reflecting the practical situation in pig breeding (e.g. parent information, performance of 5 full sibs and 80 half sibs). The standard deviation of a genomic index with a reference population of 5'000 animals cannot be achieved by any realistic conventional index. Our conclusions also hold for the comparison of accuracies of the indices.

Genomic breeding values for several traits might be correlated due to an overlap in the training sets used to derive the GBVs. We have shown that this will have an effect on expected genetic progress, and how this can be quantified as a function of a parameter Δ reflecting the deviation of the correlation of prediction errors of GBVs from its expectation. The actual magnitude of Δ remains to be quantified in empirical studies.

Acknowledgements

We would like to thank the German ministry of education and research for the financial support of the program FUGATO+brain (0315136), which was also supported by the Förderverein Biotechnologieforschung e.V., Lohmann Tierzucht and five German horse breeding associations (Hanoverian, Trakehner, Oldenburger, Holsteiner and Westfalian horses). Furthermore, this research was financed by the European Union and the federal state of North Rhine-Westphalia within the project pigGS (005-NA02-018 C). We wish to thank M. Tietze as well as two anonymous reviewers for their helpful comments. We especially thank Peter Amer for challenging us to think more about the possible dependency of GBVs in multiple trait scenarios.

Investing in **your** future!



References

- Buch LH, Sørensen MK, Berg P, Pedersen LD and Sørensen AC 2012. Genomic selection strategies in dairy cattle: Strong positive interaction between use of genotypic information and intensive use of young bulls on genetic gain. *J Anim Breed Genet* 129, 138-151.
- 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. *J Anim Breed Genet* 124, 331-341.
- Erbe M, Reinhardt F and Simianer H 2011. Empirical determination of the number of independent chromosome segments based on cross-validated data. In: Proc. of 62th Annu. Mtg. Eur. Assoc. Anim. Prod., Stavanger, Norway, August 29 - September 2, 2011. Wageningen Academic Publishers, The Netherlands.
- Goddard ME, Hayes BJ, Meuwissen THE 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J Anim Breed Genet* 128, 409-421.
- Haberland AM, Ytournal F, Luther H and Simianer H 2010. Evaluation of selection strategies including genomic breeding values in pigs. In: Proc. of 61th Annu. Mtg. Eur. Assoc. Anim. Prod., Heraklion, Greece, August 23-27, 2010. Wageningen Academic Publishers, The Netherlands.
- Haberland AM, König von Borstel U, Simianer H and König S 2012. Integration of genomic information into sport horse breeding programs for optimization of accuracy of selection. *Animal* 6, 1369-1376.
- Hansen Axelsson H, Johansson K, Eriksson S, Petersson K-J, Rydhmer L and Philipsson J 2011. Selection of bull dams for production and functional traits in an open nucleus herd. *J Dairy Sci* 94, 2592-2600.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci* 92, 433-443.
- Hazel LN 1943. The genetic basis for constructing selection indexes. *Genetics* 28, 476-490.
- König S, Simianer H and Willam A 2009. Economic evaluation of genomic breeding programs. *J Dairy Sci* 92, 382-391.
- König S and Swalve HH 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. *J Dairy Sci* 92, 5292-5303.
- Meuwissen THE, Hayes BJ and Goddard ME 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Pimentel ECG and König S 2012. Genomic selection for the improvement of meat quality in beef. *J Anim Sci* 90, 3418-3426.

- Rohrer GA, Alexander LJ, Hu Z, Smith TP, Keele JW and Beattie CW 1996. A comprehensive map of the porcine genome. *Genome Res* 6, 371-391.
- Schaeffer LR 2006. Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet* 123, 218-223.
- Simianer H 2009. The potential of genomic selection to improve litter size in pig breeding programs. In: Proc. of 60th Annu. Mtg. Eur. Assoc. Anim. Prod., Barcelona, Spain, August 24-27, 2009. Wageningen Academic Publishers, The Netherlands.
- Smith C 1983. Effects of changes in economic weights on the efficiency of index selection. *J Anim Sci* 56, 1057-1064.
- Van Eenennaam AL, van der Werf JHJ and Goddard ME 2011. The value of using DNA markers for beef bull selection in the seedstock sector. *J Anim Sci* 89, 307-320.

Appendix

According to Figure 1, we have for instance for trait i : $G_i = H_i + \hat{Q}_i + R_i$. From this, the following correlations can be derived

$$r_{G_i H_j} = r_{G_i G_j} \sqrt{1 - q_j^2} \quad (1)$$

$$r_{G_i R_j} = r_{G_i G_j} q_j \sqrt{1 - r_{\hat{Q}_i \hat{Q}_j}^2} \quad (2)$$

$$r_{H_i H_j} = \sqrt{1 - q_i^2} \sqrt{1 - q_j^2} r_{G_i G_j} \quad (3)$$

$$r_{R_i R_j} = \sqrt{1 - r_{\hat{Q}_i \hat{Q}_j}^2} q_i r_{G_i G_j} q_j \sqrt{1 - r_{\hat{Q}_i \hat{Q}_j}^2} \quad (4)$$

$$r_{R_i H_j} = \sqrt{1 - r_{\hat{Q}_i \hat{Q}_j}^2} q_i r_{G_i G_j} \sqrt{1 - q_j^2} \quad (5)$$

Furthermore, one can express the covariance between \hat{Q}_i and \hat{Q}_j as:

$$\begin{aligned} \text{cov}(\hat{Q}_i, \hat{Q}_j) &= \text{cov}(G_i - H_i - R_i, G_j - H_j - R_j) \\ \text{cov}(\hat{Q}_i, \hat{Q}_j) &= \text{cov}(G_i, G_j) - \text{cov}(G_i, H_j) - \text{cov}(G_i, R_j) - \text{cov}(H_i, G_j) - \text{cov}(H_i, H_j) - \\ &\quad - \text{cov}(H_i, R_j) - \text{cov}(R_i, G_j) - \text{cov}(R_i, H_j) - \text{cov}(R_i, R_j) \end{aligned}$$

One thus has:

$$\begin{aligned} r_{\hat{Q}_i \hat{Q}_j} &= \frac{1}{\sigma_{\hat{Q}_i} \sigma_{\hat{Q}_j}} \left(r_{G_i G_j} \sigma_{G_i} \sigma_{G_j} - r_{G_i H_j} \sigma_{G_i} \sigma_{H_j} - r_{G_i R_j} \sigma_{G_i} \sigma_{R_j} - r_{H_i G_j} \sigma_{H_i} \sigma_{G_j} + r_{H_i H_j} \sigma_{H_i} \sigma_{H_j} \right. \\ &\quad \left. + r_{H_i R_j} \sigma_{H_i} \sigma_{R_j} - r_{R_i G_j} \sigma_{R_i} \sigma_{G_j} + r_{R_i H_j} \sigma_{R_i} \sigma_{H_j} + r_{R_i R_j} \sigma_{R_i} \sigma_{R_j} \right) \\ r_{\hat{Q}_i \hat{Q}_j} &= r_{G_i G_j} \frac{\sigma_{G_i} \sigma_{G_j}}{\sigma_{\hat{Q}_i} \sigma_{\hat{Q}_j}} - r_{G_i H_j} \frac{\sigma_{G_i} \sigma_{H_j} \sigma_{G_j}}{\sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} - r_{G_i R_j} \frac{\sigma_{G_i} \sigma_{R_j} \sigma_{G_j}}{\sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} - r_{H_i G_j} \frac{\sigma_{H_i} \sigma_{G_i} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{\hat{Q}_j}} + r_{H_i H_j} \frac{\sigma_{H_i} \sigma_{G_i} \sigma_{H_j} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} \\ &\quad + r_{H_i R_j} \frac{\sigma_{H_i} \sigma_{G_i} \sigma_{R_j} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} - r_{R_i G_j} \frac{\sigma_{R_i} \sigma_{G_i} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{\hat{Q}_j}} + r_{R_i H_j} \frac{\sigma_{R_i} \sigma_{G_i} \sigma_{H_j} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} + r_{R_i R_j} \frac{\sigma_{R_i} \sigma_{G_i} \sigma_{R_j} \sigma_{G_j}}{\sigma_{G_i} \sigma_{\hat{Q}_i} \sigma_{G_j} \sigma_{\hat{Q}_j}} \end{aligned} \quad (6)$$

Using a regression, one can derive that

$$\begin{aligned} q_i r_{\hat{Q}_i \hat{Q}_j} &= \frac{\sigma_{\hat{Q}_i}}{\sigma_{G_i}} \\ q_i \sqrt{1 - r_{\hat{Q}_i \hat{Q}_j}^2} &= \frac{\sigma_{R_i}}{\sigma_{G_i}} \\ \sqrt{1 - q_i^2} &= \frac{\sigma_{H_i}}{\sigma_{G_i}} \end{aligned}$$

Including these equations and the equations (1) to (5) in (6) provides:

$$r_{\hat{Q}_i \hat{Q}_j} = r_{G_i G_j} \left(\frac{1}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} - \frac{1 - q_j^2}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} - \frac{q_j^2 (1 - r_{Q_j \hat{Q}_j}^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} - \frac{1 - q_i^2}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} + \frac{(-q_i^2)(-q_j^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} \right. \\ \left. + \frac{q_j^2 (-q_i^2) (1 - r_{Q_j \hat{Q}_j}^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} - \frac{q_i^2 (-r_{Q_i \hat{Q}_i}^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} + \frac{q_i^2 (-r_{Q_i \hat{Q}_i}^2)(-q_j^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} + \frac{q_i^2 (-r_{Q_i \hat{Q}_i}^2) q_j^2 (1 - r_{Q_j \hat{Q}_j}^2)}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} \right)$$

$$r_{\hat{Q}_i \hat{Q}_j} = \frac{r_{G_i G_j}}{q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}} \left(1 - (-q_j^2) - q_j^2 (1 - r_{Q_j \hat{Q}_j}^2) - (-q_i^2) + (-q_i^2)(-q_j^2) \right. \\ \left. + q_j^2 (-q_i^2) (1 - r_{Q_j \hat{Q}_j}^2) - q_i^2 (-r_{Q_i \hat{Q}_i}^2) + q_i^2 (-r_{Q_i \hat{Q}_i}^2)(-q_j^2) + q_i^2 (-r_{Q_i \hat{Q}_i}^2) q_j^2 (1 - r_{Q_j \hat{Q}_j}^2) \right)$$

$$r_{\hat{Q}_i \hat{Q}_j} = r_{G_i G_j} q_i r_{Q_i \hat{Q}_i} q_j r_{Q_j \hat{Q}_j}$$

