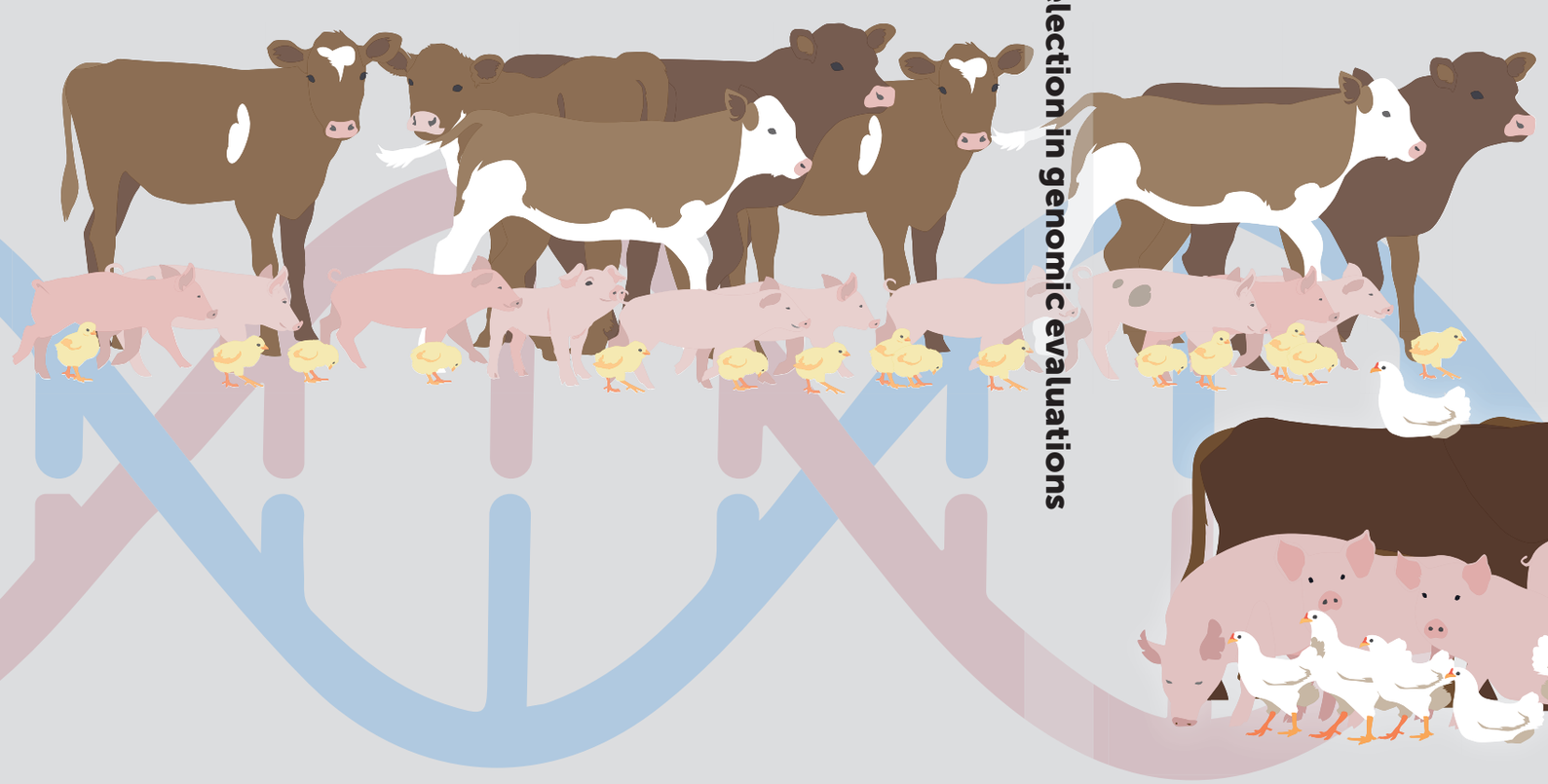
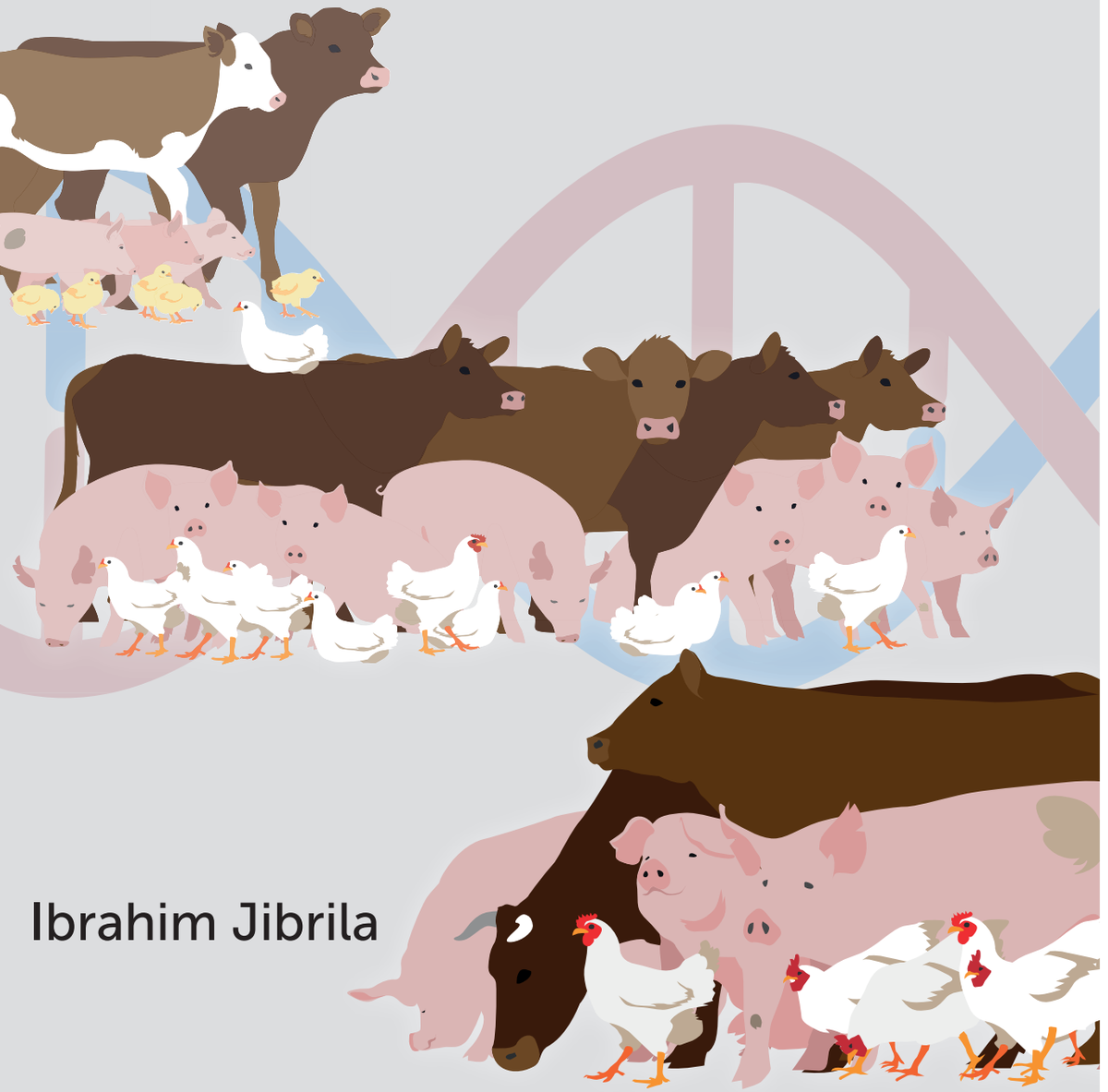
Impact of preselection in genomic evaluations



Impact of preselection in genomic evaluations

Ibrahim Jibrila

Ibrahim Jibrila



Propositions

1. With the advent of single-step genomic evaluation, preselection bias no longer exists.
   (this thesis)

2. Genotypes of animals culled prior to phenotyping are not useful in single-step genomic evaluations.
   (this thesis)

3. Merits of publishing the outcome of a PhD project in only one scientific article outweigh its demerits.

4. The greatest threat to science are the unscientific decisions made by politicians on scientific issues.

5. Cultural exchange period in higher education is essential for world peace.

6. Allowing managers of public services like healthcare to use the same services from private or foreign providers is a disservice to the society.

Propositions belonging to the thesis, entitled

Impact of preselection in genomic evaluations

Ibrahim Jibrila

Wageningen, 25 May 2022

Impact of preselection in genomic evaluations

Ibrahim Jibrila

Thesis committee

Promotors
Prof. Dr R.F. Veerkamp
Special Professor of Numerical Genetics, Animal Breeding and Genomics Group
Wageningen University & Research

Dr M.P.L. Calus
Associate Professor, Animal Breeding and Genomics Group
Wageningen University & Research

Co-promotor
Dr J. ten Napel
Researcher, Animal Breeding and Genomics Group
Wageningen University & Research

Other members
Prof. Dr M.C.M. de Jong, Wageningen University & Research
Dr D. Lourenco, University of Georgia, Athens, USA
Prof. Dr G. Thaller, Christian-Albrecht University of Kiel, Kiel, Germany
Dr K. Peeters, Hendrix Genetics Research Centre B.V., Boxmeer, the Netherlands

This research was conducted under the auspices of the Graduate School Wageningen Institute of Animal Sciences (WIAS)

Impact of preselection in genomic evaluations

Ibrahim Jibrila

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board,
to be defended in public
on Wednesday May 25, 2022
at 4 p.m. in the Omnia Auditorium

Ibrahim Jibrila
Impact of preselection in genomic evaluations
144 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With summary in English

ISBN: 978-94-6447-102-1
DOI: https://doi.org/10.18174/563471

*I dedicate this thesis to my brother **Usman**, who gave me a second chance to live.*

Contents

Abstract

Jibrila, I. (2022). Impact of preselection in genomic evaluations. PhD thesis, Wageningen University, Wageningen, the Netherlands. ISBN: 978-94-6447-102-1, DOI: https://doi.org/10.18174/563471.

The development of genomic evaluation models over the last two decades has resulted in faster genetic improvement of animals, compared to when only pedigree-based genetic evaluation models were used. In large animal breeding programs, selection of parents of the next generation usually takes place in multiple stages, and the initial stages of this selection are collectively called preselection. Preselection takes place when selection candidates are young, sometimes even before they have records for any breeding goal trait. As the preselected animals grow older, they generally get records for more breeding goal traits, and they are re-evaluated in subsequent evaluations to select the final set of parents of the next generation. Impact of preselection on accuracy and bias of subsequent genomic evaluation of preselected animals is poorly understood. The same applies for the role of genotypes of preculled animals (i.e. animals removed from the breeding program at preselection stage) in subsequent genomic evaluation of their preselected sibs. In this thesis, I used single-step genomic best linear unbiased prediction (ssGBLUP) as the representative genomic evaluation model, and used simulated and real datasets to investigate the impact of i) types and intensities of preselection and ii) genotypes and phenotypes of different groups of animals, on accuracy and bias in ssGBLUP evaluation of preselected animals. I showed that preselection, regardless of its type and intensity, results in some accuracy loss in subsequent ssGBLUP evaluation of preselected animals, compared to a scenario without preselection. I explained that the accuracy loss is mainly due to loss of relatives with records. I also showed that ssGBLUP evaluates preselected animals without preselection bias, regardless of type and intensity of preselection. I further showed that genotypes of preculled animals are only needed in subsequent ssGBLUP evaluation of their genomically preselected sibs if some of their parents are not genotyped. The results of this thesis also showed that if ssGBLUP is used in subsequent evaluation of genomically preselected animals, realized genetic gain is only slightly lower compared to a scenario without preselection. To minimize this loss of genetic gain as a result of genomic preselection, I recommended that commercial animal breeding programs genotype as many young selection candidates as economically possible.

This page intentionally left blank

Chapter 1

General Introduction

1.1 Development of breeding value estimation models

Accurate and unbiased estimation of breeding values of animals is of utmost importance in animal breeding. This is to ensure that animals that best fit the breeding goal are as much as possible selected to produce the next generation, and genetic gain is estimated as accurately as possible. Traditionally, breeding values of animals were mainly estimated from phenotypes and pedigree relationships, using methods with a property known as best linear unbiased prediction (BLUP; e.g. Henderson et al., 1959; Henderson, 1975, 1976). In this thesis, I refer to BLUP utilizing only pedigree and phenotypes as pedigree-based BLUP (PBLUP), and the breeding values obtained therefrom as estimated breeding values (EBV). PBLUP uses a pedigree relationship matrix (**A**) to establish relationships among animals, thereby enabling related animals to benefit from records of each other in genetic evaluations. The inverse of the pedigree relationship matrix ($\mathbf{A}^{-1}$) or elements of $\mathbf{A}^{-1}$ are directly used in PBLUP. Efficient methods of building $\mathbf{A}^{-1}$ or its elements without having to firstly build **A** have been developed (e.g. Henderson, 1976; Misztal & Gianola, 1988), and this makes PBLUP generally easy to implement and to run fast. The main limitation of PBLUP is that it provides low accuracies for animals without own or progeny performance (e.g. Vitezica et al., 2011; Wolc et al., 2016).