4TH CHAPTER**Efficiency of different selection strategies against boar taint in pigs**

A.M. Haberland¹, H. Luther², A. Hofer², E. Tholen³, H. Simianer¹ B. Lind⁴ and C.
Baes^{2,5}

¹Department of Animal Sciences, Georg-August-University Goettingen, Goettingen, Germany

²SUISAG, Sempach, Switzerland

³Department of Animal Breeding, Rheinische Friedrich-Wilhelms-University Bonn, Bonn, Germany

⁴Förderverein Biotechnologieforschung e.V., Bonn, Germany

⁵School of Agricultural, Forest and Food Sciences, Bern University of Applied Sciences, Zollikofen, Switzerland

Published online in Animal

on 1st of November 2013

© 2013 The Animal Consortium

doi: 10.1017/S1751731113001857

Abstract

The breeding scheme of a Swiss sire line was modeled to compare different target traits and information sources for the selection against boar taint. The impact of selection against boar taint on production traits was assessed for different economic weights of boar taint compounds. Genetic gain and breeding costs were evaluated using ZPlan+, a software based on selection index theory, gene flow method and economic modeling. Scenario I reflected the currently practiced breeding strategy as a reference scenario without selection against boar taint. Scenario II incorporated selection against the chemical compounds of boar taint, androstenone (AND), skatole (SKA) and indole (IND) with economic weights of -2.74, -1.69 and -0.99 Euro per unit of the log transformed trait, respectively. As information sources, biopsy-based performance testing of live boars (BPT) was compared with genomic selection (GS) and a combination of both. Scenario III included selection against the subjectively assessed human nose score (HNS) of boar taint. Information sources were either station testing of full and half sibs of the selection candidate or GS against HNS of boar taint compounds. In scenario I, annual genetic gain of log transformed AND (SKA; IND) was 0.06 (0.09; 0.02) Euro, which was due to favorable genetic correlations with lean meat percentage and meat surface. In scenario II, genetic gain increased to 0.28 (0.20; 0.09) Euro per year when conducting BPT. Compared with BPT, genetic gain was smaller with GS. A combination of BPT and GS only marginally increased annual genetic gain, whereas variable costs per selection candidate augmented from 230 Euro (BPT) to 330 Euro (GS) or 380 Euro (both). The potential of GS was found to be higher when selecting against HNS, which has a low heritability. Annual genetic gain from GS was higher than from station testing of 4 full sibs and 76 half sibs with one or two measurements. The most effective strategy to reduce HNS was selecting against chemical compounds by conducting BPT. Because of heritabilities higher than 0.45 for AND, SKA and IND and high genetic correlations to HNS, the (correlated) response in units of the trait could be increased by 62% compared with scenario III with GS and even by 79% compared with scenario III, with station testing of siblings with two measurements. Increasing the economic weights of boar taint compounds amplified negative effects on average daily gain, drip loss and intramuscular fat percentage.

Keywords: boar taint, biopsy, androstenone, genomic selection, skatole

Implications

Because the European pig producers agreed to omit the practice of surgical castration by 2018, new strategies for reducing the amount of tainted carcasses are required. Using the practical breeding program of a Swiss sire line as a reference, we compare different approaches to reduce the occurrence of boar taint by means of selection. Selection against the chemical compounds of boar taint as measured in liquid fat is compared with selection against boar taint as measured by test persons in the abattoir. The efficiency of different information sources, including genomic information, is evaluated in terms of genetic gain and breeding costs.

Introduction

Until recently, surgical castration as a reliable means for producing meat free of boar taint has been a common practice in pig production in many European countries. To improve animal welfare (von Borell *et al.*, 2009), the European pig industry collectively and voluntarily agreed to discontinue surgical castration of piglets by 2018 (European Commission, 2010). A ban on surgical castration, including that performed using anaesthesia or analgesia, will likely be anchored in the legislation of many European countries in the foreseeable future; feasible alternatives are required as soon as possible.

Alternatives to surgical castration have been the topic of intense research in Europe. Three main possibilities exist: (1) sexing semen, which would allow the production of only female animals, totally circumventing the problem of boar taint, (2) immunocastration, involving the immunization of young pigs against gonadotropin-releasing hormone (GnRH) (Prunier *et al.*, 2006; Fàbrega *et al.*, 2010; Rydhmer *et al.*, 2010); and (3) raising intact boars. While common in cattle breeding, sexed swine semen is not likely to become available on a commercial scale in the near future because of various technical limitations (Vazquez *et al.*, 2009). Although registered in most of Europe, immunocastration is not widely used, because of image concerns of retailers. Breeding against the main compounds of boar taint (skatole, androstenone and indole) seems promising because of high consumer acceptance, favorable effects on various production traits, high heritabilities and a more efficient food conversion of intact male boars (Walstra, 1974; Sellier *et al.*, 2000; Windig *et al.*, 2012). Before intact male boars can be produced on a large scale, however, the frequency of tainted carcasses must be reduced and a reliable means of identifying carcasses with organoleptic anomalies must be implemented. Management practices adapted to rearing intact boars (i.e. feeding regimes, housing facilities, etc.) will also be necessary.

Incorporating selection against boar taint into practical breeding programs requires a reliable system for recording the target traits. Those can be either the amount of boar taint compounds, for example, in liquid fat, or the human nose score (HNS) being the intensity of odor as perceived by trained test individuals (Windig *et al.*, 2012). Levels of chemical boar taint compounds can be measured either in the abattoir, for example, in siblings of the selection candidate, or by conducting a biopsy-based performance test in live boars, as proposed by Baes *et al.* (2013). Assessing the HNS requires a trained panel of testers (Mathur *et al.*, 2012).

Accuracy of selection and therefore response to selection may be improved by additionally considering genomic information. The gain in accuracy will depend on whether boar taint compounds or HNS are considered in the breeding goal owing to the considerable differences in heritability (Windig *et al.*, 2012). Genomic selection (GS) is defined as the estimation of breeding values based on genome-wide dense marker maps (Meuwissen *et al.*, 2001). The development of a 60K SNP array for *Sus scrofa* (Ramos *et al.*, 2009) enables a routine assessment of a large number of markers that, in addition to conventional pedigree-based information, should help to partition the genetic variance observed in the population. Estimation of linkage disequilibrium (LD) carried out by Uimari and Tapio (2011) and Badke *et al.* (2012) showed high values of r^2 between adjacent SNPs in pigs; these r^2 values were comparative with those in North American Holstein cattle, indicating that the estimation of accurate genomic breeding values (GBVs) for pigs should be feasible using a 60K SNP array. Accuracies of GBVs for traits with low heritability (female reproduction traits) were found to be clearly higher than the accuracy of conventional information normally available at the time of selection (Cleveland *et al.*, 2010). Next to LD, the number of animals in the reference population is an important factor determining the accuracy of GBVs. Haberland *et al.* (2013) estimated a lower limit of about 1'000 animals to increase genetic gain of a pig breeding program using GS.

The aim of this study was to model a terminal sire line breeding program to assess the potential of selection against boar taint as reflected in different target traits (HNS or chemical compounds) using selection index theory. The Swiss terminal sire line PREMO[®] was used as an example for comparing different information sources: (i) biopsy-based performance testing (BPT) of live boars; (ii) assessment of HNS on station; and (iii) GS against either chemical compounds or HNS. The economic weights of boar taint components were varied to assess the effects on monetary genetic gain of production traits, and on time needed to reduce boar taint within the examined pig population.

Material and methods

Within the three-way crossbreeding scheme of the Swiss pig production company SUISAG, the terminal sire line PREMO[®] is mated to F1 crossbreed sows (Swiss Large White x Swiss Landrace). In this study, we focus on the selection scheme of the sire line. Genetically, the breed originates from a Swiss Large White line and has been selected for high fattening performance and meat quality for about 10 years. Because of the low average androstenone content in PREMO[®] boars compared with other breeds such as Duroc, Landrace or Large White (Grindflek *et al.*, 2011; Windig *et al.*, 2012; Baes *et al.*, 2013), the use of this terminal sire line in a breeding program should provide a good starting point for reducing the number of carcasses with organoleptic anomalies. Heritabilities, phenotypic standard deviations and economic weights of the breeding goal traits in the current population are given in Table 1; phenotypic and genetic correlations are shown in Table 2.

Table 1. Heritabilities (h^2), phenotypic standard deviations (σ_P) and economic weights (w) per unit of considered traits (SUISAG, 2012)

<i>Trait</i>	h^2	σ_P	w	<i>unit</i>
ADG_S	0.27	85.33	0.05	g/day
FCR	0.35	0.16	-40.00	kg/kg
SUR	0.61	4.08	0.7	cm ²
IMF	0.60	0.53	9.25	%
pH ₁	0.17	0.19	20.00	pH
PIGM	0.27	0.17	12.00	score
DL	0.30	1.71	-3.30	%
ADG_F	0.29	40.77	0.06	g/day
BFT	0.40	2.46	-	cm
ADG_SI	0.37	48.23	0.12	g/day
LMP	0.34	2.45	1.65	%
AND	0.45	0.95	-2.74	ln(μ g/g liquid fat)
SKA	0.49	0.73	-1.69	ln(μ g/g liquid fat)
IND	0.55	0.59	-0.99	ln(μ g/g liquid fat)
HNS	0.12	0.95	-2.93	score

ADG_S= average daily gain (station test), FCR= feed conversion ratio, SUR= surface, IMF= percentage of intramuscular fat, pH₁= acidity 1 hour after slaughtering), PIGM= pigmentation, DL= drip loss, ADG_F= average daily gain (field test), BFT= backfat thickness, ADG_SI= average daily gain (at slaughtering), LMP= lean meat percentage, AND= androstenone, SKA= skatole, IND= indole.