With progress in DNA technology, large-scale genotyping of animals became affordable, and genomic information could be included in genetic evaluations of animals (Meuwissen *et al.,* 2001). In this thesis, I refer to genetic evaluation models utilizing genomic information as genomic evaluation models. Genomic evaluation models such as single nucleotide polymorphism BLUP (SNPBLUP; Meuwissen et al., 2001; Kolbehdari et al., 2007) and genomic BLUP (GBLUP; e.g. VanRaden, 2008; Hayes et al., 2009) evaluate animals mainly using only genotypes and phenotypes. SNPBLUP estimates breeding values from SNP effects, and such breeding values are usually called direct genomic values (DGV). On the other hand, GBLUP uses a genomic relationship matrix (**G**) in its estimation of breeding values, and such breeding values are usually called genomic EBV (GEBV). SNPBLUP and GBLUP perform equivalent genomic evaluations, and generally estimate breeding values more accurately than PBLUP, but have the limitation of estimating breeding values only for genotyped animals (e.g. VanRaden et al., 2009; Vitezica et al., 2011; Wolc et al., 2016). In this thesis, I use GEBV to refer to both DGV and GEBV, i.e. all breeding values estimated using genomic information.

In genomic evaluations, animals are divided into reference animals and selection candidates (e.g. Goddard & Hayes, 2007; Hayes et al., 2009). Reference animals are

usually from previous generations, already selected, and with genotypes and phenotypes available. Selection candidates on the other hand are usually young animals with genotypes but without phenotypes. Because it is still too expensive to genotype all animals in a breeding program, usually only better-than-average animals are genotyped. Sometimes, selection of animals to genotype is within families, and this results in biased GEBV (e.g. Vitezica et al., 2011; Wolc et al., 2016).

In the early years of genomic evaluation in dairy cattle, only males were genotyped. Milk production and fertility traits are among the most important traits in dairy cattle. Milk production traits and most fertility traits cannot be measured on males. However, each male has a mother, some sisters, and/or many daughters with records for these traits. Even for those traits that can be measured on males, such as stature, the males also have female relatives with these traits measured, and utilizing these records on relatives can increase accuracy of breeding values of the genotyped males. Since genomic evaluation models such as SNPBLUP and GBLUP could only evaluate genotyped animals, there was a need to find a way of utilizing records of ungenotyped relatives in the genomic evaluation of genotyped males. This gave rise to multi-step genomic evaluation, which involves the following steps (e.g. Goddard & Hayes, 2007; Hayes et al., 2009; VanRaden et al., 2009): 1. A PBLUP evaluation including all animals, where phenotypes of ungenotyped relatives are used to estimate EBV for genotyped males. 2. The EBV of the genotyped males are converted to pseudo-phenotypes. 3. Then genotypes and pseudo-phenotypes of the genotyped males are used in SNPBLUP or GBLUP to estimate GEBV for the genotyped males. 4. As multi-step genomic evaluation results in genotyped animals having two sets of breeding values, one set being pedigree based (EBV) and the other set being genomic based (GEBV), the two sets of breeding values need to be blended. The blending is often not straight forward due to differences in the assumptions that pedigree and genomic evaluation models make (Goddard & Hayes, 2007; Hayes et al., 2009; VanRaden et al., 2009).

To solve the above problems associated with genomic evaluation models like SNPBLUP and GBLUP, single-step genomic evaluation models were introduced (Legarra et al., 2009; Aguilar et al., 2010; Christensen & Lund, 2010). Single-step models utilize all available pedigree, genotypes and phenotypes in one analysis and provide GEBV for all animals, irrespective of whether the animals have genotypes and/or phenotypes. Single-step genomic evaluation models can be in the form of single-step GBLUP (ssGBLUP; an improvement over GBLUP) or single-step SNPBLUP (ssSNPBLUP; an improvement over SNPBLUP). ssGBLUP combines pedigree

relationships for both genotyped and ungenotyped animals (**A**) and genomic relationships for genotyped animals (**G**), into blended pedigree-genomic relationships for all animals (**H**), and estimates GEBV for all animals based on **H** (Misztal et al., 2009; Legarra et al., 2009). Because it is the inverse of **H** (i.e. $\mathbf{H}^{-1}$) that ssGBLUP uses and not **H** itself, methods to directly obtain $\mathbf{H}^{-1}$ without having to firstly make **H** have been developed (Aguilar et al., 2010; Christensen & Lund, 2010), and this has made ssGBLUP more efficient. On the other hand, ssSNPBLUP uses pedigree relationships between genotyped and ungenotyped animals to explicitly or implicitly impute genotypes for ungenotyped animals, and uses SNP effects to estimate GEBV for all animals (Fernando et al., 2014; Liu et al., 2014). It has been shown that single-step models produce more accurate and less biased breeding values than pedigree-based and genomic models, even when reference animals were selected (e.g. Misztal et al., 2013; Legarra et al., 2014). Single-step models are now commonly used in breeding programs for all livestock species. The basic presentation of BLUP procedure (including PBLUP, SNPBLUP, GBLUP, ssGBLUP and ssSNPBLUP) is explained in Box 1.

Box 1: Basic presentation and development of BLUP

The basic presentation of BLUP model in matrix notation is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Wp} + \mathbf{Zu} + \mathbf{e},$$

where **y** is the vector of phenotypes; **b** is the vector of fixed effects, with incidence matrix **X** linking observations to fixed effects; **p** is the vector of non-genetic random effects, with incidence matrix **W** linking observations to non-genetic random effects; **u** is the vector of estimated breeding values, with incidence matrix **Z** linking observations to estimated breeding values, and **e** is the vector of residuals.

1.2 Preselection

In animal breeding, selection of parents of the next generation usually takes place in two or more stages (e.g. Henderson, 1975; Appel et al., 1998; Schrooten et al., 2005; Janhunen et al., 2014), and the term 'preselection' is used to refer to the early stages

of this selection (e.g. Patry & Ducrocq, 2011a,b; Janhunen et al., 2014; Masuda et al., 2018). Preselection takes place in breeding programs for a number of reasons. For example, a small percentage of animals will be removed from breeding program at an early age, because these animals show some undesired traits such as lameness, deformity or dwarfism. Animal breeders therefore avoid the unnecessary cost of further raising such animals by removing them from the breeding program as early as possible. Another reason to perform preselection is to reduce phenotyping costs, and this is especially relevant for traits that are difficult or expensive to measure, e.g. feed intake of individual animals. For other traits, preselection is performed because cost of raising the animals until phenotyping is high, as the phenotypes can only be measured at advanced stages of life, e.g. litter size and other reproduction traits. Animal breeders therefore preselect the young animals that will be raised further and evaluated for these traits. Preselection can be based on raw phenotypes or breeding values for early-recorded breeding goal traits such as birth weight, leg soundness and number of teats, and this type of preselection is loosely called phenotypic preselection (PPS). Preselection can also be based on average parental breeding value for the entire breeding goal (i.e. parent average preselection, PAPS) or based on GEBV of selection candidates for the entire breeding goal (i.e. genomic preselection, GPS). Preselection can be at random, and is then called random preselection (RPS). RPS is rarely applied, and when applied it is mainly due to lack of informative criteria that can be used to preselect animals. Animals that survive preselection are called 'preselected animals' (e.g. Patry & Ducrocq, 2011a,b; Patry et al., 2013; Masuda et al., 2018). In this thesis, I refer to those animals that are removed from the breeding program at preselection stage as 'preculled animals'. Table 1.1 summarizes common types of preselection and their definitions.

Table 1.1 Common types of preselection and their definitions

Preselection type	Definition
Genomic preselection (GPS)	Preselection based on the genomic estimated breeding values of selection candidates for the entire breeding goal
Phenotypic preselection (PPS)	Preselection based on selection candidates' phenotypes or breeding values for early-recorded breeding goal trait(s)
Parent average preselection (PAPS)	Preselection based on selection candidates' average parental breeding values for the entire breeding goal
Random preselection (RPS)	Preselection performed randomly

In the pregenomic era, preselection was mainly PPS, based on traits such as body shape, leg soundness, birth and weaning weights, number of teats, and survival to a particular age. GPS is now popular among most commercial animal breeding programs. This popularity of GPS is undoubtedly related to the fact that genotyping

is becoming cheaper by the day, and the reasonable reliabilities achieved for GEBV (e.g. Hayes *et al.,* 2009, Su *et al.,* 2010). Because it is still too expensive to genotype all selection candidates, animals to be genotyped are still preselected mainly using PPS or PAPS. Then after genotyping, the remaining selection candidates are further preselected using GPS. Regardless of the type of preselection used and in how many stages preselection is implemented, preselection is always aimed at keeping animals with better-than-average genetic merit to be raised further for subsequent evaluations and final selection of parents of the next generation. Figure 1.1 illustrates how the distribution of true genetic merits (e.g. Mendelian sampling (MS) terms or breeding values) of selection candidates change from birth to final selection of parents of the next generation.

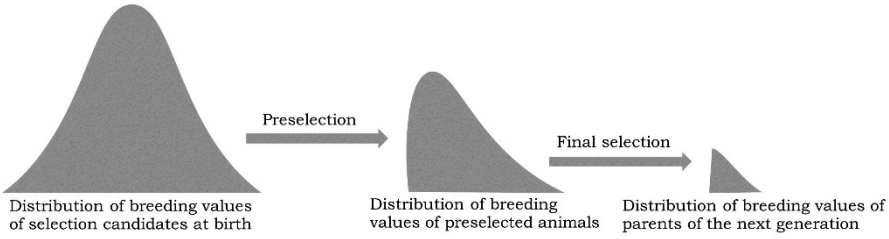


Figure 1.1 Distribution of breeding values of selection candidates from birth to final selection of parents of the next generation.

For a typical quantitative trait that undergoes positive selection, genetic merit of the young animals at birth are approximately normally distributed (as in the distribution at extreme left of the figure). When preselection is applied, depending on type and intensity of the preselection, the distribution of genetic merit of preselected animals looks more or less like the distribution in the middle of the figure (with a positive mean). When the final selection of the parents of the next generation is done, also depending on type and intensity of the selection, the distribution of genetic merit of the selected animals looks more or less like the distribution at the extreme right of the figure (with a large positive mean).

In Figure 1.1 and in this thesis as a whole, I represented all possible preselection stages across livestock species by only one preselection stage. For preselection to have impact on accuracy and bias in subsequent evaluation of preselected animals, the preselection needs to be able to skew the distribution of genetic merits of the preselected animals, as will be seen in Subsection 1.4. Skewing the distribution of genetic merits of selection candidates by the same magnitude is expected to result

in the same impact, regardless of the number of stages in which the preselection was implemented.

1.3 Subsequent genetic evaluation and selection of parents of the next generation

After preselection, the preselected animals are raised further until they have own or progeny records for some or more breeding goal traits. When these records become available, the preselected animals are re-evaluated, and parents of the next generation are then selected. In this thesis, we refer to this re-evaluation as subsequent genetic evaluation. In chickens and pigs, preselected animals go into performance testing, where recordings of production traits such as feed intake and average daily gain are made. After performance testing, the preselected animals are subsequently evaluated, and parents of the next generation selected. When the selected parents produce progeny, some more selection steps may be applied based on e.g. reproduction traits. In dairy cattle breeding programs where selection has mainly been in males, in the pregenomic era preselected young bulls underwent progeny testing. When daughters of these young bulls finished their first lactation, their records were used in the final evaluation of their progeny-tested fathers, where some of the bulls were kept and the remaining culled. Due to high reliability of GEBV, it is nowadays common to have GPS as the final selection stage in dairy cattle, and the genomically preselected young bulls end up as the sires of the next generation. In situations where GPS is the final selection stage, GPS is also referred to as genomic selection. However, even in the dairy breeding programs where GPS is the final selection stage, the (pre)selected young bulls are usually re-evaluated when their daughters have records, by comparing the bulls' GEBV to their daughter yield deviations or deregressed proofs (e.g. Mäntysaari et al., 2020). This re-evaluation is mainly done to check whether the genetic gain predicted from GEBV of the young bulls was correctly predicted.

1.4 Problem statement

Genomic preselection has been reported to cause a positive average MS term for the preselected animals (e.g. Patry & Ducrocq, 2011a; Sullivan, 2018; Tyrisevä et al., 2018). Because selection of parents of the next generation takes place in multiple stages, preselected animals are the selection candidates in subsequent genetic evaluation. Selection candidates having a positive average MS term is a violation of one of the assumptions of genetic evaluation models (i.e. that the expectation of the average MS of the observed offspring is zero). Impact of violation of this assumption

due to preselection has been a subject of study in animal breeding for a long time. In the pregenomic era, Henderson (1975) showed that following PPS, subsequent evaluations with PBLUP are biased, unless the subsequent evaluations are made multi-trait, with phenotypes of both the trait(s) on which preselection was conducted and subsequently recorded trait(s) included. In the genomic era, Patry & Ducrocq (2011a,b) and Patry et al. (2013) showed that following GPS, subsequent evaluations with PBLUP produce biased and less accurate EBV. Patry & Ducrocq (2011b) showed that the bias and accuracy losses caused by GPS can be prevented by including genomic information in the form of genomic pseudo-performances (e.g., deregressed genomic proofs) of both preselected and preculled animals in the subsequent evaluations with PBLUP.

Although genomic evaluation models are the genetic evaluation models of the present and likely of the future as well, impact of preselection on subsequent genomic evaluation of preselected animals is still not clearly understood. Aguilar et al. (2010) hypothesized that single-step genomic evaluation models should at least to some extent be able to overcome the impact of preselection, but so far no study has tested this hypothesis. There are also some unpublished reports coming from commercial animal breeding programs attributing some of the bias observed from subsequent single-step genomic evaluations to preselection. It is important to know whether and to what extent preselection affects the ability of genomic evaluation models to accurately and unbiasedly evaluate preselected animals. This thesis will investigate the accuracy loss and bias in subsequent genomic evaluation of preselected animals that is attributable to preselection.

1.5 Objectives of this thesis

The main aim of this thesis was to investigate the impact of preselection on accuracy and bias in subsequent genomic evaluations. This main aim was split into two research questions. The first question was what is the impact of preselection on accuracy and bias in subsequent genomic evaluation of preselected animals? As preculled animals usually have neither progeny nor phenotypes for subsequently recorded and evaluated traits, they are by default not included in subsequent genomic evaluation. So, the second research question was what is the impact of including (or excluding) genotypes of preculled animals in subsequent genomic evaluations. Throughout the thesis, I used ssGBLUP as genomic evaluation model, as it is now the commonest genomic evaluation model used in animal breeding.

1.6 Outline of this thesis

A number of factors may modify the impact of preselection on accuracy and bias in subsequent evaluation of preselected animals. Such factors studied in this thesis, and in which chapter they were studied, are listed in Table 1.2.

Table 1.2 Possible modifiers of impact of preselection and where they were covered in this thesis

Possible modifier of impact of preselection	Covered in Chapter
Type of preselection	2
Intensity of preselection	2, 4, 5
Heritability	2, 4, 5
Trait weight in the breeding goal	4
Whether preselection is in a single generation or in multiple generations	4, 5
Whether the traits studied are widely-recorded or scarcely-recorded	4, 5
Whether records are available or not on selection candidates at the time of subsequent evaluation	4, 5

I used simulated datasets in Chapters 2 and 3. In Chapter 2, I investigated, for various heritabilities, the impact of intensity and type of preselection on accuracy and bias in subsequent ssGBLUP evaluation of preselected animals. In Chapter 3, I investigated the roles of genotypes and phenotypes from various groups of animals in preventing bias due to preselection, when estimating GEBV of genomically preselected animals in subsequent ssGBLUP evaluation. This was done to establish the minimum information required in subsequent ssGBLUP evaluation to estimate GEBV of genomically preselected animals without bias associated with preselection.

In Chapters 4 and 5, I used real datasets to verify whether what I found using simulated datasets holds in reality as well. In Chapter 4, I investigated the impact of genomic preselection on accuracy and bias in subsequent ssGBLUP evaluation, using real data from a commercial pig breeding program. I studied widely-recorded traits in Chapter 4 (i.e. traits that are recorded on most animals that survive to the point of recording such traits). As opposed to widely-recorded traits, there are traits that are only recorded on a small number of animals that survive to the point of recording such traits, and those are called scarcely-recorded traits. So in Chapter 5, I investigated the impact of genomic preselection on accuracy and bias in subsequent ssGBLUP evaluation of animals for scarcely-recorded traits. In Chapter 6 (General Discussion), I made inferences on situations not directly covered in my thesis, and discussed the implications of my thesis results for the animal breeding industry.

1.7 References

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic

information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2), 743–752. https://doi.org/10.3168/jds.2009-2730

Appel, L. J., Strandberg, E., Danell, B., & Lundeheim, N. (1998). Adjusting for Missing Data Due to Culling before Testing in Genetic Evaluations of Swine. *Journal of Animal Science*, 76(7), 1794–1802. https://doi.org/10.2527/1998.7671794x

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, 42(2). https://doi.org/https://doi.org/10.1186/1297-9686-42-2

Fernando, R. L., Dekkers, J. C. M., & Garrick, D. J. (2014). A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics Selection Evolution*, 46(50). https://doi.org/10.1186/1297-9686-46-50

Goddard, M. E., & Hayes, B. J. (2007). Genomic selection. *Journal of Animal Breeding and Genetics*, 124, 3794–3805. https://doi.org/10.1016/b978-0-12-820595-2.00010-2

Hayes, B.J., Bowman, P. J., Chamberlain, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, 92(2), 433–443. https://doi.org/10.3168/jds.2008-1646

Hayes, Ben J., Visscher, P. M., & Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research*, 91(1), 47–60. https://doi.org/10.1017/S0016672308009981

Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometris*, 31(2), 423–447. https://doi.org/10.2307/2529430

Henderson, C. R. (1976). A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. *Biometrics*, 32(1), 69–83. https://doi.org/10.2307/2529339

Henderson, C. R., Kempthorne, O., Searle, S. R., & Krosigk, C. M. (1959). The Estimation of Environmental and Genetic Trends from Records Subject to Culling. *Biometrics*, 15(2), 192–218. https://doi.org/10.2307/2527669

Janhunen, M., Kause, A., Vehviläinen, H., Nousiainen, A., & Koskinen, H. (2014). Correcting within-family pre-selection in genetic evaluation of growth-A simulation study on rainbow trout. *Aquaculture*, 434, 220–226. https://doi.org/10.1016/j.aquaculture.2014.08.020

Kolbehdari, D., Schaeffer, L. R., & Robinson, J. A. B. (2007). Estimation of genome-wide haplotype effects in half-sib designs. *Journal of Animal Breeding and Genetics*, 124(6), 356–361. https://doi.org/10.1111/j.1439-0388.2007.00698.x

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, 92(9), 4656–4663. https://doi.org/10.3168/jds.2009-2061

Legarra, Andres, Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, 166(1), 54–65. https://doi.org/10.1016/j.livsci.2014.04.029

Liu, Z., Goddard, M. E., Reinhardt, F., & Reents, R. (2014). A single-step genomic

model with direct estimation of marker effects. *Journal of Dairy Science*, 97(9), 5833–5850. https://doi.org/10.3168/jds.2014-7924

Mäntysaari, E. A., Koivula, M., & Strandén, I. (2020). Symposium review: Single-step genomic evaluations in dairy cattle. *Journal of Dairy Science*, 103(6), 5314–5326. https://doi.org/10.3168/jds.2019-17754

Masuda, Y., VanRaden, P. M., Misztal, I., & Lawlor, T. J. (2018). Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. *Journal of Dairy Science*, 101(6), 5194–5206. https://doi.org/10.3168/jds.2017-13310

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Map. *Genetics*, 157(4), 1819–1829. https://doi.org/11290733

Misztal, I., Legarra, A., & Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*, 92(9), 4648–4655. https://doi.org/10.3168/jds.2009-2064

Misztal, I, & Gianola, D. (1988). Indirect Solution of Mixed Model Equations. *Journal of Dairy Science*, 71(Supplementary 2), 99–106. https://doi.org/https://doi.org/10.1016/S0022-0302(88)79987-7

Misztal, Ignacy, Aggrey, S. E., & Muir, W. M. (2013). Experiences with a single-step genome evaluation. *Poultry Science*, 92(9), 2530–2534. Retrieved from https://doi.org/10.3382/ps.2012-02739

Patry, C., & Ducrocq, V. (2011). Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. *Journal of Dairy Science*, 94(2), 1011–1020. https://doi.org/10.3168/jds.2010-3804

Patry, Clotilde, & Ducrocq, V. (2011). Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. *Genetics Selection Evolution*, 43(30). https://doi.org/10.1186/1297-9686-43-30

Patry, Clotilde, Jorjani, H., & Ducrocq, V. (2013). Effects of a national genomic preselection on the international genetic evaluations. *Journal of Dairy Science*, 96(5), 3272–84. Retrieved from https://doi.org/10.3168/jds.2011-4987

Schrooten, C., Bovenhuis, H., van Arendonk, J. A. M., & Bijma, P. (2005). Genetic Progress in Multistage Dairy Cattle Breeding Schemes Using Genetic Markers. *Journal of Dairy Science*, 88(4), 1569–1581. https://doi.org/10.3168/jds.S0022-0302(05)72826-5

Su, G., Guldbrandtsen, B., Gregersen, V. R., & Lund, M. S. (2010). Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. *Journal of Dairy Science*, 93(3), 1175–1183. https://doi.org/10.3168/jds.2009-2192

Sullivan, P. G. (2018). Mendelian Sampling variance tests with genomic preselection. *INTERBULL BULLETIN*, 53.

Tyrisevä, A.-M., Mäntysaari, E. A., Jakobsen, J., Aamand, G. P., Dürr, J., Fikse, W. F., & Lidauer, M. H. (2018). Detection of evaluation bias caused by genomic

preselection. *Journal of Dairy Science*, 101(4), 3155–3163. https://doi.org/10.3168/jds.2017-13527

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., & Schenkel, F. S. (2009). Invited Review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science*, 92(1), 16–24. https://doi.org/10.3168/jds.2008-1514

Vitezica, Z. G., Aguilar, I., Misztal, I., & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, 93(5), 357–366. https://doi.org/10.1017/S001667231100022X

Wolc, A., Kranis, A., Arango, J., Settar, P., Fulton, J. E., O'Sullivan, N. P., … Dekkers, J. C. M. (2016). Implementation of genomic selection in the poultry industry. *Animal Frontiers*, 6(1), 23–31. https://doi.org/10.2527/af.2016-0004

Chapter 2

Investigating the impact of preselection on subsequent single-step genomic BLUP evaluation of preselected animals

Ibrahim Jibrila, Jan ten Napel, Jeremie Vandenplas, Roel F Veerkamp and Mario P L Calus

Wageningen University and Research Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708PB Wageningen, the Netherlands

Genet Sel Evol 52, 42 (2020)

Abstract

Background

Preselection of candidates, hereafter referred to as preselection, is a common practice in breeding programs. Preselection can cause bias and accuracy loss in subsequent pedigree-based best linear unbiased prediction (PBLUP). However, the impact of preselection on subsequent single-step genomic BLUP (ssGBLUP) is not completely clear yet. Therefore, in this study, we investigated, across different heritabilities, the impact of intensity and type of preselection on subsequent ssGBLUP evaluation of preselected animals.

Methods

We simulated a nucleus of a breeding program, in which a recent population of 15 generations was produced with PBLUP-based selection. In generation 15 of this recent population, the parents of the next generation were preselected using several preselection scenarios. These scenarios were combinations of three intensities of preselection (no, high or very high preselection) and three types of preselection (genomic, parental average or random), across three heritabilities (0.5, 0.3 or 0.1). Following each preselection scenario, a subsequent evaluation was performed using ssGBLUP by excluding all the information from the preculled animals, and these genetic evaluations were compared in terms of accuracy and bias for the preselected animals, and in terms of realized genetic gain.

Results

Type of preselection affected selection accuracy at both preselection and subsequent evaluation stages. While preselection accuracy decreased, accuracy in the subsequent ssGBLUP evaluation increased, from genomic to parent average to random preselection scenarios. Bias was always negligible. Genetic gain decreased from genomic to parent average to random preselection scenarios. Genetic gain also decreased with increasing intensity of preselection, but only by a maximum of 0.1 additive genetic standard deviation from no to very high genomic preselection scenarios.

Conclusions

Using ssGBLUP in subsequent evaluations prevents preselection bias, irrespective of intensity and type of preselection, and heritability. With GPS, in addition to reducing the phenotyping effort considerably, the use of ssGBLUP in subsequent evaluations realizes only a slightly lower genetic gain than that realized without preselection. This is especially the case for traits that are expensive to measure (e.g. feed intake of individual broiler chickens), and traits for which phenotypes can only be measured at advanced stages of life (e.g. litter size in pigs)

2.1 Background

Selection of the parents of the next generation usually takes place in two or more stages (e.g.[1–3]), and the term 'preselection' is used to refer to the early stages of selection (e.g. [3–5]). Preselection is a common practice in the nuclei of breeding programs, where only a few hundred to a few thousand replacement animals are required per generation. In order to have a large pool of animals to select from, many more young animals are produced than the numbers required for producing the next generation. Preselection is done for different reasons for different traits. For traits that are difficult or expensive to measure (e.g. feed intake of individual broiler chickens), preselection is used to reduce phenotyping costs. For traits for which phenotypes can be measured only at advanced stages of life (e.g. litter size in pigs), preselection is used to reduce the cost of raising the animals until phenotyping. Traditionally, preselection has mostly been based on correlated trait(s) that can be measured easily and cheaply early in life (e.g. [1, 3, 6–8]). In the genomic era, preselection is often based on genomic estimated breeding values (GEBV) of young selection candidates, and in the literature this type of preselection is called genomic or genotypic preselection (GPS; e.g. [4, 5, 9]).

Before the introduction of genomic prediction [10], models for the genetic evaluation of animals were based on phenotypic and pedigree data. These models are generally easy to implement and run fast, but their limitation is that they provide low accuracies for animals without own phenotype (e.g.[11, 12]). With the progress in DNA technology, large-scale genotyping of animals became affordable and genomic information can now be included in genetic evaluations of animals, e.g. by using multi-step genomic evaluation models, where genomic and pedigree information are used in two separate steps [13]. Generally, multi-step genomic evaluation models estimate breeding values more accurately than pedigree-based models, but have the disadvantage of estimating breeding values for genotyped animals only (e.g.[11, 12]). Because the required reference population (animals with genotypes and phenotypes) for multi-step models are usually already selected, the breeding values obtained are biased (e.g.[11, 12]). In 2010, single-step genomic evaluation models were introduced as improvements over both pedigree-based and multi-step genomic models [14, 15]. Single-step models combine all available pedigree, genomic and phenotypic information and provide GEBV for all the animals regardless of whether the animals have phenotypes and/or genotypes. It has been shown that single-step models produce more accurate and less biased breeding values than pedigree-based and multi-step genomic models, even in the presence of

selective genotyping and phenotyping (e.g. [16, 17]).

Preselection is known to result in a positive average Mendelian sampling (MS) term for the selected animals (e.g. [4, 18, 19]). Selection candidates that have a positive average MS term represent a violation of one of the assumptions of genetic evaluation models (i.e. that the expectation of the average MS of the observed offspring is zero). This has been reported to result in biased and less accurate estimated breeding values (EBV) in subsequent evaluations that are done using pedigree-based best linear unbiased prediction (PBLUP, e.g. [4, 18–22]). It is also known that when all the information on which preselection is based is included in the subsequent PBLUP evaluations, the impact of the violation of this assumption is usually alleviated, e.g. [20–22].