Table 2. Heritabilities (diagonal), phenotypic (above diagonal) and genotypic (below diagonal) correlations between considered traits (SUISAG, 2012; Frieden (personal communication), 2013; Windig *et al.*, 2012)

	ADG_S	FCR	SUR	IMF	pH ₁	PIGM	DL	ADG_F	BFT	ADG_SI	LMP	AND	SKA	IND	HNS
ADG_S	0.27	-0.45	-0.09	0.06	0.04	-0.02	0.03	0.41	0.13	0.19	-0.06	Na	Na	Na	Na
FCR	-0.32	0.35	-0.13	0.28	-0.01	0.04	-0.12	-0.10	0.23	-0.11	-0.29	0.13	0.14	0.16	Na
SUR	-0.05	-0.12	0.61	-0.12	-0.04	-0.10	0.08	-0.08	-0.20	0.00	0.25	-0.23	-0.16	-0.20	Na
IMF	0.06	0.37	-0.07	0.60	0.02	-0.03	-0.21	-0.01	0.22	-0.04	-0.23	0.19	-0.04	0.14	Na
pH ₁	-0.06	0.07	-0.13	0.23	0.17	-0.02	-0.47	0.01	0.03	0.04	-0.02	Na	Na	Na	Na
PIGM	-0.13	0.05	-0.03	-0.11	0.02	0.27	-0.01	-0.03	0.06	-0.06	-0.04	Na	Na	Na	Na
DL	0.22	-0.32	0.27	-0.50	-0.59	-0.13	0.30	0.05	-0.08	-0.06	0.08	-0.05	0.06	-0.10	Na
ADG_F	0.48	-0.13	-0.22	0.02	0.00	0.01	0.16	0.29	0.16	0.15	-0.08	0.19	-0.05	0.02	Na
BFT	-0.13	0.56	-0.13	0.31	0.05	0.12	-0.24	0.10	0.40	-0.09	-0.43	0.27	0.01	0.15	Na
ADG_SI	0.57	-0.33	0.02	-0.06	0.23	-0.14	-0.13	0.42	-0.17	0.37	-0.03	Na	Na	Na	Na
LMP	-0.05	-0.51	0.28	-0.27	-0.07	-0.09	0.18	-0.16	-0.81	0.02	0.34	-0.22	-0.12	-0.21	Na
AND	Na	0.13	-0.23	0.19	Na	Na	-0.05	0.19	0.27	Na	-0.22	0.45	0.28	0.26	0.27
SKA	Na	0.14	-0.16	-0.04	Na	Na	0.06	-0.05	0.01	Na	-0.12	0.11	0.49	0.74	0.36
IND	Na	0.16	-0.20	0.14	Na	Na	-0.10	0.06	0.15	Na	-0.21	0.35	0.90	0.55	0.32
HNS	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	0.65	0.90	0.84	0.12

ADG_S= average daily gain (station test), FCR= feed conversion ratio, SUR= meat surface, IMF= percentage of intramuscular fat, pH₁= acidity 1 hour after slaughtering), PIGM= pigmentation, DL= drip loss, ADG_F= average daily gain (field test), BFT= backfat thickness, ADG_SI= average daily gain (at slaughtering), LMP= lean meat percentage, AND= androstenone in liquid fat, SKA= skatole in liquid fat, IND= indole in liquid fat, HNS= human nose score, Na= not available

The breeding program was modeled using ZPlan+ (Täubert *et al.*, 2010). This software combines selection index theory (Hazel, 1943), gene flow method (Elsen and Mocquot, 1974; Hill, 1974) and economic modeling, enabling deterministic simulation of livestock breeding programs (Willam *et al.*, 2008). Breeding schemes can be compared in terms of generation interval, monetary genetic gain, breeding costs, returns and discounted profit. The selection index is implemented in ZPlan+ as described by Hazel (1943).

In the genomic scenarios, GBVs were integrated into the selection index considering them as auxiliary traits with a heritability of 1, as proposed by Dekkers (2007). Phenotypic and genetic correlations between these ‘genomic traits’ and the traits of the breeding goal were calculated in accordance with Dekkers (2007). Only our formula to determine the genetic correlation between GBVs of two different traits differs from Dekkers (2007) in that it was derived assuming the proportion of genetic variance associated with markers (q^2) not being necessarily identical (Haberland *et al.*, 2013). We adopted the value of $q=0.9$ suggested by Erbe *et al.* (2011) who used cross-validated data to empirically determine q for genotyped Holstein Friesian bulls. To the best of our knowledge, such empirical data are not yet available for pigs. The accuracy of the GBVs r_{GBV} was calculated using a formula derived by Daetwyler *et al.* (2010):

$$r_{GBV} = \sqrt{\frac{N_p r^2}{N_p r^2 + M_e}}$$

where N_p is the number of individuals in the reference population. In our calculations, we assumed $N_p = 1,000$, which may be assumed a minimum for GS in pigs (Haberland *et al.*, 2013). For the reliability of the quasi-phenotypes, that is, of the conventional EBVs of the animals in the reference population, we assumed $r^2 = 0.49$ for all traits. M_e is the effective number of chromosome segments segregating in the population and can be approximated with $2N_e Lk / \log(N_e L)$ as proposed by Goddard *et al.* (2011). N_e denotes the effective population size, L is the average length of a chromosome in Morgan and k is the number of chromosome pairs. Assuming $N_e = 100$, $k = 19$ and $L = 1.2$ Morgan (with length of the porcine genome being 23 Morgan, Rohrer *et al.*, 1996), the value of M_e was $\sim 1'000$.

The following scenarios were compared in terms of annual genetic gain of log-transformed (ln) boar taint components AND, SKA and IND and in terms of variable breeding costs per selection candidate. To correct for skewness, boar taint phenotypes AND, SKA and IND were log transformed to achieve a normal distribution of the data (Baes *et al.*, 2013). The monetary genetic gain per year was calculated as $\Delta G / a = i r_{TI} \sigma_T / \Delta T$, where i is the selection intensity, r_{TI} is the accuracy of the index, σ_T is the standard deviation of the breed-

ing goal and ΔT is the generation interval. The expected change in the amount of boar taint prevalence per year was estimated using the average amount of boar taint compounds in the current PREMO[®] population (0.70, 0.03 and 0.03 $\mu\text{g/g}$ liquid fat for AND, SKA and IND, respectively) as a starting point. The selection was intensified by increasing the economic weighting factors for boar taint compounds in three steps. The scenarios were also assessed with respect to the impact on production traits, provided that the genetic correlation between production traits and boar taint compounds was known. Table 3 shows an overview over the scenarios and the respective information sources.

Table 3. Breeding goals and information sources of the different scenarios

	Breeding goal	Information sources			Station testing
		BPT	GS	BPT+GS	
Scenario I	no selection against boar taint (reference scheme)				
Scenario II	chemical compounds (AND, SKA, IND)	a)	b)	c)	
Scenario III	HNS		b)		a)

BPT= biopsy-based performance testing; GS= genomic selection; AND= androstenone; SKA= skatole; IND= indole

Scenario I, conservative scheme (base scenario)

This base scenario models the current breeding program. The breeding nucleus consists of 270 sows with an annual replacement rate of 75%. The breeding sows are mated to 60 AI boars, 35% of which are progeny tested. The rather large number of young boars is maintained to control inbreeding within the small breeding nucleus and to increase genetic gain more rapidly. Genetic gain is transferred to the production units by a larger pool of 150 AI boars, which is assumed to be selected with a lower intensity than the breeding boars, but in which the breeding boars are included. In ZPlan+, we split the breeding sows and breeding boars into two groups according to two selection steps. The first selection step is based on field performance testing of 1,200 male and 1,200 female selection candidates per year at a live weight of 100 to 130 kg. A total of 200 young breeding sows and 42 young boars are selected according to their own and 60 half sib performances in the traits average daily gain and backfat thickness (measured using ultra sound). In addition, two full sibs and 12 half sibs of every selection candidate are tested on station for average daily gain, feed conversion ratio, intramuscular fat, pH₁, pigmentation, drip loss and lean meat percentage. The productive lifetime of the young breeding animals selected in the first selection step is 1 year. Field performance testing was assumed to cost 180 Euro. In the second selection step, 70 sows and 20 boars are selected to be kept for another two years according to their progeny re-

records. Progeny testing is carried out by testing six purebred progeny on station, and by recording about 40 crossbred end-products for lean meat percentage and average daily gain. Progeny testing was assumed to cost 1535 Euro. The larger pool of boars used for matings within the production unit is also split into two groups, namely, 105 younger and 42 older boars with a productive lifetime of 1 or 2 years, respectively. Hence, including the production unit consisting of 60,000 crossbred sows, there are seven groups involved in the breeding program modeled in ZPlan+. The transmission matrix (gene flow) within the modeled population is shown in Table 4.

Fixed costs of the breeding program were not accounted for because of the complexity of their determination and because only variable costs have an impact on the efficiency of the breeding strategy. Boar taint compounds were included in scenario I with an economic weighting of zero; this was done to assess the correlated response because of their correlations with production traits.

Scenario II, breeding against boar taint compounds

Log-transformed boar taint compounds AND, SKA and IND were included in the breeding goal. Because no genetic correlation between boar taint compounds and fattening traits were available, these relationships were partly adopted from the German Piétrain Herdbook Organisations (Tholen, personal communication). The underlying data set of these parameters comprises information from 1,010 stations tested, Piétrain-sired commercial crossbreds (Tholen *et al.*, 2011). AND and IND show favorable genetic correlations to lean meat percentage, meat surface and feed conversion ratio (Table 2). Undesirable correlations exist between AND and average daily gain as well as between intramuscular fat percentage and the boar taint compounds AND and IND.

Three kinds of information sources for the selection index were compared: BPT in live boars (scenario IIa), GS (scenario IIb) or a combination of both (scenario IIc).

(IIa) A biopsy was assumed to be taken from the neck region of 1,200 live male selection candidates during the field test (Baes *et al.*, 2013). Thus, boar taint compounds could be quantified in addition to the currently measured traits average daily gain and backfat thickness. Each selection candidate had information on boar taint compounds from itself (only if male), its sire and its 30 male half sibs. The regular costs of the field test (180 Euro) and the costs for biopsy and analysis (50 Euro) added up to 230 Euro per animal. Expected long-term change in boar taint prevalence was calculated assuming different economic weighting factors of boar taint compounds. There is no established payment system for carcasses of intact male boars with respect to boar taint that would allow the derivation of economic weights. Therefore, we arbitrarily have defined relative weights for the three boar taint components with 75% for SKA and IND relative to AND per genetic standard deviation of the trait, and all three components together accounting to 5% of the standard deviation of the overall breeding goal. These assumptions have resulted in economic weights of -2.74, -1.69 and -0.99 Euro per unit of log-transformed AND, SKA and IND. To investigate the effect of higher economic weights of boar taint components on genetic gain, these values were increased in such a way that they represented a proportion of 10%, 20% and 30% of the variance of the overall breeding goal.

(IIb) For the genomic scenario, GBVs were assumed to be available for boar taint compounds AND, SKA and IND. The presumed genotyping costs were 150 Euro. Together with the regular field test, costs added up to 330 Euro.

(IIc) In a third scheme, conventional information from the biopsy-based field test was combined with the genomic information. Consequently, information sources within the selection index were own and half sib performances from the field test, performance of two full sibs and 12 half sibs tested on station, information on the parent's performance and on the

genomic traits. Costs of genotyping and performance testing added up to 380 Euro per selection candidate.

Scenario III, breeding against HNS

The HNS was included in the breeding goal instead of boar taint compounds. Heritability, repeatability and phenotypic standard deviation, and phenotypic and genetic correlations between boar taint compounds and HNS of AND, SKA and IND were adopted from Windig *et al.* (2012) and are displayed in Tables 1 and 2. For the derivation of the economic weight of HNS, we first assumed a new trait H as an index of the chemical compounds AND, SKA and IND, each weighted by their respective index weights (b-values according to selection index theory). The phenotypic variance of H was calculated as $\sigma_p^2 = \mathbf{b}' \mathbf{P} \mathbf{b}$, where \mathbf{b} is a vector of the index weights of AND, SKA, IND and HNS (= 0) and \mathbf{P} is the phenotypic (co)variance matrix of these traits. The phenotypic covariances between H and AND, SKA, IND and HNS, respectively, were calculated as $\mathbf{cov}_H = \mathbf{b}' \mathbf{P}$. Subsequently, the economic weight of trait H was calculated as $\mathbf{w}_H = \mathbf{w}' \mathbf{cov}_H / \sigma_p^2$, where \mathbf{w} is a vector of the economic weights of AND, SKA, IND and HNS (= 0). The economic weight of HNS was then calculated by performing a regression of H on HNS and dividing the economic weight of H by the resulting regression coefficient, which resulted in -2.93 Euro per unit of the trait.

A performance test on station (scenario IIIa) was compared with GS (scenario IIIb).

(IIIa) Information sources for station testing were chosen in accordance with Windig *et al.* (2012). Hence, 4 full sibs and 76 half sibs of the selection candidate were slaughtered and tested by one trained test individual. Additionally, we assessed the effect of a second test individual. Information on station testing of siblings was assumed to cost 50 Euro per selection candidate.

(IIIb) For the genomic strategy, the GBV of the HNS was included in the selection index according to the explanations above.

Results and discussion

Annual genetic gain

Annual genetic gain in log-transformed AND, SKA and IND achieved in scenarios I and II is depicted in Figure 1. Even for the case where no information on boar taint was included in the index (scenario I), we observed a decrease in boar taint compounds. The genetic gain in log-transformed AND (SKA; IND) was 0.06 (0.09; 0.02) Euro per year. This correlated response is because of the selection on favorably correlated production traits such as lean

meat percentage (cp. Table 2). Scenario I only involved the variable costs for regular field testing of 180 Euro per selection candidate (Figure 2).

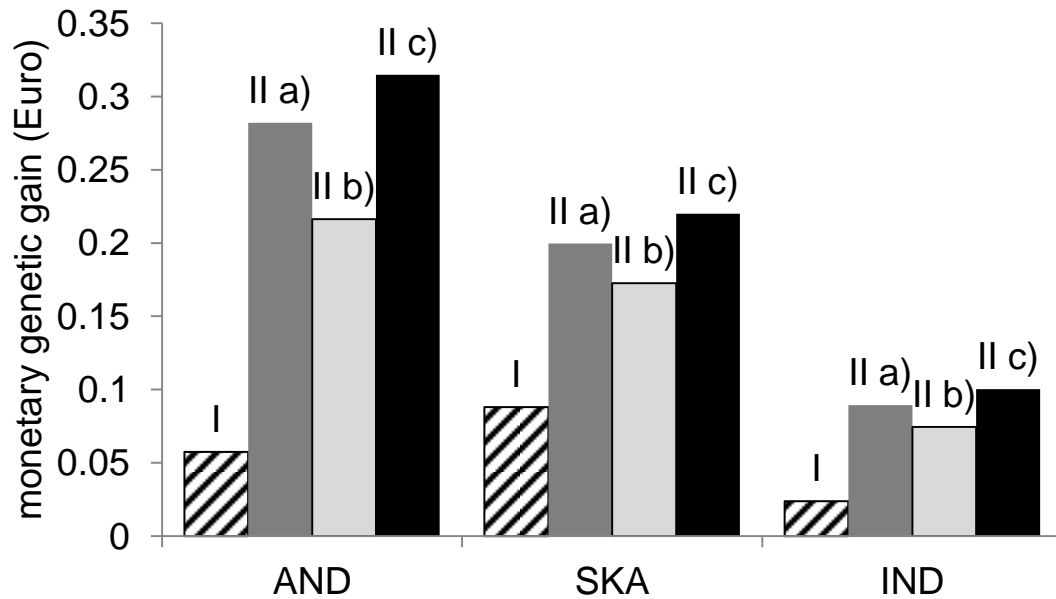


Figure 1. Annual genetic gain of boar taint compounds androstenone (AND), skatole (SKA) and indole (IND) for different information sources: scenario I (hatched); scenario II with biopsy-based performance testing (dark grey); genomic selection (GS, light grey); and a combination of biopsy and genomic selection (black).

In scenario IIa, with information from BPT, genetic gain in AND (SKA; IND) was 0.28 (0.20; 0.09) Euro per year. The high level of genetic gain is because of heritabilities greater than 0.45 and the availability of a high number of half sib performances in addition to an own performance of the selection candidate. Variable costs per selection candidate of scenario IIa were 230 Euro for the field test and the biopsy. When using genomic selection (scenario IIb), the genetic gain was reduced by 23% (14%; 17%) for AND (SKA; IND) compared with scenario IIa. In contrast, variable costs per selection candidate for GS were 100 Euro higher than for BPT (Figure 2). When combining the information sources BPT and GS (scenario IIc), genetic gain only marginally exceeded the gain achievable from BPT alone, whereas variable costs per selection candidate added up to 380 Euro. Consequently, this economically demanding scheme with small additional gain may not be considered for practical application if only considering its benefits for the selection against boar taint compounds. However, the potential of GS with regard to production traits was found to be higher for the same population of pigs (Haberland *et al.*, 2010). If assuming the introduction of GS for selection on production traits, the variable costs per selection candidate could be partly refunded by additional profit in the production traits. In this case, the consideration of genomic information on boar taint compounds in addition to BPT could be worthwhile. Nevertheless, the build-up of a reference population with $N_p > 500$ is challenging for regional lines such as PREMIO[®]. One

possibility to increase N_p would be a joint analysis of genetically close lines within a larger reference population. Investigations of Badke *et al.* (2012) showed high prediction accuracies across breeds (Landrace and Yorkshire) if markers were not spaced more than 100 kb apart. For the PREMO[®] population originating from a Large White line, an even closer relationship with other Large White populations can be expected. Own calculations of the genetic differentiation (Wright, 1951) between the PREMO[®] population and a German Large White line resulted in F_{ST} being in a range with populations that were selected separately for about 50 years. An even more promising approach would be using progeny-tested sows of the same population for increasing N_p . The common genetic background ensures high accuracies of the predicted GBVs. Nevertheless, breeding values for progeny-tested boars and sows will likely differ in accuracy. These differences have to be accounted for by approaches, for example, as proposed by Garrick (2009), in which the residual term of the mixed model is weighted according to the difference in variance.

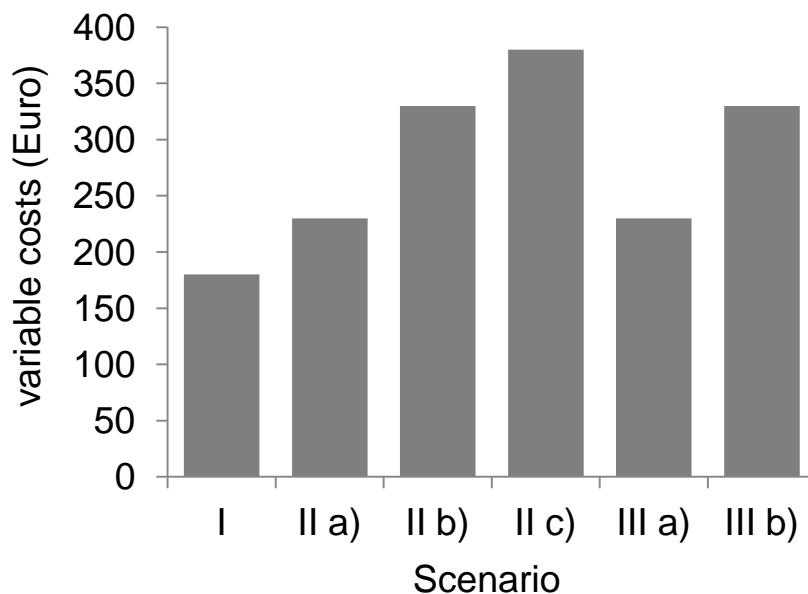


Figure 2. Variable costs in Euro for the reference scenario (I), selection against boar taint compounds via biopsy-based performance testing (II a)), genomic selection (II b)) or both (II c)) as well as for the selection against the human nose score via test persons (III a)) or genomic selection (III b)).

GS is considered to be mainly beneficial if selecting for traits with a low heritability and those that cannot be measured in the animal itself (Goddard and Hayes, 2007). As HNS is a carcass trait, which can only be measured in sibs of the selection candidate, we expected a high potential for GS in scenario III. The annual genetic gain in units of the trait for scenario III is depicted in Figure 3. In comparison with station testing of 4 full and 76 half sibs with one or two measurements, annual genetic gain could be increased by factor 2.8 or 1.8, respectively, when using genomic information on the selection candidate. However, the variable costs of station testing are considerably lower than the costs of GS (Figure 2). We also assessed the correlated response of HNS when using scenario IIa, that is, when selecting against the chemical compounds of boar taint via BPT. Because of the high genetic correlations between boar taint compounds and HNS (cp. Table 2), the (correlated) response in HNS could be further increased by factor 2.6 compared with scenario IIIb (Figure 3). Thus, the best strategy for reducing the HNS was breeding against chemical compounds (scenario IIa).