Single-step genomic BLUP (ssGBLUP) has been reported to handle GPS better than PBLUP. For example, Masuda et al. [5] reported lower genetic trends in milk, fat, and protein yields in genomically-preselected US Holsteins when the subsequent evaluations were performed with PBLUP than with ssGBLUP. These authors ([5]) used these differences in genetic trends between PBLUP and ssGBLUP as an evidence of preselection bias in PBLUP evaluations following GPS. Although Aguilar et al. [14] hypothesized that ssGBLUP could completely prevent preselection bias, to date, there is no study in the literature that compared results using the same data with and without preselection to investigate this hypothesis. The study of Masuda et al. [5] evaluated preselection bias in subsequent PBLUP and ssGBLUP evaluations, but did not include a scenario based on the complete data (without preselection), with which the other scenarios could be compared. Furthermore, the benefit of including the genotypes of the selection candidates discarded at the preselection stage - hereafter referred to as preculled animals - in subsequent ssGBLUP evaluations is still not clear. On the one hand, Shabalina et al. [23] concluded that including the genotypes of preculled animals in subsequent ssGBLUP evaluations improves accuracy in situations where (some of the) parents of the genotyped selection candidates are not genotyped. On the other hand, Koivula et al. [24] reported larger biases and losses in reliability in subsequent ssGBLUP evaluations when genotypes of preculled animals were included and most of the parents of the selection candidates were genotyped. Thus, our aim was to investigate the impact of preselection on subsequent evaluations of preselected animals, using ssGBLUP with all the information from the preculled animals excluded.

2.2 Methods

2.2.1 Data simulation

To achieve our aim, we simulated a nucleus of a breeding program with inputs from the international breeding companies that operate in the Netherlands, using QMSim [25]. The QMSim parameter file, with all the details of the simulation, are in Additional file 2.1. For each animal in the breeding program, a genome of 30 chromosomes each 100 cM long was simulated. Sixty thousand single nucleotide polymorphisms (SNPs) and 3000 quantitative trait loci (QTL) were evenly distributed across the entire genome, and the QTL effects were randomly drawn from a gamma distribution with a shape parameter of 0.4. The simulation started with a historical population, to establish mutation-drift equilibrium and linkage disequilibrium among markers and QTL. The historical population had 3000 generations of random mating, starting with 2500 female and 2500 male animals (both sexes were equally represented throughout the simulation). The size of the historical population decreased linearly until it reached 50 animals at generation 2997, and then increased and reached 5000 animals again at generation 3000. The founder population, which comprised 100 males and 1000 females, was randomly selected from the 3000th historical generation. Then, from this founder population, 15 (recent) generations of artificial selection were simulated. In each generation, 100 males and 1000 females were selected and mated to produce the next generation of 16,000 animals. Within sex, all selected parents contributed equally to the next generation. Selection was based on EBV, and the mating design aimed at minimizing inbreeding by using minimum co-ancestry matings as described in [26], which minimize the average relationship among all sires and dams, and therefore also among their offspring. There was no preselection during the production of these 15 generations, thus information on all the animals (including the culled animals) was used to inform selection decisions. The breeding goal consisted of a single quantitative trait that was measured in both sexes. Simulations were carried out with heritabilities of 0.5, 0.3 and 0.1, to represent breeding goal traits with high, medium and low heritabilities, respectively. Pedigree of all animals (from generations 0 to 15), genotypes of all animals in generations 13 to 15 and phenotypes of all animals in generations 11 to 15 were used in this study.

2.2.2 Implementation of preselection

Preselection was implemented in generation 15 by performing several scenarios, which were combinations of three intensities of preselection and three types of preselection, across the three simulated heritabilities. An overview of these

preselection scenarios is in Table 2.1. The three intensities of preselection investigated were no preselection (control), high preselection, and very high preselection. With no preselection, all the selection candidates (animals produced in generation 15) were kept until the subsequent genetic evaluation; thus this scenario mimicked single-stage selection. With high preselection, 10% of the male and 15% of the female selection candidates were preselected. With very high preselection, 5% of the male and 12.5% of the female selection candidates were preselected. The choice of these intensities of preselection was informed by the information that we obtained from the international breeding companies operating in the Netherlands. The three types of preselection were GPS, parent average preselection (PAPS) and random preselection (RPS). Details of the information used in each preselection type are in Table 2.2. Briefly, with GPS, GEBV of the selection candidates were used, which were estimated by ssGBLUP, with the phenotypes of the selection candidates excluded from the model. With PAPS, average parental GEBV of the selection candidates were used, which were estimated by ssGBLUP, with the genotypes and phenotypes of the selection candidates excluded from the model. As the name implies, RPS preselects the selection candidates randomly, and in this study, we used it to investigate the impact of reducing the number of selection candidates per se. The GEBV used in performing preselection in all scenarios of GPS and PAPS were estimated by the ssGBLUP procedure of MiXBLUP [27].

2.2.3 Subsequent genetic evaluation

Following each preselection scenario, we performed a subsequent genetic evaluation with ssGBLUP. The subsequent evaluations included pedigree information of all the animals from generation 0 to preselected generation 15, genotypes of all the animals from generation 13 to preselected generation 15 and phenotypes of all the animals from generation 11 to preselected generation 15. This means that no information from the preculled animals was used in the subsequent evaluations. These (subsequent) evaluations provided the breeding values that were used to finally select the 100 males and 1000 females in generation 15 that become the parents of the next generation. MiXBLUP [27] was also used in these (subsequent) evaluations. Each step (simulation of the breeding program, implementation of preselection and subsequent genetic evaluations) was replicated 10 times.

Table 2.1 Overview of the various preselection scenarios implemented

Heritability of the breeding goal trait	Type of preselection	Intensity of preselection	Scenario name
0.5	-	No[a]	No preselection with high heritability
0.5	Genomic	High[b]	High genomic preselection with high heritability
0.5	Genomic	Very high[c]	Very high genomic preselection with high heritability
0.3	-	No[a]	No preselection with medium heritability
0.3	Genomic	High[b]	High genomic preselection with medium heritability
0.3	Genomic	Very high[c]	Very high genomic preselection with medium heritability
0.1	-	No[a]	No preselection with low heritability
0.1	Genomic	High[b]	High genomic preselection with low heritability
0.1	Genomic	Very high[c]	Very high genomic preselection with low heritability
0.1	Parental average	High[b]	High parental average preselection with low heritability
0.1	Parental average	Very high[c]	Very high parental average preselection with low heritability
0.1	Random	High[b]	High random preselection with low heritability
0.1	Random	Very high[c]	Very high random preselection with low heritability

[a] No preselection: all selection candidates were kept until the subsequent genetic evaluation. [b] High preselection: 10% of the male and 15% of the female selection candidates were preselected. [c] Very high preselection: 5% of the male and 12.5% of the female selection candidates were preselected.

Table 2.2 Details of the information used in the different types of preselection

Type of preselection	Preselection was based on	Information used in preselection model
Genomic	GEBV of the selection candidates[a]	Complete pedigree[b], genotypes of all the animals in generations 13 to 15, phenotypes of all the animals in generations 11 to 14
Parent average	Average parental GEBV of the selection candidates[a]	Complete pedigree[b], genotypes of all the animals in generations 13 and 14, phenotypes of all the animals in generations 11 to 14
Random	Random	Random numbers

[a] Selection candidates were the animals in generation 15. [b] The complete pedigree consisted of all the animals from the founder generation (generation 0) to the most recent generation (generation 15).

2.2.4 Implementation of single-step GBLUP

In order to make sure that any observed bias and loss in accuracy in our results were due to preselection, all other known possible sources of bias and loss in accuracy in ssGBLUP evaluations were accounted for. Thus, the inverse of our combined pedigree-genomic relationship matrix ($\mathbf{H^{-1}}$) was as follows:

$$\mathbf{H^{-1}} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & (0.9\mathbf{G_t} + 0.1\mathbf{A}_{22})^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{A}^{-1}$ is the inverse of pedigree relationship matrix, and $\mathbf{A}_{22}$ is the pedigree relationship matrix among genotyped animals. To avoid the bias that is caused by not considering inbreeding in the construction of $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$ [28], we considered inbreeding in both $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$, and the inbreeding coefficients were calculated using the algorithm of Meuwissen and Luo [29]. $\mathbf{G_t}$ is the adjusted genomic relationship matrix that was obtained according to the F_{ST} method described by Powell et al. [30] and Vitezica et al. [12] and aimed at setting the average genomic inbreeding equal to the average pedigree inbreeding as follows:

$$\mathbf{G_t} = \left(1 - \bar{f}_p\right)\mathbf{G_r} + 2\bar{f}_p\mathbf{J},$$

where $\bar{f}_p$ is the average pedigree inbreeding coefficient across genotyped animals, $\mathbf{G_r}$ is the raw genomic relationship matrix computed following the first method of VanRaden [31], and $\mathbf{J}$ is a matrix of 1s. To obtain $\mathbf{G_r}$, we calculated allele frequencies using all the available genotypic data, and set the minor allele frequency threshold at 0.005.

The additive genetic and residual variances supplied to MiXBLUP (per heritability, per replicate) were estimated by fitting an animal model in ASReml [32]. To obtain these variances, we used the pedigree of all the animals in generations 0 to 14 and the phenotypes of all the animals in generations 11 to 14 (i.e. the available pedigree and phenotypic information at the time the selection candidates were born). The full MiXBLUP instruction file for the ssGBLUP analysis is included in Additional file 2.2.

2.2.5 Indicators of model performance across preselection scenarios

The following indicators of model performance were estimated for each preselection scenario and compared among the scenarios.

2.2.5.1 (Pre)selection accuracy

Accuracy was calculated as the correlation between (G)EBV and true breeding values (TBV). After running the preselection model, preselection accuracy was calculated based on all the selection candidates, whereas after running the subsequent genetic evaluation model, the subsequent selection accuracy was computed based only on the preselected animals.

2.2.5.2 Bias

Bias was measured in two ways. First, the absolute bias was calculated as the difference between mean TBV and mean (G)EBV of all the preselected animals, and expressed in additive genetic standard deviation (SD) units. If there is no absolute bias, the difference is 0. A negative difference means on average (G)EBV overestimate TBV, and a positive difference means that on average (G)EBV underestimate TBV. In order to make TBV comparable to (G)EBV, we subtracted the mean TBV and the mean (G)EBV of the animals in generations 11 to 14 from the TBV and the (G)EBV of each of the preselected animals, respectively. Second, dispersion bias was measured as the regression coefficient of TBV on (G)EBV ($b_{TBV,(G)EBV}$) of all preselected animals. If there is no dispersion bias, $b_{TBV,(G)EBV}$ is 1. A value of $b_{TBV,(G)EBV}$ lower than 1 means that variance of (G)EBV is inflated compared to variance of TBV, and a value of $b_{TBV,GEBV}$ higher than 1 means that variance of (G)EBV is deflated compared to variance of TBV.

2.2.5.3 Realised genetic gain (RGG)

The realised genetic gain (RGG) is the difference between the average TBV of the selected individuals in two subsequent generations, provided that each of the selected animal (per sex) contributes equally to the next generation. In this study, RGG is the difference between the average TBV of the 100 males and 1000 females that were subsequently selected in generation 15 and the average TBV of the 100 males and 1000 females selected in generation 14. For each generation, we computed averages within selected males and females, separately, and then took the average of these two averages. Here, we assumed that just as in the previous generations, all the subsequently selected animals of generation 15 would have equal contributions (per sex) to the next generation. To give RGG a reference point, it was expressed in units of additive genetic SD. In reality, RGG is estimated using (G)EBV, because TBV are not known. Any bias in (G)EBV could lead to bias in estimated RGG. Thus, we calculated RGG based on (G)EBV as well. These two parameters were named true realised genetic gain (TRGG) and estimated realised genetic gain (ERGG), respectively.

2.3 Results

Results of the genetic evaluations in which ssGBLUP was used in the subsequent evaluations are in Tables 2.3 and 2.4. The results in Table 2.3 are from the evaluations that were obtained with different intensities of GPS and different heritabilities. The results in Table 2.4 are from the evaluations that were obtained with different

intensities and types of preselection, all with a heritability of 0.1.

Table 2.3 ssGBLUP performance [a], with different heritabilities and GPS[b] intensities

Measure/intensity of GPS	No[c]	High[d]	Very high[e]
Heritability of 0.5			
Preselection accuracy[f]	Not applicable	0.81 (0.81-0.81)	0.81 (0.81-0.81)
Subsequent selection accuracy[g]	0.88 (0.88-0.88)	0.66 (0.64-0.68)	0.65 (0.63-0.67)
Absolute bias[h]	0.01 (0.01-0.01)	0.03 (0.01-0.05)	0.04 (0.02-0.06)
Dispersion bias[i]	1.02 (1.02-1.02)	1.06 (1.04-1.08)	1.05 (1.03-1.07)
True realized genetic gain[j]	1.53 (1.49-1.57)	1.48 (1.44-1.52)	1.45 (1.41-1.49)
Estimated realized genetic gain[k]	1.50 (1.48-1.52)	1.45 (1.43-1.47)	1.43 (1.41-1.45)
Heritability of 0.3			
Preselection accuracy[f]	Not applicable	0.78 (0.78-0.78)	0.78 (0.78-0.78)
Subsequent selection accuracy[g]	0.86 (0.86-0.86)	0.59 (0.57-0.61)	0.58 (0.56-0.60)
Absolute bias[h]	0.01 (0.01-0.01)	0.02 (0.00-0.04)	0.02 (0.00-0.04)
Dispersion bias[i]	1.01 (1.01-1.01)	1.01 (0.97-1.05)	0.98 (0.94-1.02)
True realized genetic gain[j]	1.46 (1.42-1.50)	1.39 (1.35-1.43)	1.37 (1.33-1.41)
Estimated realized genetic gain[k]	1.45 (1.41-1.49)	1.41 (1.37-1.45)	1.39 (1.35-1.43)
Heritability of 0.1			
Preselection accuracy[f]	Not applicable	0.71 (0.69-0.73)	0.71 (0.69-0.73)
Subsequent selection accuracy[g]	0.80 (0.78-0.82)	0.48 (0.44-0.52)	0.48 (0.44-0.52)
Absolute bias[h]	0.01 (0.01-0.01)	0.03 (0.01-0.05)	0.03 (0.01-0.05)
Dispersion bias[i]	1.02 (1.00-1.04)	1.01 (0.95-1.07)	1.00 (0.94-1.06)
True realized genetic gain[j]	1.38 (1.32-1.44)	1.26 (1.18-1.34)	1.24 (1.16-1.32)
Estimated realized genetic gain[k]	1.36 (1.30-1.42)	1.28 (1.22-1.34)	1.26 (1.20-1.32)

[a] Results shown only for selection candidates in the most recent generation (i.e. generation 15), and the results are means of 10 replicates (and 95% confidence intervals). [b] Genomic preselection. [c] No preselection. [d] 10% of the male and 15% of the female selection candidates were preselected. [e] 5% of the males and 12.5% of the female selection candidates were preselected. [f] Correlation between true and genomic estimated breeding values of all selection candidates. [g] Correlation between true and genomic estimated breeding values of all preselected animals' [h] Difference between average true breeding value and average genomic estimated breeding value of preselected animals, expressed in additive genetic standard deviation. [i] Coefficient of the regression of true on genomic estimated breeding values of the preselected animals. [j] Difference between average true breeding value of the subsequently selected animals and average true breeding value of their parents, expressed in additive genetic standard deviation. [k] Difference between average genomic estimated breeding value of the subsequently selected animals and average genomic estimated breeding value of their parents, expressed in additive genetic standard deviation.