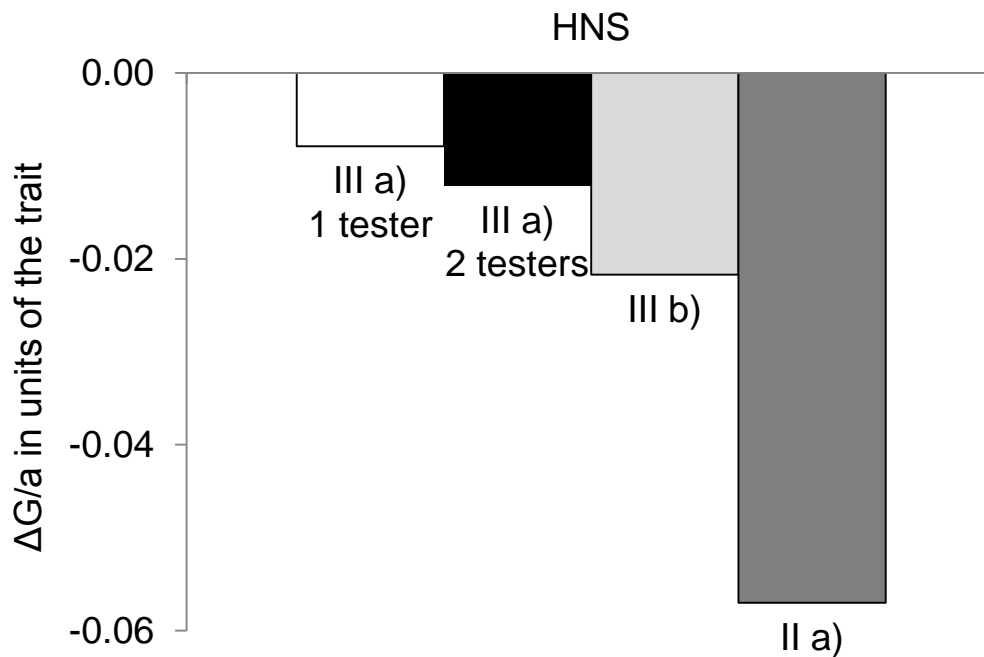


Figure 3. Annual genetic gain ($\Delta G/a$) in units of the trait for the human nose score (HNS) for different information sources: station testing of siblings with one test individual (white); station testing of siblings with two measurements (black); genomic selection (light grey); biopsy-based performance testing for boar taint compounds (correlated response, dark grey).

Expected change in average amount of boar taint:

The average amounts of chemical compounds AND, SKA and IND in the current PREMO[®] population were 0.70, 0.03 and 0.03 $\mu\text{g/g}$ liquid fat, respectively. Starting with these values, we calculated the expected change in the actual amount of untransformed AND, SKA and IND for a period of 10 years (Figure 4). In scenario I, the amount of AND (SKA; IND) de-

creased by 19% (41%; 21%) within ten years, because of favorable correlations to lean meat percentage and meat surface.

Because genetic gain was found to be greatest for scenario IIa, the calculation of the following trends was only performed for this scenario. The amount of AND could be reduced by 50% within 7 years if assuming an economic weight of -2.74 Euro per unit of the trait (Figure 4a). Desmoulin and Bonneau (1982) proposed a threshold of 0.5 µg AND/g liquid fat, below which consumers found no more difference between boar meat and meat from gilts or castrates. The average amount of AND in the PREMO[®] population could be reduced to this threshold within 4 years, which is in accordance with Merks *et al.* (2009) and Windig *et al.* (2012). If selecting more intensively, that is, increasing the economic weight up to a proportion of 10% or 20% of the variance of the overall breeding goal, this threshold could be reached even within 3 or 2 years. The average amount of SKA in the PREMO[®] population is currently 0.029 µg/g liquid fat, which is already very close to the threshold of 0.026 µg/g liquid fat proposed by Annor-Frempong *et al.* (1997). Assuming an economic weight of -1.69 Euro per unit of the trait, the amount of SKA could be reduced by 50% within 6 years (Figure 4b). A further increase in economic weighting only had marginal effects. The amount of IND could be reduced by 50% within 8 years with the original economic weighting of -0.99 Euro per unit of the trait (Figure 4c). If economic weighting was increased up to a proportion of 10% or 20% of the variance of the overall breeding goal, the amount of IND could be halved within 6 or 4 years. For all three chemical compounds, a further increase of economic weights only provided marginal improvements but amplified negative effects, which will be discussed in the following sections.

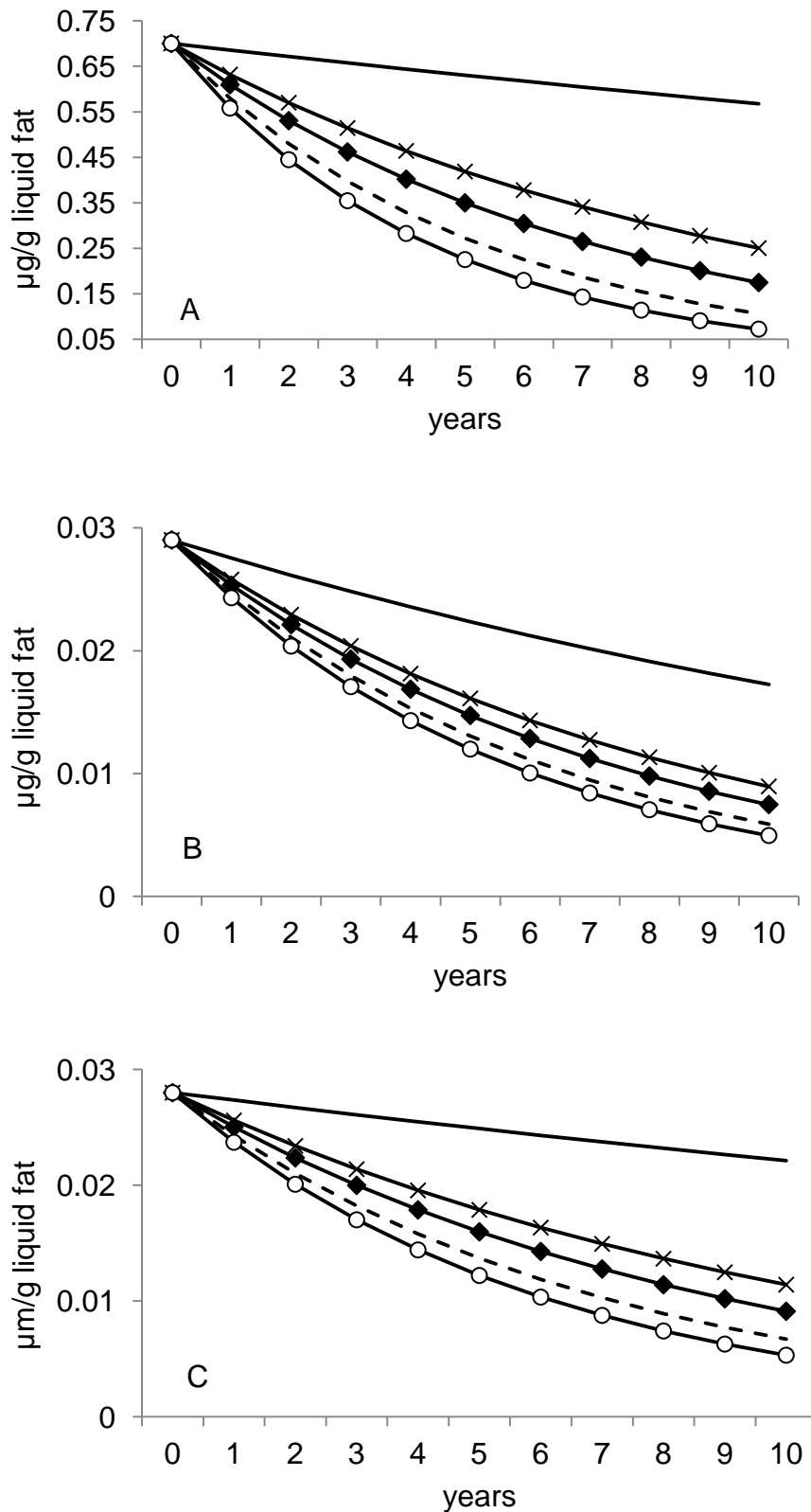


Figure 4. Expected changes in the prevalence of boar taint compounds androstenone (A), skatole (B) and indole (C) as a function of economic weighting: scenario I (solid line); scenario II and biopsy-based performance testing with weighting of -2.74, -1.69 or -0.99 (crosses); economic weighting increased up to a proportion of 10% (diamonds); 20% (dashed line); 30% (circles) of the variance of the overall breeding goal.

Correlated effects on other traits

Figure 5 depicts the annual monetary genetic gain in the production traits for scenario I and scenario IIa. Some breeding goal traits were left out of the comparison because information on their genetic correlation with boar taint compounds was not available (Table 2). Scenario III had to be left out for the same reason. The initial economic weights for AND, SKA and IND of -2.74, -1.69 and -0.99 Euro per unit of the trait, respectively, were increased up to a proportion of 10% and 20% of the variance of the overall breeding goal within this comparison. Selection against boar taint entailed positive effects on lean meat percentage, meat surface and feed conversion ratio because of favorable genetic correlations (cp. Table 2). For example, when conducting scenario IIa with economic weights of -2.74, -1.69 and -0.99 Euro per unit of the trait for AND, SKA and IND, respectively, annual genetic gains in these traits increased by 0.10 Euro, 0.12 Euro and 0.09 Euro, respectively, in comparison with scenario I (Figure 5). Negative effects were found for average daily gain (station test), drip loss and intramuscular fat percentage. When conducting scenario IIa with the initial economic weights, monetary genetic gain in these traits changed by -0.05 Euro, 0.06 Euro and -0.16 Euro, respectively, in comparison with scenario I. The negative effects on these traits increased when the economic weighting of boar taint compounds increased (Figure 5). A negative trend for growth rate in a Large White line selected for low AND was also reported by Sellier and Bonneau (1988) and Sellier *et al.* (2000), but is in contrast to findings of Windig *et al.* (2012). The different information sources within scenario II had no noticeable impact on annual monetary genetic gain in the production traits (results not shown).

Strong genetic correlations between AND and other sex steroids, like for example, testosterone have been reported by, for example, Willeke *et al.* (1987) and Grindflek *et al.* (2011). Moreover, the level of AND has been found to be strongly correlated with testes size (Sellier and Bonneau, 1988) and size of the bulbo-urethral gland (Sellier *et al.*, 2000). Therefore, selection against AND may entail problems such as delayed sexual maturity in male boars as reported for females (Willeke *et al.*, 1987; Sellier and Bonneau, 1988). One possibility to prevent negative effects on other sex hormones would be assessing single genes rather than conventional selection without molecular information. A genome-wide association study by Grindflek *et al.* (2011) showed breed-specific QTL associated with SKA, but most QTL affecting AND also showed associations with other sex steroids. Contrary to these findings, Sellier and Bonneau (1988), Bergsma *et al.* (2007) and Merks *et al.* (2010) found no or even positive effects of selection against boar taint compounds on male fertility traits. The relationship between boar taint and fertility is not yet conclusive and requires further investigation. Nevertheless, our results show that breeding against chemical compounds measured

by BPT is an effective and powerful way to reduce the occurrence of boar taint in finishing pigs.

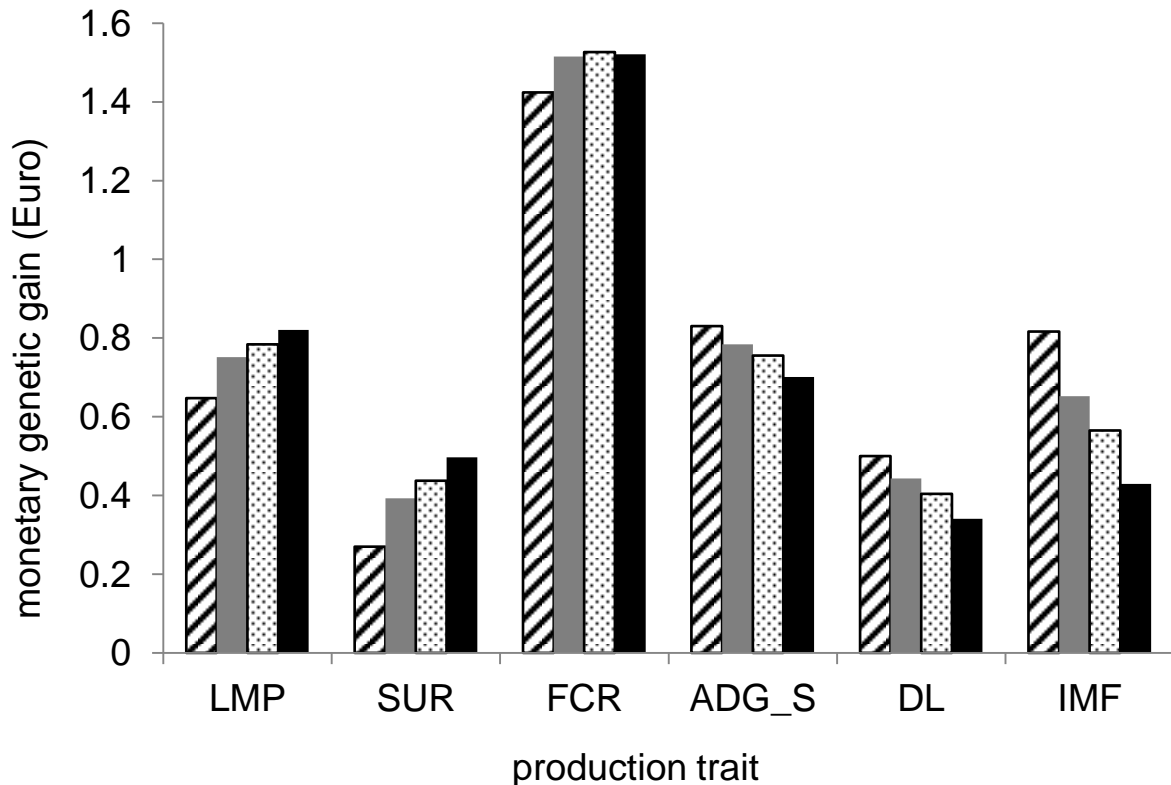


Figure 5. Annual monetary genetic gain in the production traits lean meat percentage (LMP), surface (SUR), feed conversion ratio (FCR), average daily gain measured on station (ADG_S), drip loss (DL) and percentage of intramuscular fat (IMF) for the reference scenario (hatched) and scenario II a) with economic weights of -2.74, -1.69 or -0.99 Euro per unit of log transformed androstenone, skatole and indole, respectively (dark grey) and for economic weights increased up to a proportion of 10% (light grey); and 20% (black) of the variance of the overall breeding goal.

Conclusions

On the basis of our results, breeding against boar taint by conducting BPT is an effective method for optimizing both the selection against the chemical compounds AND, SKA and IND (scenario II), as well as the selection against HNS of boar taint (scenario III) in terms of genetic gain per year and variable costs per selection candidate. By using economic weights of -2.74, -1.69 and -0.99 Euro per unit of log transformed AND, SKA and IND, the average amount in the PREMIO[®] population could be reduced by 50% within 7, 6 or 8 years, respectively; an average amount of 0.5 µg AND/g liquid fat could be reached within 4 years. Despite these advantages, the introduction of boar taint as a selection trait should be undertaken with caution owing to possible negative effects on average daily gain, drip loss and intramuscular fat percentage, as well as possible negative effects on fertility traits.

ACKNOWLEDGMENTS

The authors thank the European Union and the federal state of North Rhine-Westphalia for the financial support within the project pigGS (005-NA02-018 C). Financial support of the Swiss Commission for Innovation and Technology is also gratefully acknowledged.

Investing in **your** future!



References

- Annor-Frempong IE, Nute GR, Whittington FW and Wood JD 1997. The problem of taint in porc: 1. Detection thresholds and odour profiles of androstenone and skatole in a model system. *Meat Science* 46, 45-55.
- Badke YM, Bates RO, Ernst CW, Schwab C and Steibel JP 2012. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* 13, 24.
- Baes C, Mattei S, Luther H, Ampuero Kragten S, Sidler X, Bee G, Spring P and Hofer A 2012. A performance test for boar taint compounds in live boars. *animal* doi:10.1017/S1751731112002273.
- Bergsma R, Knol E and Feitsma H 2007. Parameters of AI boars and predicted correlated responses of selection against boar taint. In: *Book of abstracts of the 58th Annual Meeting of the European Association for Animal Production*. Wageningen Academic Publishers, Dublin.
- Cleveland MA, Forni S, Garrick DJ and Deeb N 2010. Prediction of Genomic Breeding Values in a Commercial Pig Population. In: *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, Germany: August 2010 47.
- 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. *J Anim Breed Genet* 124, 331-341.
- Desmoulin B and Bonneau M 1982. Consumer testing of pork and processed meat from boars: The influence of fat androstenone level. *Livest Prod Sci* 9, 707-715.
- Elsen JM and Mocquot JC 1974. Méthode de prévision de l'évolution du niveau génétique d'une population soumise à une opération de sélection et dont les générations se chevauchent. *INRA Bull. tech. Dépt. Génét. anim.*, 17, 30-54.
- Erbe M, Reinhardt F, Simianer H 2011. Empirical determination of the number of independent chromosome segments based on cross-validated data. *Book of abstracts of the 62th Annual Meeting of the European Association for Animal Production*, August 29 - September 2, 2011, Stavanger, Norway, p. 115. Wageningen Academic Publishers, The Netherlands.
- European Commission 2010. European Declaration on alternatives to surgical castration of pigs. http://ec.europa.eu/food/animal/welfare/farm/initiatives_en.htm.
- Fàbrega E, Velarde A, Cros J, Gispert M, Suárez P, Tibau J and Soler J 2010. Effect of vaccination against gonadotrophin-releasing hormone, using Improvac[®], on growth performance, body composition, behaviour and acute phase proteins. *Livest Sci* 132, 53-59.

- Garrick DJ, Taylor JF and Fernando RL 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genetics Selection Evolution* 41, 55.
- Goddard ME and Hayes BJ 2007. Genomic selection (Review article), *J Anim Breed Genet* 124, 323-330. Goddard ME, Hayes BJ, Meuwissen THE 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *Journal of Animal Breeding and Genetics* 128, 409-421.
- Grindflek E, Meuwissen THE, Aasmundstad T, Hamland H, Hansen MHS, Nome T, Kent M, Torjesen P and Lien S 2011. Revealing genetic relationships between compounds affecting boar taint and reproduction in pigs. *J Anim Sci* 89, 680-692.
- Haberland AH, Ytournal F, Luther H and Simianer H 2010. Evaluation of selection strategies including genomic breeding values in pigs. In: *Book of abstracts of the 61th Annual Meeting of the European Association for Animal Production*. Wageningen Academic Publishers, Dublin.
- Haberland AH, Pimentel, ECG, Ytournal F, Erbe M and Simianer H 2013. Interplay between heritability, genetic correlation and economic weighting in a selection index with and without genomic information. *J Anim Breed Genet*, published online on 29th of August 2013. doi: 10.1111/jbg.12051.
- Hazel LN 1943. The genetic basis for constructing selection indexes. *Genetics* 28, 476-490.
- Hill WG 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim Prod* 18, 117-139.
- Karras K, Niebel E, Nitter G and Bartenschlager H 1993. ZPLAN- a PC computer program to optimise livestock selection programs. University of Hohenheim: Hohenheim.
- Merks JWM, Bloemhof S, Mathur PK and Knol EF 2010. Quantitative genetic opportunities to ban castration. In: *Book of abstracts of the 61th Annual Meeting of the European Association for Animal Production*. Wageningen Academic Publishers, Dublin.
- Mathur PK, ten Napel J, Bloemhof S, Heres L, Knol EF and Mulder HA 2012. A human nose scoring system for boar taint and its relationship with androstenone and skatole. *Meat Science* 91, 414-422.
- Merks JWM, Hanenberg EHAT, Bloemhof S, Knol EF 2009. Genetic opportunities for pork production without castration. *Animal Welfare* 18, 539-544.
- Meuwissen THE, Hayes BJ and Goddard ME 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Prunier A, Bonneau M, von Borell EH, Cinotti S, Gunn M, Fredriksen B, Giersing M, Morton DB, Tuytens FAM and Velarde A 2006. A review of the welfare consequences of surgical castration in piglets and the evaluation of non-surgical methods. *Animal Welfare* 15, 277-289.

- Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu Z-L, Kerstens HH, Law AS, Megens H-J, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TPL, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB and Groenen MAM 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE* 4, e6524. doi:10.1371/journal.pone.0006524.
- Rohrer GA, Alexander LJ, Hu Z, Smith TP, Keele JW, Beattie CW 1996. A comprehensive map of the porcine genome. *Genome Research* 6, 371-391.
- Rydmer L, Lundström K and Andersson K 2010. Immunocastration reduces aggressive and sexual behavior in male pigs. *animal* 4, 965-972.
- Sellier P and Bonneau M 1988. Genetic relationships between fat androstenone level in males and development of male and female genital tract in pigs. *J Anim Breed Genet* 105, 11–20.
- Sellier P, Le Roy P, Fouilloux MN, Gruand J and Bonneau M 2000. Responses to restricted index selection and genetic parameters for fat androstenone level and sexual maturity status of young boars. *Livest Prod Sci* 63, 265-274.
- Täubert H, Reinhardt F and Simianer H 2010. ZPLAN+, a new software to evaluate and optimize animal breeding programs. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production. 1-6 August 2010, Leipzig, Germany.
- Uimari P and Tapio M 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *J Anim Sci* 89, 609-614.
- Vazquez JM, Parrilla I, Roca J, Gil MA, Cuello C, Vazquez JL and Martínez EA 2009. Sex-sorting sperm by flow cytometry in pigs: Issues and perspectives. *Theriogenology* 71, 80-88.
- von Borell E, Baumgartner J, Giersing M, Jäggin N, Prunier A, Tuytens FAM and Edwards SA 2009. Animal welfare implications of surgical castration and its alternatives in pigs. *animal* 3, 1488-1496.
- Walstra P 1974. Fattening of young boars: Quantification of negative and positive aspects. *Livest Prod Sci* 1, 187-196.
- Willam A, Nitter G, Bartenschlager H, Karras K, Niebel E and Graser H-U 2008. ZPLAN- Manual for a PC-program to optimize livestock selection schemes.
- Willeke H, Claus R, Müller E, Pirchner F, Karg H 1987. Selection for high and low level of 5 α -androst-16-en-3-one in boars. *Journal of Animal Breeding and Genetics* 104, 64-73.
- Windig JJ, Mulder HA, ten Napel J, Knol EF, Mathur PK and Crump RE 2012. Genetic parameters for androstenone, skatole, indole, and human nose scores as measures of boar taint and their relationship with finishing traits. *J Anim Sci* 90, 2120-2129.