Table 2.4 ssGBLUP performance [a], with different preselection types[b] and intensities[c], all with a heritability of 0.1

Measure	Type[b] and intensity[c] of preselection						
	No preselection	High GPS	Very high GPS	High PAPS	Very high PAPS	High RPS	Very high RPS
Preselection accuracy[d]	Not applicable	0.71 (0.69-0.73)	0.71 (0.69-0.73)	0.44 (0.42-0.46)	0.44 (0.42-0.46)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Subsequent selection accuracy[e]	0.80 (0.78-0.82)	0.48 (0.44-0.52)	0.48 (0.44-0.52)	0.69 (0.67-0.71)	0.68 (0.66-0.70)	0.73 (0.71-0.75)	0.73 (0.73-0.73)
Absolute bias[f]	0.01 (-0.01-0.03)	0.03 (0.01-0.05)	0.03 (0.01-0.05)	0.03 (0.01-0.05)	0.04 (0.02-0.06)	0.05 (0.01-0.09)	0.05 (0.01-0.09)
Dispersion bias[g]	1.02 (1.00-1.04)	1.01 (0.95-1.07)	1.00 (0.94-1.06)	1.01 (0.99-1.03)	1.01 (0.97-1.05)	1.01 (0.97-1.05)	1.02 (0.98-1.06)
True realized genetic gain[h]	1.38 (1.32-1.44)	1.26 (1.18-1.34)	1.24 (1.16-1.32)	1.11 (1.03-1.19)	1.00 (0.92-1.08)	0.58 (0.54-0.62)	0.39 (0.37-0.41)
Estimated realized genetic gain[i]	1.36 (1.30-1.42)	1.28 (1.22-1.34)	1.26 (1.20-1.32)	1.12 (1.06-1.18)	1.01 (0.95-1.07)	0.58 (0.56-0.60)	0.37 (0.35-0.39)

[a] Results shown only for the selection candidates in the most recent generation (i.e. generation 15), and the results are means over 10 replicates (and 95% confidence intervals). [b] Types of preselection: GPS - genomic preselection; PAPS - parent average preselection; RPS – random preselection. [c] Intensities of preselection: No preselection; high preselection - 10% of the male and 15% of the female selection candidates preselected; very high preselection - 5% of the male and 12.5% of the female selection candidates preselected. [d] Correlation between true and genomic estimated breeding values of all the selection candidates. [e] Correlation between true and genomic estimated breeding values of the preselected animals. [f] Difference between average true breeding value and average genomic estimated breeding value of preselected animals, expressed in additive genetic standard deviation. [g] Coefficient of regression of true on genomic estimated breeding values of the preselected animals. [h] Difference between average true breeding value of subsequently selected animals and average true breeding value of their parents, expressed in additive genetic standard deviation. [i] Difference between average genomic estimated breeding value of the subsequently selected animals and average genomic estimated breeding value of their parents, expressed in additive genetic standard deviation.

2.3.1 (Pre)selection accuracy

2.3.1.1 Preselection accuracy: Within the same heritability and type of preselection, preselection accuracy was the same for the high and very high intensities of preselection (Tables 2.3 and 2.4). GPS provided a higher preselection accuracy (0.71) than PAPS (0.44), and as expected, RPS provided a preselection accuracy equal to zero (Table 2.4).

2.3.1.2 Subsequent selection accuracy: For a given heritability, subsequent selection accuracy was always highest without preselection. It decreased with preselection (ranging from 0.80 to 0.48 for the scenarios with a heritability of 0.1), but within the same type of preselection, it remained similar across high and very high intensities of preselection (Tables 2.3 and 2.4). For a given heritability, subsequent selection accuracy increased from GPS to PAPS, and from PAPS to RPS (Table 2.4).

2.3.2 Bias

Both absolute and dispersion bias were always numerically very small, and often not statistically significant. The highest observed absolute bias was 0.05 genetic SD units, and the highest deviation of the $b_{TBV,GEBV}$ from 1 (indicator of dispersion bias) was 0.06. Thus, the impacts of intensity of preselection and type of preselection on bias are considered negligible across all heritabilities.

2.3.3 Realised genetic gain

With the same heritability and type of preselection, RGG (both TRGG and ERGG) always decreased with increasing intensity of preselection (Tables 2.3 and 2.4). With the same intensity of preselection, RGG decreased from GPS to PAPS and from PAPS to RPS (Table 2.4), and ranged from 0.39 to 1.38 genetic SD (TRGG) and 0.37 to 1.36 genetic SD (ERGG) for the scenarios with heritability of 0.1. Irrespective of intensity of preselection, type of preselection, and heritability, ERGG was never statistically different from its corresponding TRGG (Tables 2.3 and 2.4).

2.4 Discussion

In this study, we investigated, for different heritabilities, the impact of intensity and type of preselection on the subsequent evaluation of preselected animals in terms of selection accuracy, bias and genetic gain, using ssGBLUP with all the information from preculled animals excluded. We implemented only one stage of preselection and only one type of preselection at a time, to clearly identify the impact of each type and intensity of preselection. However, in reality, most breeding programs

involve at least two stages of preselection, i.e. a first preselection of elite families using PAPS and then genotyping some members of these elite families for performing GPS. In addition, female selection candidates may not be genotyped in all cases. It is expected that, in the near future, genotyping costs will become so cheap that breeding companies will decide to genotype all their selection candidates [33]. In addition, based on our findings (i.e. that GPS hardly leads to any significant loss of genetic gain whereas PAPS does), breeding companies may become more inclined to genotype all their selection candidates so that they can perform GPS as the only type of preselection.

2.4.1 Bias

We observed negligible bias in our subsequent evaluations with ssGBLUP. Patry and Ducrocq [4] have shown that PBLUP following GPS underestimates the genetic trend and decreases the accuracy of EBV of young bulls and of their daughters. Therefore, we hypothesized that our observed lack of bias was due to using ssGBLUP in the subsequent evaluations. To show this, we repeated the subsequent evaluations for our preselection scenarios with a heritability of 0.1, this time using PBLUP, with all the other parameters left unchanged. The results of the PBLUP evaluations are in Table 2.5. Subsequent evaluations with ssGBLUP (Table 2.4) resulted in higher accuracies, lower or at least similar biases, and higher realized genetic gains than the corresponding PBLUP evaluations (Table 2.5). Without preselection or with RPS, bias (in both absolute and dispersion forms) was absent with PBLUP, just as with ssGBLUP. Without preselection, or with an ineffective preselection such as RPS (as shown from preselection accuracies in Tables 2.3, 2.4 and 2.5), no preselection bias is expected. However, with GPS and PAPS, where preselection was effective (as shown from preselection accuracies in Tables 2.3, 2.4 and 2.5), bias was always statistically significant with PBLUP (absolute bias ranging from 0.20 to 0.50 additive genetic SD, and $b_{TBV,EBV}$ ranging from 0.71 to 0.46), as opposed to being insignificant with ssGBLUP (absolute bias ranging from 0.03 to 0.04 additive genetic SD, and $b_{TBV,EBV}$ always not statistically different from 1). This comparison indeed confirms that with preselection, the observed bias in subsequent genetic evaluations based on PBLUP, is removed by using ssGBLUP.

Table 2. 5 PBLUP performance [a], with different preselection types[b] and intensities[c], all with heritability of 0.1

Measure	Type[b] and intensity[c] of preselection						
	No preselection	High GPS	Very High GPS	High PAPS	Very high PAPS	High RPS	Very high RPS
Preselection accuracy[d]	Not applicable	0.71 (0.69-0.73)	0.71 (0.69-0.73)	0.44 (0.42-0.46)	0.44 (0.42-0.46)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Subsequent selection accuracy[e]	0.46 (0.44-0.48)	0.25 (0.21-0.29)	0.24 (0.20-0.28)	0.32 (0.28-0.36)	0.30 (0.26-0.34)	0.33 (0.31-0.35)	0.31 (0.29-0.33)
Absolute bias[f]	0.00 (-0.02-0.02)	0.42 (0.38-0.46)	0.50 (0.46-0.54)	0.20 (0.16-0.24)	0.25 (0.23-0.27)	0.00 (-0.04-0.04)	0.00 (-0.04-0.04)
Dispersion bias[g]	0.98 (0.94-1.02)	0.47 (0.41-0.53)	0.46 (0.38-0.54)	0.71 (0.65-0.77)	0.67 (0.61-0.73)	0.99 (0.91-1.07)	0.98 (0.90-1.06)
True realized genetic gain[h]	0.79 (0.73-0.85)	1.12 (1.06-1.18)	1.16 (1.10-1.22)	0.84 (0.76-0.92)	0.82 (0.74-0.90)	0.26 (0.24-0.28)	0.15 (0.13-0.17)
Estimated realized genetic gain[i]	0.81 (0.75-0.87)	0.47 (0.43-0.51)	0.39 (0.35-0.43)	0.54 (0.50-0.58)	0.46 (0.42-0.50)	0.27 (0.25-0.29)	0.17 (0.15-0.19)

[a] Results shown only for the selection candidates in the most recent generation (i.e. generation 15), and the results are means over 10 replicates (and 95% confidence intervals). [b] Types of preselection: GPS - genomic preselection; PAPS - parent average preselection; RPS – random preselection. [c] Intensities of preselection: no preselection; high preselection - 10% of the male and 15% of the female selection candidates preselected; very high preselection - 5% of the male and 12.5% of the female selection candidates preselected. [d] Correlation between true and genomic estimated breeding values of all the selection candidates. [e] Correlation between true and estimated breeding values of the preselected animals. [f] Difference between average true breeding value and average estimated breeding value of preselected animals, expressed in additive genetic standard deviation. [g] Coefficient of regression of true on estimated breeding values of the preselected animals. [h] Difference between average true breeding value of subsequently selected animals and average true breeding value of their parents, expressed in additive genetic standard deviation. [i] Difference between average estimated breeding value of the subsequently selected animals and average estimated breeding value of their parents, expressed in additive genetic standard deviation.

2.4.2 Subsequent selection accuracy

Subsequent selection accuracy decreased with preselection, and this is in line with the findings reported by Patry and Ducrocq [4] in PBLUP evaluations following GPS in dairy cattle breeding schemes. It is important to note that without preselection, the subsequent selection accuracy was calculated across many more animals (16,000) compared to the 2000 and 1400 animals, respectively, used to for high and very high preselection scenarios. Even when the subsequent selection accuracy in the scenario without preselection was calculated using only these 2000 or 1400 preselected animals, it was still higher than in the scenarios with preselection (see Additional file 2.3). The explanation for this result is that each selection candidate had, on average, more full and half sibs at the subsequent genetic evaluation without preselection than in the high and with very high preselection scenarios, and the phenotypes of these additional full and half sibs added to the accuracy of the scenario without preselection. With different types of preselection, contrary to the trend that we observed with preselection accuracy, the subsequent selection accuracy increased from GPS to PAPS, and from PAPS to RPS, because the more accurate the preselection was, the lower the additive genetic variance left in the preselected animals [34, 35], which in turn reduced selection accuracy [35].

2.4.3 Realised genetic gain (RGG)

We observed a decrease in RGG (both TRGG and ERGG) as intensity of preselection increased. The reason for this is that as intensity of preselection increased, more of the best animals (in terms of TBV) were lost during preselection, since preselection was never 100% accurate. Other studies have reported a similar trend, i.e. a reduction in genetic gain with an increasing intensity of preselection, and offered similar explanations (e.g. [2,36–38]). With different types of preselection, we observed that RGG depended more on preselection accuracy than on subsequent selection accuracy, and therefore RGG had a trend that was more similar to the trend of preselection accuracy than to that of subsequent selection accuracy (Table 2.4). The reason is that among preselection types, variation in preselection accuracy was larger than that in subsequent selection accuracy (Table 2.4), due to different sources of information used in each preselection type (Table 2.2). In the subsequent genetic evaluations, irrespective of the type of preselection, the model used all three sources of information, i.e. pedigree, genotypes and phenotypes of the preselected candidates. This explains why RGG was always higher with GPS, than with PAPS, and why the lowest genetic gain was recorded with RPS. Schrooten et al. [2] also reported a larger impact of preselection accuracy than of subsequent selection accuracy on genetic gain in dairy cattle breeding schemes.