Wright S 1951. The genetical structure of populations. *Ann Eugen* 15, 323–354.

5TH CHAPTER

General Discussion

Prospects of genomic selection for horse breeding

Sport horse breeding is characterized by long generation intervals because riding performance can be tested at the age of three years at earliest. As the development of the bones only stops at the age of about five years (Voswinkel, 2009), selection early in life is to some extent unsure. Nevertheless, a large proportion of male selection candidates is already culled after a first selection step at the age of six months, on the occasion of the inspection of foals, long before their riding performance can be assessed (von Lengerken and Schwark, 2002). Results presented in chapter two of this thesis show that selection decisions at this young age could be carried out more accurately by genotyping the foals prior to castration. This strategy could save male selection candidates for a later, more accurate, selection step according to the target traits (riding performance).

The same applies for the second selection step, the licensing of stallions at the age of two and a half years. Riding performance still cannot be measured at this occasion and conformation traits only show low to moderate correlations to riding performance (Koenen *et al.*, 1995). Thus, the prospects of GS for increasing the accuracy of selection at this selection step are promising. Recently, a new procedure for licensing has been implemented by the breeding association of Oldenburger horses. Stallions are presented only at the age of three but therefore already ridden. This practice will most probably increase the accuracy of selection in regard to riding performance.

The possible size of a reference population for the implementation of GS in German sport horses largely depends on the willingness of the different breeding associations to cooperate in this matter. Research will be needed to investigate to which extent the accuracy of GBVs can be increased by assembling horses of different breeding organizations in one larger reference population. This will depend on the actual genetic exchange between different sport horse populations. McCue *et al.* (2012) calculated genetic distances, inbreeding coefficients and LD for 14 horse breeds, thereof two sport horse breeds (Hanoverian and Swiss Warmblood). The Hanoverian was one of the breeds with the lowest inbreeding coefficients and the largest genetic distances between pairs of individuals within the breed, which according to the authors has resulted from admixture. This admixture could be explained by an interchange of genetic material between different German or European sport horse populations (Koenen *et al.*, 2004; Thorén Hellsten *et al.*, 2009). For example, the Holsteiner is often used for increasing the jumping ability of the Hanoverian, and several stallions with excellent dressage performance from the Netherlands are approved for matings within the Hanoverian population. Therefore, genetic distances between sport horse breeds in Germany or even within Europe might be shorter than the ones between diverse breeds like Icelandic, Arabian

or Quarter Horse, estimated by McCue et al. (2012). Studies will be needed to confirm (or not) this hypothesis.

In German sport horse breeding, there is still unused potential of conventional selection strategies, which could be exploited before introducing genomic breeding strategies. The stationary performance test for stallions was shown to have high correlations with competition results (Huizinga, 1991; Thorén Hellsten *et al.*, 2006; Olsson *et al.*, 2008). In Germany, its value was further increased by estimating and publishing a breeding value (SEBV) based on the results of the selection candidates and their male relatives (Figure 1). Unfortunately, the concept of breeding values is not well accepted by horse breeders.

Like in other livestock species, breeding values for horses are estimated using a multiple-trait repeatability model (von Velsen-Zerweck, 1998). But the integrated breeding values (IEBVs) incorporating all performances from station tests and riding competitions of all relatives of a horse are only published when achieving a reliability of 0.7 ($r_{TI} = 0.84$) and when they are based on performance data of at least five progeny (Figure 1).

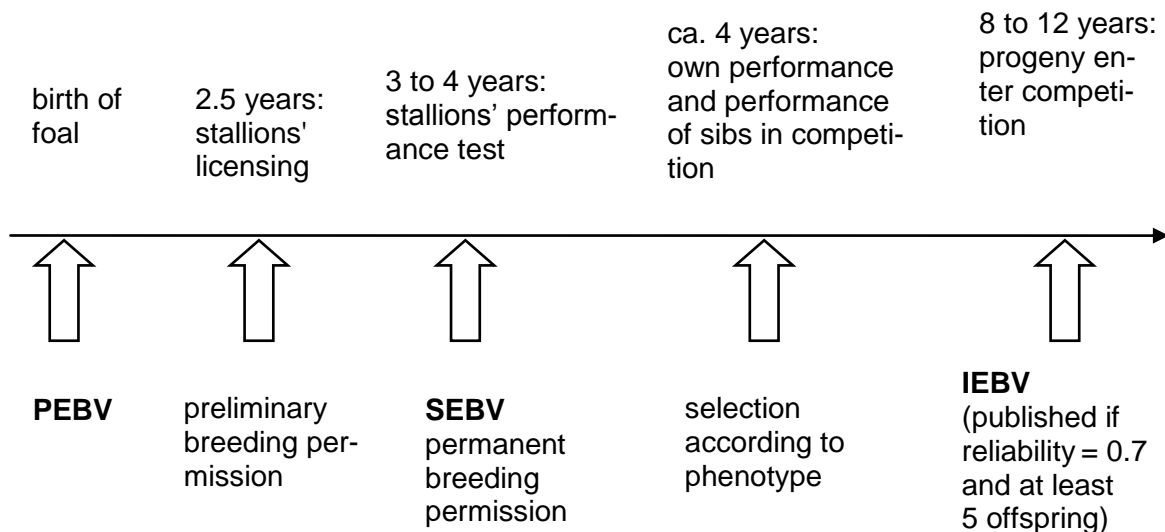


Figure 1. Availability of different breeding values in sport horse breeding (PEBV= pedigree-based breeding value; SEBV= breeding value for stationary performance testing of stallions; IEBV= integrated breeding value).

This is clearly too late, as stallions are at the age of eight to twelve years when their IEBVs are published and a large proportion of matings is performed by young stallions (Niemann, 2009). The generation interval of the Hanoverian has been decreasing within the last 40 years (Niemann, 2009), which would be desirable if the selection at an earlier age would be conducted with the same accuracy. But the opposite seems to be true, as breeders tend

to choose the winners of stallions' licensing or stallions' performance tests, as well as stallions with good results in popular competitions like the Bundeschampionat. One reasonable strategy would be to publish the IEBVs instead of the SEBVs on the occasion of the stallions' performance test to provide breeders earlier with extensive information on performances of all relatives of a stallion.

Another possibility for conventionally increasing response to selection would be to increase the selection intensity of the mare side. As more than 90% of mares are registered as breeding animals in the studbook (Sitzenstock, 2012) and the traits measured during the studbook registration show very low correlations to competition later in life (Koenen *et al.*, 1995), selection occurring during this event has negligible effects. But from a commercial perspective, it is an important event for breeding associations in order to get breeders attached to their organization. Nevertheless, an effective selection strategy is also in the breeders' interest because only the best horses achieve high market prices. However, a large proportion of horse breeders does not focus on profit, but take horse breeding as their hobby (Klunker and Barth, 2008). Horse breeding organizations therefore are forced to meet the diverging interests of their breeders, which often involves suboptimal selection strategies compared to other livestock breeding programs.

Linear scoring of conformation traits has been proposed for optimizing the assessment of conformation traits in German sport horse breeding (von Lengerken and Schwark, 2002; Hartmann, 2006). In horse breeding, this system has been successfully practiced in the Netherlands and in Switzerland. The three main advantages of the linear scoring system are: (1) a more objective assessment of traits; (2) mating decisions by the breeder can be based on precise information on the characteristics of their mares; (3) a more precise assessment of definite traits, which is a prerequisite in regard to GS.

Thus, prior to implementation of GS, there is still a great potential of conventional tools to increase response to selection in sport horse breeding. The implementation of GS is currently discussed in the Netherlands for the selection against *osteochondrosis*, which will force German breeding organizations to react in order to stay competitive.

Prospects of genomic selection for pig breeding

Because the economic importance of traits is known in pig breeding, assessing the efficiency of selection strategies is much easier than in horse breeding. It was shown in chapter three of this thesis that the size of a reference population has to be at least 1'000 animals, if GS is to be beneficial in pig breeding. However, building up a reference population consisting of 1'000 progeny tested boars cannot be accomplished by most of the breeding organizations. Mostly, the genetic exchange between the regionally operating enterprises is rare even within the breeds. F_{ST} -values, as a measure of genetic differentiation (Wright, 1951) between a German and a Swiss Large White population were in the range of F_{ST} -values between two simulated populations, which were selected separately for about 50 years (Wiebelitz and Erbe, 2012). For this reason, the benefit of combining different populations of the same breed within a larger reference set, similar to the EuroGenomics project in Holstein cattle (Lund *et al.*, 2010), may be small with the currently available 60k chip for pigs. Badke *et al.* (2012) used the 60k chip to estimate LD in four US pig breeds and reported relatively low persistence of phase across breeds, and suggested limited use of multi-breed panels for genomic selection. Nevertheless, increased persistence of phase across populations may be achieved with higher marker densities in the future. De Roos *et al.* (2008) suggested that to achieve persistence of phase across breeds such as Holstein-Friesian, Jersey and Angus would require ~300k markers.

Considering the current marker density available, a more promising approach seems to be including sows with records on one or two litters in the reference population. Even if the accuracy of their EBVs is lower, the additional information is based on the same genetic background and therefore is more valuable than information on distantly related animals with a higher accuracy of their EBVs. A further strategy to enlarge the size of the reference population could be to perform imputation of dams, as proposed by Pimentel *et al.* (2013). This approach focuses on the imputation of genomic information of un-genotyped dams based on the information on already genotyped progeny, their sires and maternal grandsires. Further research would be needed to investigate the feasibility and potential benefit of these approaches. Because the productive life time of sows is much shorter than of cows, there may be not enough progeny available, which are already genotyped.