2.4.4 GPS and RGG

The decrease in RGG from no preselection to high and very high GPS scenarios was always small. Specifically, TRGG and ERGG decreased by 3.3 to 8.7% and 2.8 to 5.9%, respectively, from no preselection to high GPS, depending on heritability (Table 2.3). With the very high intensity of preselection, the number of females required to produce the next generation in this study (1000 females) was already reached at the preselection stage, thus there was no selection in females at the subsequent selection stage. TRGG and ERGG decreased, by 5.2 to 10.1% and 4.1 to 7.4%, respectively, from no preselection to very high GPS, depending on heritability (Table 2.3). These results show that, with ssGBLUP evaluations following GPS, it is possible to achieve a level of genetic gain that is similar to that achieved without preselection. This is especially important for traits that are expensive to measure (e.g. feed intake of individual broiler chickens), and traits for which phenotypes can only be measured at advanced stages of life (e.g. litter size in pigs). For such traits, GPS enables saving on the cost of phenotyping the preculled animals, and on the cost of raising the preculled animals in the expensive nucleus environments of breeding programs.

2.5 Conclusions

Using ssGBLUP in subsequent genetic evaluations prevents preselection bias, irrespective of intensity and type of preselection, and heritability. With GPS, in addition to reducing the phenotyping effort considerably, the use of ssGBLUP in subsequent genetic evaluations realizes only a slightly lower genetic gain than that realized without preselection. This is especially the case for traits that are expensive to measure (e.g. feed intake of individual broiler chickens), and traits for which phenotypes can only be measured at advanced stages of life (e.g. litter size in pigs).

2.6 Acknowledgements

The authors thank Marco Bink, Katrijn Peeters, Jeroen Visscher and Abe Huisman from Hendrix Genetics; Egbert Knol, Rob Bergsma and Egiel Hanenberg from Topigs Norsvin; John Henshall and Randy Borg from Cobb; and André M. Hidalgo, Chris Schrooten, and Gerben de Jong from CRV for their inputs towards the data simulation and implementation of preselection.

2.7 References

1. Appel LJ, Strandberg E, Danell B, Lundeheim N. Adjusting for missing data due to culling before testing in genetic evaluations of swine. J Anim Sci. 1998;76:1794–802.
2. Schrooten C, Bovenhuis H, van Arendonk JAM, Bijma P. Genetic progress in

multistage dairy cattle breeding schemes using genetic markers. J Dairy Sci. 2005;88:1569–81.
3. Janhunen M, Kause A, Vehviläinen H, Nousiainen A, Koskinen H. Correcting within-family pre-selection in genetic evaluation of growth-a simulation study on rainbow trout. Aquaculture. 2014;434:220–6.
4. Patry C, Ducrocq V. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci. 2011;94:1011–20.
5. Masuda Y, VanRaden PM, Misztal I, Lawlor TJ. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. J Dairy Sci. 2018;101:5194–206.
6. Meyer K, Thompson R. Bias in variance and covariance component estimators due to selection on a correlated trait. Zeitschrift für Tierzüchtung und Züchtungsbiologie. 1984;101:33–50.
7. Jensen J, Mao IL. Estimation of genetic parameters using sampled data from populations undergoing selection. J Dairy Sci. 1991;74:3544–51.
8. Árnason T, Albertsdóttir E, Fikse WF, Eriksson S, Sigurdsson Á. Estimation of genetic parameters and response to selection for a continuous trait subject to culling before testing. J Anim Breed Genet. 2012;129:50–9.
9. Mäntysaari EA, Liu Z, VanRaden P. Interbull validation test for genomic evaluations. Interbull Bull. 2010;41:17-22.
10. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157:1819–29.
11. Wolc A, Kranis A, Arango J, Settar P, Fulton JE, O'Sullivan NP, et al. Implementation of genomic selection in the poultry industry. Anim Front. 2016;6:23–31.
12. Vitezica ZG, Aguilar I, Misztal I, Legarra A. Bias in genomic predictions for populations under selection. Genet Res (Camb). 2011;93:357–66.
13. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, et al. Invited review: Reliability of genomic predictions for North American Holstein bulls. J Dairy Sci. 2009;92:16–24.
14. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci. 2010;93:743–52.
15. Christensen OF, Lund MS. Genomic prediction when some animals are not genotyped. Genet Sel Evol. 2010;42:2.
16. Misztal I, Aggrey SE, Muir WM. Experiences with a single-step genome evaluation. Poult Sci. 2013;92:2530–4.
17. Legarra A, Christensen OF, Aguilar I, Misztal I. Single step, a general approach for genomic selection. Livest Sci. 2014;166:54–65.
18. Sullivan PG. Mendelian sampling variance tests with genomic preselection. Interbull Bull. 2018;54:1-4.
19. Tyrisevä A-M, Mäntysaari EA, Jakobsen J, Aamand GP, Dürr J, Fikse WF, et al. Detection of evaluation bias caused by genomic preselection. J Dairy Sci.

2018;101:3155–63.
20. Henderson CR. Best linear unbiased estimation and prediction under a selection model. Biometris. 1975;31:423–47.
21. Pollak EJ, van der Werf J, Quaas RL. Selection bias and multiple trait evaluation. J Dairy Sci. 1984;67:1590–5.
22. Patry C, Ducrocq V. Accounting for genomic pre-selection in national blup evaluations in dairy cattle. Genet Sel Evol. 2011;43:30.
23. Shabalina T, Pimentel ECG, Edel C, Plieschke L, Emmerling R, Götz K-U. Short communication: The role of genotypes from animals without phenotypes in single-step genomic evaluations. J Dairy Sci. 2017;100:8277–81.
24. Koivula M, Strandén I, Aamand GP, Mäntysaari EA. Reducing bias in the dairy cattle single-step genomic evaluation by ignoring bulls without progeny. J Anim Breed Genet. 2018;00:1-9.
25. Sargolzaei M, Schenkel FS. QMSim: a large-scale genome simulator for livestock. Bioinformatics. 2009;25:680–1.
26. Sonesson AK, Meuwissen THE. Mating schemes for optimum contribution selection with constrained rates of inbreeding. Genet. Sel. Evol. 2000; 32:231–48.
27. ten Napel J, Vandenplas J, Lidauer M, Stranden I, Taskinen M, Mäntysaari E, et al. MiXBLUP: a user-friendly software for large genetic evaluation systems. 2017. https://mixblup.eu/download.html.
28. Tsuruta S, Lourenco DAL, Misztal I, Lawlor TJ. Possible causes of inflation in genomic evaluations for dairy cattle. In Proceedings of the 11th World Congress on Genetics Applied to Livestock Production: 11-16 February 2018; Auckland. 2018. p. 1–6.
29. Meuwissen THE, Luo Z. Computing inbreeding coefficients in large populations. Genet Sel Evol. 1992;24:305–13.
30. Powell JE, Visscher PM, Goddard ME. Reconciling the analysis of ibd and ibs in complex trait studies. Nat Rev Genet. 2010;11:800–5.
31. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.
32. Gilmour AR, Gogel BJ, Cullis BR, Thompson R. ASReml user guide release 3.0. Hemel Hempstead: VSN Int. Ltd. 2009. p. 275.
33. Knol EF, Nielsen B, Knap PW. Genomic selection in commercial pig breeding. Anim Front. 2016;6:15–22.
34. Bulmer MG. The effect of selection on genetic variability. Am Nat. 1971;105:201–11.
35. Gomez-Raya L, Burnside EB. The effect of repeated cycles of selection on genetic variance, heritability, and response. Theor Appl Genet. 1990;79:568–74.
36. Martinez V, Kause A, Mäntysaari E, Mäki-Tanila A. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: ii. two-stage selection. Aquaculture. 2006;254:195–202.
37. Campo JL, Fuente MB de la. Efficiency of two-stage selection indices in tribolium. J Hered. 1991;82:228–32.

38. Xu S, Martin TG, Muid WM. Multistage selection for maximum economic return with an application to beef cattle breeding. J Anim Sci. 1995;73:669–710.

Supplementary information

Additional file 2.1 The QMSim parameter file used to simulate the data used in this study

```
/*** Global parameters ***/
title = "title";
seed = "seed_main.prv"; // use the file "seed_main.prv" as the seed file
nrep = 10; // replicate 10 times
h2 = 0.3; // total h2, also 0.5 and 0.1
qtlh2 = 0.27; // proportion of h2 explained by QTL
phvar = 100; // total phenotypic variance

/*** Historical population ***/
begin_hp;
hg_size = 5000 [0] 50 [2997] 5000 [3000]; // start with 5000 animals at generation 0, keep reducing the population size until it reaches 50 animals at generation 2997, and then start raising it until it reaches 5000 animals again at generation 3000. Maintain equal sex ratio across all the generations.

/*** Recent population ***/
begin_pop = "rp";
    begin_founder;
        male [n = 100, pop = "hp"];
    female [n = 1000, pop = "hp"]; // select the founder population from the last generation of historical population. Select 100 males and 1000 females randomly to form the founder population. Also, select 100 males and 1000 females per generation to produce the next generation, based on the selection and mating criteria below:
    end_founder;
    ng = 15; // simulate 15 generations of recent population
    ls= 16; // litter size is 16 offspring per dam
    pmp = 0.5 /fix_litter; // maintain equal sex ratio per litter (so each litter has 8 male and 8 female offspring)
    md = minf; // mating design is to minimize inbreeding
    sd = ebv /h; // selection is based on highest EBV
    ebv_est = blup; // estimate EBV using pedigree BLUP
    begin_popoutput;
        data; // save all individual's data except their genotypes
        stat; // save brief statistics on simulated data
        allele_freq /mafbin 50; // save allele frequencies, with 50 bins for minor allele frequency distribution
        genotype /gen 13 14 15; // save genotype data for generations 13, 14, and 15 only
    end_popoutput;
end_pop;

/*** Genome section ***/
begin_genome;
    begin_chr = 30; // simulate 30 chromosomes
```

```
        chrlen = 100; // each should be 100cM in length
        nmloci = 5000; // simulate 5000 markers per chromosome. This will ensure that at the
        end of the historical population, we have about 2000 markers per chromosome with
        MAF ≥ 0.005, as about 60% of the markers end up with MAF < 0.005.
        mpos = even /start 0.5 /end 99.5; // place the markers evenly between 0.5 cM and
        99.5 cM
        nma = all 2; // all markers should have two alleles each at the beginning of the
        historical population
        maf = eql; // at the 1st generation of historical population, all marker alleles should
        have equal frequency (0.5)
        nqloci = 250; // simulate 250 QTL per chromosome. This will ensure that at the end of
        the historical population, we have about 100 QTL per chromosome with MAF >= 0.005,
        as about 60% of the QTL end up with a MAF < 0.005.
        qpos = even /start 1 /end 99; // place the QTL between 1st and the 99th cM
        nqa = all 2; // all QTL should have two alleles each at the beginning of the historical
        population
        qaf = eql; // at the 1st generation of historical population, all QTL alleles should have
        equal frequency (0.5)
        qae = rndg 0.4; // QTL allele effects should be randomly drawn from a gamma
        distribution with shape parameter of 0.4
    end_chr;
    select_seg_loci /maft 0.005; // at the end of the historical population, consider only
    markers and QTL with MAF ≥ 0.005 in producing the recent population
    mmutr = 2.5e-5 /recurrent; // marker mutation rate, and the mutation should be
    recurrent. This means that mutation is only possible among existing alleles and no new
    alleles are formed.
    qmutr = 2.5e-5 /recurrent; // QTL mutation rate, and the mutation should be recurrent
end_genome;

/*** General output ***/
begin_output;
    linkage_map; // save linkage map
    allele_effect; // save marker and QTL allele substitution effects
    hp_stat; // save statistics of the historical population
end_output;
```

Additional file 2.2 The MiXBLUP instruction file used for data analysis in this study

mixblup_ssGBLUP # title of the analysis

 # Observations & systematic effects
DATAFILE pheno # name of the file with phenotypic data
animal I # the ID of the animals is in the first column, and it is in numeric form
sire I # the sires of the animals are in the second column, and they are in numeric form
dam I # the sires of the animals are in the second column, and they are in numeric form
sex A # sex of the animal, and it is in alphanumeric form
phen T # the trait of interest

 # Genetic similarity among individuals
ERMFILE geno !CONSTRUCT SSmat # the genotype file; make a new weighted $\mathbf{G}^{-1}$ in the MiXBLUP parser from the genotype file
animal I # the ID of the animals is in the first column, and it is in numeric form
!METHOD VanRaden # use the first method of VanRaden in making **G**
!DENSE #markers provided arevin dense format
!MAF 0.005 # Minor allele frequency threshhold
!STORE_GINV # store $\mathbf{G}^{-1}$ in the right format to be re-used by calc_grm.
!NUMPROC 5 # the number of threads to be used by calc_grm.
!LAMBDA 1 # weighing factor for $\mathbf{G}^{-1}$ in making $\mathbf{H}^{-1}$
!ALPHA 0.9 # weighing factor for $\mathbf{G}$ in making $\mathbf{G}^{-1}$
!BETA 0.1 # weighing factor for $\mathbf{A}_{22}^{-1}$ in making $\mathbf{G}^{-1}$
!OMEGA 1 # weighing factor for $\mathbf{A}_{22}^{-1}$in making $\mathbf{H}^{-1}$
!SINGLESTEP # MiXBLUP kernel should calculate the $\mathbf{H}^{-1}$ from a $\mathbf{G}^{-1}$, the pedigree file and a file with inbreeding coefficients
PEDFILE pedi !CALCINBR # pedigree file, calculate inbreeding coefficients from it
animal I # the ID of the animals is in the first column, and it is in numeric form
sire I # the sires of the animals are in the second column, and they are in numeric form
dam I # the sires of the animals are in the second column, and they are in numeric form

 # Components of variance and covariance among traits
PARFILE VCfile # the file containing additive genetic and residual variances

 # Statistical models
MODEL
phen ~ mu !RANDOM G(animal) # the phenotype of interest is a function of overall mean and genomic differences among the animals

 # Control of analysis and output
SOLVING # use the below information to control the process and the output of the analysis.
!STOPCRIT 1.0E-06 # convergence criterion
!NOPEEK # do not store preliminary results
TMPDIR /destination # working directory

Additional file 2. 3 Accuracy[a] of the no preselection (control) scenario in subsequent ssGBLUP evaluations, calculated across different animals[b]

Calculated across	Accuracy[a]
all the 16,000 selection candidates	0.80 (0.78-0.82)
the 2000 selection candidates preselected under high GPS scenario	0.62 (0.60-0.64)
the 1400 selection candidates preselected under very high GPS scenario	0.62 (0.60-0.64)
the 2000 selection candidates preselected under high PAPS scenario	0.75 (0.73-0.77)
the 1400 selection candidates preselected under very high PAPS scenario	0.75 (0.75-0.75)
the 2000 selection candidates preselected under high RPS scenario	0.80 (0.78-0.82)
the 1400 selection candidates preselected under very high RPS scenario	0.80 (0.80-0.80)

[a] Correlation between true and genomic estimated breeding values of all the preselected candidates, and results are means of 10 replicates (and 95% confidence intervals). [b] The different animals are the animals preselected under the different types and intensities of preselection. The types of preselection are GPS - genomic preselection, PAPS - parent average preselection and RPS – random preselection. The different intensities of preselection are no preselection, high preselection - 10% of the male and 15% of the female selection candidates preselected, and very high preselection - 5% of the male and 12.5% of the female selection candidates preselected.