A further aspect to be considered in genomic selection programs for pigs is the fact that selection within the purebred lines should increase performance on their crossbred descendants. Dekkers (2007) proposed a selection strategy where breed-specific marker effects on crossbred performance are estimated using a reference population of crossbred end-products. These effects would then be used to predict GBVs for crossbred performance of purebred animals. The number of end-product animals in the crossbred population should

usually be larger than the number of boars used for matings within the nucleus, which would result in larger reference populations. As the aim of purebreeding in pigs is to improve the performance of crossbred end-products, results by Dekkers (2007) showed that including information on crossbred relatives for the estimation of GBVs increases the response to selection in the nucleus with regard to crossbreeding performance. Ibáñez-Escriche *et al.* (2009) compared two models for prediction of merit of purebreds for performance of their crossbred descendants: one model assumed breed specific marker allele effects and the other assumed the same marker effects across breeds. They concluded that with high marker density estimation of breed specific marker effects may not be necessary.

Despite small reference populations of ~300 animals, high accuracies in the range of 0.60 to 0.75 were found for GBVs estimated for the trait piglets born alive (Gertz *et al.*, 2013; Simianer *et al.*, 2013). As similar findings were reported for laying hens (Erbe, personal communication), the reason for these high accuracies may be due to LD patterns diverging from those in cattle populations. LD in pigs was found to be higher than in American Holstein cattle, especially for higher marker distances (Uimari and Tapio, 2011; Badtke *et al.*, 2012).

GS is known to be especially useful when selecting on traits with a low heritability (Lande and Thompson, 1990; Goddard and Hayes, 2007; König and Swalve, 2009). This finding was again confirmed by the results presented in chapter three. For traits with a low heritability, an index including genomic information was advantageous over an index incorporating the normally available conventional performances (own and parent performances as well as full and half sib performances). However, for some applications, classical selection on an indicator trait may be more efficient than GS as was found in chapter four. The best strategy for selection against HNS of boar taint, which has a low heritability, was selecting against the chemical compounds AND, SKA and IND. The correlated response in units of the trait was higher than the response to selection when selecting directly on HNS. This was due to the strong genetic correlations between the breeding goal trait and the indicator traits, as well as to the high heritability of the indicator traits. Additionally, the variable costs per selection candidate were much higher for GS than for biopsy-based performance testing.

For the same population of pigs, the potential of GS was found to be higher in regard to production traits (Haberland *et al.*, 2010). If assuming the introduction of GS for selection on production traits, the variable costs per selection candidate could be partly refunded by additional profit in the production traits. In this case, the consideration of genomic information on boar taint compounds in addition to BPT could be worthwhile and yet increase response to selection. As long as castration is still in practice, another possibility for refunding of genotyping costs would be an earlier castration of a higher proportion of selection candidates based on their GBVs. Due to the lower risk of being tainted, prices for castrates are higher

than for entire boars. Apportioning this additional gain over all selection candidates considerably lowers the genotyping costs per animal. Wellmann (2013) proposed to genotype the selection candidates with a low-density SNP chip of only 384 SNP markers and in a second step to use imputing procedures (as proposed by Weigel *et al.*, 2010 for dairy cattle) to increase the number of markers up to 60'000 SNPs. The authors found that SNPs for the reduced panel should be chosen with respect to an equal spacing across the genome rather than to their estimated effects. More selection candidates could be genotyped due to the reduced genotyping costs. On the other hand, the imputation on the basis of such a small number of known SNPs involves an error rate of around 18%. When using this strategy, the selected boars used for matings have to be genotyped again with a high-density SNP chip (60K), to prevent an accumulation of errors (Wellmann, 2013).

Generally, GS seems to have good prospects for pig breeding. However, its potential for a special application has to be assessed in detail by considering all the interacting parameters. The potential of GS largely depends on the characteristics of the respective breeding program. A software like ZPlan+ offers an effective way to increase the understanding of interactions occurring for specific selection applications.

References

- Badke YM, Bates RO, Ernst CW, Schwab C and Steibel JP 2012. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics* 13, 24.
- Dekkers JCM 2007. Marker-assisted selection for commercial crossbred performance. *J Anim Sci* 85, 2104-2114.
- de Roos APW, Hayes BJ, Spelman RJ and Goddard ME 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus Cattle. *Genetics* 179, 1503-1512.
- Gertz M, Edel C, Dodenhoff J, Götz K-U and Thaller G. Genomische Selektion bei Mutterlinien – Projekt FROGS. In: 9th Schweineworkshop Uelzen, DGfZ-Schriftenreihe Heft 62, Bonn, Germany.
- Haberland AM, Ytournel F, Luther H and Simianer H 2010. Evaluation of selection strategies including genomic breeding values in pigs. In: Proc. of 61th Annu. Mtg. Eur. Assoc. Anim. Prod., Heraklion, Greece, August 23-27, 2010. Wageningen Academic Publishers, The Netherlands.
- Huizinga HA, van der Werf JHJ, Korver S and van der Meij GJW 1991. Stationary performance testing of stallions from the Dutch Warmblood riding horse population. 1. Estimated genetic parameters of scored traits and the genetic correlation with dressage and jumping competition from offspring of breeding stallions. *Livest Prod Sci* 27, 231-244.
- Ibáñez-Escriche N, Fernando RL, Toosi A and Dekkers JCM 2009. Genomic selection of purebreds for crossbred performance. *Genet Sel Evol* 41, 12.
- Klunker M and Barth I 2008. Wünsche der Praxis an eine Zuchtorganisation – Ergebnis einer Befragung. In: 5th Pferdeworkshop Uelzen, 14-21, Uelzen, Germany.
- Koenen EPC, Aldridge LI and Philipsson J 2004. An overview of breeding objectives for Warmblood sport horses. *Livest Prod Sci* 88, 77-84.
- Koenen EPC, van Veldhuizen AE and Brascamp EW 1995. Genetic parameters of linear scored conformation traits and their relation to dressage and show-jumping performance in the Dutch Warmblood Riding Horse population. *Livest Prod Sci* 43, 85-94.
- Lande R and Thompson R 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743-756.
- 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 M 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 2010, Leipzig, Germany.

- McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guérin G, Hasegawa T, Hill EW, Leeb T, Lindgren G, Penedo MCT, Røed KH, Ryder OA, Swinburne JE, Tozaki T, Valberg SJ, Vaudin M, Lindblad-Toh K, Wade CM and Mickelson JR 2012. A high density SNP array for the domestic horse and extant perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet* 8, e1002451. doi:10.1371/journal.pgen.1002451.
- Niemann B 2009. Untersuchungen zu Veränderungen im Zuchtgeschehen und deren Auswirkungen auf die Hannoveraner Pferdezucht. PhD, Goettingen Univ.
- Olsson E, Näsholm A, Strandberg E and Philipsson J 2008. Use of field records and competition results in genetic evaluation of station performance tested Swedish Warmblood stallions. *Livest Sci* 117, 287-297.
- Pimentel ECG, Wensch-Dorendorf M, König S, Swalve HH 2013. Enlarging a training set for genomic selection by imputation of un-genotyped animals in populations of varying genetic architecture. Under revision (*Genet Sel Evol*).
- Simianer H, Bergfelder S, Haberland AM, Große-Brinkhaus C, Ni G, Lind B and Tholen E 2013. Genomische Selektion bei Mutterlinien – das Projekt pigGS. In: 9th Schweineworkshop Uelzen, DGfZ-Schriftenreihe Heft 62, Bonn, Germany.
- Sitzenstock F 2012. Genetic progress and inbreeding rate in complex breeding programmes - applications to sport horses and laying hens. PhD thesis, Georg-August University, Goettingen, Germany.
- 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. *Livest Sci* 103, 1-12.
- Thorén Hellsten E, Näsholm A, Jorjani H, Strandberg E and Philipsson J 2009. Influence of foreign stallions on the Swedish Warmblood breed and its genetic evaluation. *Livest Sci* 121, 207-214.
- Uimari P and Tapio M 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *J Anim Sci* 89, 609-614.
- von Lengerken G and Schwark H-J 2002. Exterieur und Leistungen in der Pferdezucht – Alleskönner oder Spezialisten. *Arch Tierz Dummerstorf* 45, 68-79.
- von Velsen-Zerweck A 1998. Integrierte Zuchtwertschätzung für Zuchtpferde. PhD thesis, Georg-August University, Goettingen, Germany.
- Voswinkel L 2009. Einfluss der Bewegungsaktivität auf Wachstums- und Ausdauerparameter beim Pferd. PhD thesis, University of Kiel, Germany.

-
- Weigel KA, de los Campos G, Vazquez AI, Rosa GJM, Gianola D and Van Tassell CP 2010. Accuracy of direct genomic breeding values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. *J Dairy Sci* 93, 5423-5435.
- Wellmann R, Preuß S, Tholen E, Wimmers K and Bennewitz J 2013. Genomic selection in pig breeding using low density marker panels. Under revision.
- Wiebelitz J, Erbe M and Simianer H 2012. Genauigkeit der genomischen Zuchtwertschätzung in unterteilten Populationen. In: Gemeinschaftstagung der Deutschen Gesellschaft für Züchtungskunde, 12th to 13th September in Kiel, Germany.
- Wright S 1951. The genetical structure of populations. *Ann Eugen* 15, 323–354.

I would like to thank:

Prof. Dr. Simianer

Prof. Dr. König

Prof. Dr. Thaller

The financiers of the projects FUGATO+brain and pigGS

Mrs. Döring

Florian

Manfred

Malena

Eduardo

My family

God

-
- geboren 1981 in Berlin
 - nach dem Abitur Ausbildung zum Pferdewirt im Haupt- und Landgestüt Marbach
 - Bachelorstudium der Pferdewissenschaften an der Veterinärmedizinischen Universität Wien
(Thema der Abschlussarbeit bei Prof. Dr. J. Sölkner: *„Die Ausnutzung der Notenskala bei der Exterieurbeurteilung: Ein Vergleich zwischen Bonitursystem und Linearer Beschreibung“*)
 - Masterstudium der Nutztierwissenschaften an der Georg-August-Universität Göttingen
(Thema der Abschlussarbeit bei Prof. Dr. S. König: *„Evaluierung von Selektionsstrategien in der deutschen Reitpferdezucht“*)
 - 2009-2013 wissenschaftliche Mitarbeit und Promotion im Department für Nutztierwissenschaften, Abteilung Tierzucht und Haustiergenetik (Prof. Dr. H. Simianer)