This page intentionally left blank

Chapter 3

Avoiding preselection bias in subsequent ssGBLUP evaluations of genomically preselected animals

Ibrahim Jibrila, Jeremie Vandenplas, Jan ten Napel, Roel F Veerkamp and Mario P L Calus

Wageningen University and Research Animal Breeding and Genomics, Droevendaalsesteeg 1, 6708PB Wageningen, the Netherlands

J Anim Breed Genet. 2021;138:432–441

Abstract

In animal breeding, parents of the next generation are usually selected in multiple stages, and the initial stages of this selection are called preselection. Preselection reduces the information available for subsequent evaluation of preselected animals and this sometimes leads to bias. The objective of this study was to establish the minimum information required to subsequently evaluate genomically preselected animals without bias arising from preselection, with single-step genomic best linear unbiased prediction (ssGBLUP). We simulated a nucleus of a breeding program in which a recent population of 15 generations was produced. In each generation, parents of the next generation were selected in a single-stage selection based on pedigree BLUP. However, in generation 15, 10% of male and 15% of female offspring were preselected on their genomic estimated breeding values (GEBV). These GEBV were estimated using ssGBLUP, including the pedigree of all animals in generations 0 to 15, genotypes of all animals in generations 13 to 15 and phenotypes of all animals in generations 11 to 14. In subsequent ssGBLUP evaluation of these preselected animals, genotypes and phenotypes from various groups of animals were excluded one after another. We found that GEBV of the preselected animals were only estimated without preselection bias when genotypes and phenotypes of all animals in generations 13 and 14 and of the preselected animals were included in the subsequent evaluation. We also found that genotypes of the animals discarded at preselection only helped in reducing preselection bias in GEBV of their preselected sibs when genotypes of their parents were absent or excluded from the subsequent evaluation. We concluded that to prevent preselection bias in subsequent ssGBLUP evaluation of genomically preselected animals, information representative of the reference data used in the evaluation at preselection and genotypes and phenotypes of the preselected animals are needed in the subsequent evaluation.

Keywords: Bias, genomic preselection, multi-stage selection, single-step genomic BLUP

3.1 Background

In animal breeding programs, parents of a next generation are usually selected in multiple stages (e.g. Meyer & Thompson, 1984; Xu et al., 1995; Árnason et al., 2012), and initial stages of selection are referred to as preselection (e.g. Patry & Ducrocq, 2011; Janhunen et al., 2014; Masuda et al. 2018). Impact of preselection on subsequent genetic evaluations has been a subject of research for a long time in the field of animal breeding (e.g. Henderson, 1975; Appel et al., 1998; Patry & Ducrocq, 2011a; Masuda et al., 2018). Traditionally, preselection for target traits has mostly been based on correlated indicator traits that are easily and cheaply measurable early in lives of selection candidates. For example, piglets could be preselected based on weaning weight as an indicator trait for average daily gain during performance testing. In such situations, multi-trait evaluations are performed including both the target traits and the indicator traits based on which animals are preselected (e.g. Henderson, 1975; Pollak et al., 1984; Janhunen et al., 2014), to prevent preselection bias in the subsequent evaluations. In this case animals retained at preselection (preselected animals) will have better phenotypes for the indicator traits compared to their discarded (preculled) siblings, and this informs the subsequent evaluations using pedigree-based best linear unbiased prediction (PBLUP) model that for the target traits, preselected animals are better-than-average sets of offspring of their parents. In other words, including the indicator traits in subsequent evaluations provides the PBLUP model in the subsequent evaluations with data to better estimate the (on-average-positive) Mendelian sampling terms of preselected animals. In the genomic era, preselection is mostly based on genomic estimated breeding values (GEBV) of young selection candidates, and this form of preselection is called genomic preselection (GPS, e.g. Patry & Ducrocq, 2011a; Masuda et al., 2018; Sullivan et al., 2019). Although GPS is practiced in several livestock species, including pigs and poultry, reports on GPS in the literature so far are all focussing on dairy cattle. Subsequent PBLUP evaluations after GPS, such as the Interbull and national dairy cattle evaluations, have been reported to be biased (e.g. Patry et al., 2013; Masuda et al., 2018; Sullivan, 2018). It has been shown that this preselection bias can be prevented by including genomic information in form of genomic pseudoperformances (e.g. deregressed proofs) of both preselected and preculled animals in the subsequent PBLUP evaluations (Patry & Ducrocq, 2011b). The genomic pseudoperformances of preculled animals help to inform the PBLUP model in subsequent evaluations that preselected animals are better-than-average subsets of offspring of their parents (Patry & Ducrocq, 2011b). Jibrila et al. (2020) showed that using ssGBLUP in subsequent evaluations prevents GEBV of preselected animals

from becoming biased due to preselection, even if genotypes of preculled animals are excluded. This suggests that, in contrast with subsequent PBLUP evaluations, information from preculled animals is not strictly needed to prevent preselection bias in subsequent ssGBLUP evaluation of their preselected sibs. Based on the literature and our previous work (Shabalina et al., 2017; Koivula et al., 2018; Jibrila et al., 2020), we hypothesize that the impact of genotypes of preculled animals in subsequent ssGBLUP evaluations depend on whether genotypes of their parents are included in the subsequent evaluation. The objective of this study was to establish, through simulation, the minimum information required in subsequent ssGBLUP evaluations to estimate GEBV of genomically preselected animals without bias associated with preselection. We also investigated under which circumstances the use of genotypes of preculled animals is beneficial in subsequent evaluations of their preselected sibs. And finally, we evaluated the accuracy realized with each of the implemented scenarios of subsequent evaluation.

3.2 Materials and methods

3.2.1 Data simulation

Before designing this study, we had discussions with the industrial partners of the Breed4Food consortium (https://breed4food.com/), which are breeding companies in dairy cattle (CRV), pigs (Topigs Norsvin and Hendrix Genetics), poultry (Hendrix Genetics and Cobb Europe) and Aquaculture (Hendrix Genetics). During these discussions it became clear that breeding practices for the different species are relatively similar and can be represented by a general breeding program as simulated in our study. Therefore, we used inputs from these breeding companies and simulated a nucleus of a general breeding program. We used QMSim (Sargolzaei & Schenkel, 2009) to simulate the datasets, and the details of the simulation can be found in Jibrila et al. (2020).

Briefly, at the end of a historical population of 3000 generations of random mating, we randomly selected 100 males and 1000 females and used them as founders. From these founders, we produced a recent population of 15 generations. In each of these recent generations, 100 males and 1000 females were selected in a single stage PBLUP-based selection to produce the next generation of 16,000 animals. Within sex, all selected parents had equal contribution to the next generation. Across a simulated genome consisting of 30 chromosomes, 60,000 single nucleotide polymorphisms (SNP) and 3000 quantitative trait loci (QTL) were evenly distributed. The breeding goal was made up of a single quantitative trait that was measured in

both sexes, with heritability of 0.1. For every individual, the phenotype of the trait was simulated as the summation of random additive genetic and residual effects (so no fixed effects). The additive genetic variance was made up of QTL variance (90%) and polygenic variance (10%). In this study, we used the entire pedigree (i.e. consisting of all animals in generations 0 to 15), genotypes of the three most recent generations (i.e. consisting of all animals in generations 13 to 15) and phenotypes of the five most recent generations (i.e. consisting of all animals in generations 11 to 15).

3.2.2 Preselection and subsequent genetic evaluations

We implemented preselection only in the most recent generation (i.e. generation 15). From the selection candidates (i.e. all animals in generation 15), 10% of males and 15% of females were preselected based on their individual GEBV. These GEBV were obtained using ssGBLUP, including the pedigree of animals in generations 0 to 15, genotypes of animals in generations 13 to 15, and phenotypes of all the animals in the generations 11 to 14.

In subsequent (second stage) evaluation we implemented 13 scenarios, with varying amounts and sources of genomic and phenotypic information. Scenarios 1 to 9 included either the entire genomic information available, or a subset of it, in addition to the phenotypic information available. Details of the information included in each of these scenarios are in Table 3.1. Similarly, each of the last four scenarios included either the entire phenotypic information available or a subset of it, in addition to all the genomic information available (except for the genotypes of preculled animals). Details of the information included in each of these scenarios are in Table 3.2. Available sources of genomic and phenotypic information for subsequent evaluation of the preselected animals, based on their closeness to the preselected animals, can be grouped as follows:

- *Sources of genomic information:* i) the preselected animals themselves, ii) the preculled animals, iii) parents of the selection candidates (i.e. selected animals in generation 14), and iv) other animals with genotypes, which were, respectively, the unselected sibs of parents of the selection candidates (i.e. the rest of generation 14 animals) and the selection candidates' grandparental generation (i.e. generation 13 animals).
- *Sources of phenotypic information***:** i) the preselected animals themselves, ii) animals with both genotypes and phenotypes at the time of preselection (i.e. selection candidates' parental and grand parental generations/animals in generations 13 and 14), and iii) animals with phenotypes but no

genotypes at the time of preselection (i.e. selection candidates' (great) great grandparental generations/animals in generations 11 and 12).

All the genetic evaluations (including at the preselection stage) were performed using the ssGBLUP procedure implemented in MiXBLUP (ten Napel et al., 2017). Every step of this study (data simulation, preselection and subsequent evaluations) was replicated 10 times.

3.2.3 Implementation of single-step GBLUP

For each replicate, we used a pedigree-based animal model in ASReml (Gilmour et al., 2009) to estimate the additive genetic and residual variances that we later supplied to MiXBLUP. Pedigree information from all animals in generations 0 to 14 and phenotypic information from all animals in generations 11 to 14 were used to estimate these variances. The model used in both ASReml (for estimation of variance components) and MiXBLUP (for breeding value estimation) was:

$$\mathbf{y} = \mathbf{xb} + \mathbf{Zu} + \mathbf{e},$$

where **y** was the vector of phenotypes; **x** and **Z** were incidence vector and matrix linking phenotypes to overall mean and random animal effects, respectively; **b** was the overall mean; **u** was the vector of breeding values; and **e** was the vector of random residuals.

Genetic relationships among animals were accounted for by the inverse of the combined pedigree-genomic relationship ($\mathbf{H^{-1}}$), obtained as follows (Aguilar et al., 2010; Christensen & Lund, 2010):

$$\mathbf{H^{-1}} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & (0.9\mathbf{G_t} + 0.1\mathbf{A}_{22})^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{A}^{-1}$ was the inverse of the pedigree relationship matrix, and $\mathbf{A}_{22}$ was the pedigree relationship matrix among genotyped animals. We considered inbreeding in setting up both $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$ to avoid the bias caused by ignoring inbreeding (Tsuruta et al., 2019). The matrix $\mathbf{G_t}$ was the genomic relationship matrix calculated as follows:

$$\mathbf{G_t} = \left(1 - \bar{f}_p\right)\mathbf{G_r} + 2\bar{f}_p\mathbf{11}',$$

where $\bar{f}_p$ was the average pedigree inbreeding coefficient across genotyped animals, $\mathbf{G_r}$ was the raw genomic relationship matrix computed following the first method of VanRaden (VanRaden, 2008), and $\mathbf{11}'$ is a matrix of 1s. The transformation of $\mathbf{G_r}$ to $\mathbf{G_t}$ was done to make the average genomic inbreeding equal to the average pedigree inbreeding, i.e. to have $\mathbf{G}$ and $\mathbf{A}_{22}$ on the same scale so that they are compatible. This formula is similar to the formula $\mathbf{G_t} = \mathbf{G_r} + \alpha\mathbf{11}'$ of Vitezica et al (2011), and it

is equivalent to the F_{st}-based formula proposed by Powell et al. (2010), which can be rewritten to $\mathbf{G_t} = \left(1 - \frac{1}{2}\alpha\right)\mathbf{G_r} + \alpha\mathbf{11}'$ (Vitezica et al., 2011). The connection between these three formulas can be seen as follows. In these notations, α is the difference in average pedigree and genomic relationships (Vitezica et al., 2011). Using current allele frequencies to compute $\mathbf{G_r}$, the mean genomic relationship is expected to be zero, such that α reduces to the mean pedigree relationship, and in this case, assuming random mating, $\mathrm{E}(\bar{f}_p) = \frac{1}{2}\alpha$ (Falconer & Mackay, 1996). As there were scenarios in this study in which selective genotyping was introduced (scenarios 6 and 7), this transformation made sure that its impact was taken care of (Vitezica et al., 2011; Hsu et al., 2017). In computing $\mathbf{G_r}$, we calculated (current) allele frequencies using all available genomic data (i.e. using all the available genomic data, per scenario), and only used SNP with minor allele frequency of at least 0.005. We gave the weights of 0.9 to $\mathbf{G_t}$ and 0.1 to $\mathbf{A}_{22}$ to account for polygenic variance (which was simulated to be 10% of the genetic variance) and to ensure that **G** was invertible (Aguilar et al., 2010; Christensen & Lund, 2010). The MiXBLUP instruction file for the ssGBLUP analysis can be found in Jibrila et al. (2020).

3.2.4 Measures of bias and accuracy in the subsequent evaluations

Bias was calculated in two ways. Firstly, absolute bias was calculated as the difference between mean True Breeding Value (TBV) and mean GEBV of all preselected animals, and expressed in genetic standard deviation (SD) units. Absolute bias is a measure of whether estimated genetic gain is equal to true genetic gain. Therefore, if there is no absolute bias (i.e. when mean difference is 0), estimated genetic gain is equal to true genetic gain. A negative difference means that on average GEBV overestimate TBV, and therefore genetic gain is overestimated, and vice versa. To have TBV on the same scale as GEBV, we subtracted mean TBV and mean GEBV of the animals in generations 11 to 14 from TBV and GEBV of each preselected animal, respectively. Secondly, dispersion bias was calculated as the regression coefficient of TBV on GEBV ($b_{TBV,GEBV}$) of all preselected animals. Dispersion bias is a measure of how well differences in (G)EBV of animals represent the differences in their TBV. If $b_{TBV,(G)EBV}$ is 1, then there is no dispersion bias. A value of $b_{TBV,(G)EBV}$ less than 1 means that variance of (G)EBV of animals is inflated compared to variance of their TBV, and so differences in (G)EBV of the animals overestimate differences in their TBV, and vice versa. Accuracy was calculated as the Pearson's correlation coefficient between TBV and GEBV of all preselected animals.

3.3 Results

Results of the subsequent evaluations conducted in this study are presented in Tables 3.1 and 3.2. Results in Table 3.1 are for the scenarios with varying amounts and sources of genomic information, and those in Table 3.2 are for the scenarios with varying amounts and sources of phenotypic information.

3.3.1 Impact of genomic information from various groups of animals on bias and accuracy

With all available phenotypes included in the subsequent evaluation, negligible absolute and dispersion biases were observed when all available genotypes were included (scenario 1) and even when genotypes of preculled animals were excluded (scenario 2). For these two scenarios, absolute bias was only 0.03 genetic SDs and $b_{TBV,GEBV}$ was 1.01. The highest accuracy of GEBV of the preselected animals was achieved when all the genotypes and phenotypes available after preselection were included (0.48, scenario 1), and when the genotypes of the preculled animals were excluded (scenario 2). This means that just like with bias, accuracy too was not affected moving from scenario 1 to 2.

When genotypes of the preselected animals, of the selection candidates' parents, of the selection candidates' parents' culled sibs, or of the selection candidates' grandparental generation were excluded from the subsequent evaluation, both absolute and dispersion biases and accuracy loss were observed (scenarios 3 to 9). In scenarios 1 to 9, both absolute and dispersion biases increased and accuracy decreased with decreasing number of animals with both genotypes and phenotypes. The only exception is scenario 3, because this was the only scenario where genotypes of the preselected animals were excluded from the subsequent evaluation. (see Tables 3.1 and 3.2). We also observed that across all scenarios where preculled animals were included, their genotypes only helped in minimizing bias and accuracy loss when genotypes of the selection candidates' parents were excluded. This can be seen by comparing scenario 1 against 2, 4 against 5, and 6 against 7 on the one hand, and scenario 8 against 9 on the other hand, as described next.

Table 3.1 Implementation and results of subsequent genetic evaluations with varying sources and amounts of genomic information[a]

Scenario	Genotypes included? Yes (✓) or No (×)					Number of animals with both genotypes and phenotypes (as a proxy for reference population)	Measures of bias and accuracy		
	g13	Selected g14	Culled g14	Preselected g15	Preculled g15		Absolute bias (in genetic SD units)	Dispersion bias ($b_{TBV,GEBV}$)	Accuracy
1 (control)	✓	✓	✓	✓	✓	34000	0.03 (0.01)	1.01 (0.03)	0.48 (0.01)
2	✓	✓	✓	✓	×	34000	0.03 (0.01)	1.01 (0.03)	0.48 (0.02)
3	✓	✓	✓	×	×	32000	0.36 (0.02)	0.39 (0.03)	0.23 (0.02)
4	×	✓	✓	✓	✓	18000	0.12 (0.01)	0.67 (0.03)	0.38 (0.02)
5	×	✓	✓	✓	×	18000	0.12 (0.01)	0.67 (0.03)	0.38 (0.02)
6	×	✓	×	✓	✓	3100	0.32 (0.01)	0.51 (0.03)	0.32 (0.02)
7	×	✓	×	✓	×	3100	0.32 (0.01)	0.51 (0.03)	0.32 (0.02)
8	×	×	×	✓	✓	2000	0.32 (0.02)	0.53 (0.03)	0.32 (0.02)
9	×	×	×	✓	×	2000	0.45 (0.02)	0.45 (0.03)	0.28 (0.02)

[a] Pedigree of all animals from generation 0 up to preselected generation 15 and phenotypes of all animals from generation 11 up to preselected generation 15 included in every scenario. Pedigree of preculled animals included whenever their genotypes were included and excluded whenever their genotypes were excluded.

Table 3.2 Implementation and results of subsequent genetic evaluations with varying sources and amounts of phenotypic information[a]

Scenario	Phenotypes included? Yes (✓) or No (×)				Number of animals with both genotypes and phenotypes (as a proxy for reference population)	Measures of bias and accuracy		
	g11 & g12	g13	g14	Preselected g15		Absolute bias (in genetic SD units)	Dispersion bias ($b_{TBV,GEBV}$)	Accuracy
2 (control)[b]	✓	✓	✓	✓	34000	0.03 (0.01)	1.01 (0.03)	0.48 (0.02)
10	✓	✓	✓	×	32000	0.10 (0.03)	1.04 (0.04)	0.41 (0.01)
11	✓	✓	×	✓	18000	0.20 (0.01)	0.62 (0.03)	0.39 (0.02)
12	✓	×	×	✓	2000	0.28 (0.02)	0.41 (0.02)	0.30 (0.02)
13	×	✓	✓	✓	34000	0.01 (0.01)	0.99 (0.03)	0.48 (0.02)

[a] Pedigree of all animals from generation 0 up to preselected generation 15 and genotypes of all animals from generation 13 up to preselected generation 15 included in every scenario. [b] Scenario 2 was used as control here, as opposed to scenario 1 in Table 3.1. This means that preculled animals were completely excluded from all the scenarios in this Table.

In scenario 1, there was no bias, and the highest accuracy was achieved. Excluding genotypes of preculled animals, i.e. moving from scenario 1 to 2, did not cause any change. There was bias and accuracy loss in scenario 4, in which genotypes of the selection candidates' grandparental generation were excluded. Further exclusion of genotypes of the preculled animals, i.e. moving from scenario 4 to 5, again did not make any difference. The bias increased and accuracy dropped further, in scenario 6, as a result of exclusion of genotypes of culled sibs of the selection candidates' parents, in addition to excluding genotypes of the selection candidates' grandparental generation. Here, further exclusion of genotypes of the preculled animals, i.e. moving from scenario 6 to 7, did not make any difference either, because genotypes of the preselected animals and of all the selection candidates' parents were still in the model.

In scenario 8, in which only genotypes of all selection candidates were included in the subsequent evaluation, absolute bias was 0.32, $b_{TBV,GEBV}$ was 0.53 and accuracy was 0.32. These values are similar to those observed in scenarios 6 and 7, where the only genotypes included were those of the selection candidates and their parents. However, when the only genotypes included were those of the preselected animals (scenario 9), absolute bias increased to 0.45, $b_{TBV,GEBV}$ decreased to 0.45 and accuracy decreased to 0.28. In summary, with all available phenotypes included, including genotypes of the preselected animals and of the preselected animals' parental and grandparental generations in the subsequent evaluation appeared to be sufficient to prevent preselection bias and minimize accuracy loss due to preselection.

3.3.2 Impact of phenotypic information from various groups of animals on bias and accuracy

Because in the previous section we observed that genotypes of the preculled animals were not needed in our subsequent evaluation, we ignored them in this section. Compared to scenario 2, which was the control scenario in Table 3.2, both absolute and dispersion biases increased and accuracy decreased according to the number of animals with both genotypes and phenotypes. With genotypes of the preselected animals and of the preselected animals' parental and grandparental generations included in the subsequent evaluation, excluding phenotypes of the preselected animals (moving from scenario 2 to 10) resulted in some absolute bias (0.10 genetic SD) and some accuracy loss (from 0.48 to 0.41). A tendency towards deflation was also observed in scenario 10 ($b_{TBV,GEBV}$ of 1.04). Both absolute and dispersion biases and accuracy loss were also observed when phenotypes of the selection candidates'

parental generation were excluded (scenario 11) and even more when additionally phenotypes of the selection candidates' grandparental generation were excluded (scenario 12). Although phenotypes of animals three or four generations before the generation of the selection candidates (i.e. generations 11 and 12) were included in the evaluation at preselection stage, excluding these phenotypes from the subsequent evaluation (scenario 13) did not cause any bias or accuracy loss in GEBV of the preselected animals.

3.4 Discussion

In this study, our objective was to investigate the roles of genotypes and phenotypes from various groups of animals in preventing bias due to preselection, when estimating GEBV of genomically preselected animals in subsequent ssGBLUP evaluation. To achieve this objective, we performed simulations involving several simplifying assumptions, as discussed hereafter, that helped to assess the impact of these different sources of information, which may not be possible in data resembling the full complexity of breeding programs in practice. One of the assumptions was to have discrete generations, to enable assessing the impact of using data of different groups of ancestors of the preselected animals. We also modelled only one step of (genomic) preselection, although in reality preselection usually takes place in more than one step. For example, it is common to genotype only members of families preselected based on parent average, and then genomic preselection takes place within these families. Nevertheless, the impact of multiple steps of (different types of) low-intensity preselection is expected to be similar to that of one step of high-intensity preselection. In both cases for the subsequent ssGBLUP evaluation, phenotypes are only available for the animals that survived the last step of preselection. Although we simulated a trait whose phenotypes can be measured on both sexes, the results of our study are (in most instances) applicable to sex-limited traits as well. The main difference between the trait we simulated and sex-limited traits is availability of records on males. In practice, for sex-limited traits, progeny information serve as phenotype for males. So in our study, the fact that we performed the subsequent evaluation after preselected animals had their own records is comparable to the subsequent evaluation that takes place, in dairy cattle for example, when preselected young bulls have daughter information (though performance of many daughters is more valuable, at least for accuracy of breeding values, than a single own performance; e.g. Mrode, 2014). Overall, although the characteristics of the simulated trait more closely resemble some traits in pigs and

poultry, where phenotyping is across both sexes, our results are (in most instances) applicable to pig, poultry and dairy cattle breeding schemes.

3.4.1 Implementation of single-step GBLUP

The variance components used in ssGBLUP were estimated from our current data, rather than using the simulated values. We did this to reflect what happens in practice, where base generation variance components are not known, but estimated from current data. The trait studied in this study was simulated with heritability of 0.1 and phenotypic variance of 100. Therefore, at the base generation (generation 0) additive genetic variance was 10 and residual variance was 90. When we estimated the variance components as described in the methodology section, additive genetic and residual variances across the 10 replicates were on average 7.41 and 90.24, respectively. Note that information from the selection candidates (generation 15 animals) was not used in estimating these variance components, so all the subsequent evaluation scenarios used the same values per replicate. Using the same data as used in this study, we studied the impact of decreasing or increasing the base generation additive genetic variance by 25% while keeping the residual variance the same. We found that that does not have any statistically significant impact on accuracy and bias of ssGBLUP evaluations (results not shown).

Our scenarios 6 and 7 introduced the problem of selective genotyping, which has been reported to cause bias and reduce accuracy of ssGBLUP evaluations (e.g. Vitezica et al., 2011; Christensen, 2012; Hsu et al., 2017). The implementation of our ssGBLUP model by default takes care of this problem by making the average genomic inbreeding equal to the average pedigree inbreeding, as indicated in the Methodology section. To verify whether this correction worked, we repeated scenario 6, this time without preselection (so with phenotypes of the preculled animals included). The results we found were statistically similar as the results obtained with all available information included (i.e. pedigree of all animals in generations 0 to 14, genotypes of all animals in generations 13 to 15 and phenotypes of all animals in generations 11 to 15). This confirms that the biases and accuracy loss we observed in this study were the result of excluding, from the subsequent ssGBLUP evaluation, either some of the information used as preselection reference data (scenarios 4 to 9, 11 and 12) or information from preselected candidates themselves (scenarios 3 and 10).

3.4.2 The minimum information required in subsequent ssGBLUP evaluation to prevent preselection bias

Although phenotypes of animals three or four generations before the generation of the selection candidates (i.e. generations 11 and 12) were included in the evaluation at preselection stage, excluding these phenotypes from the subsequent evaluation did not cause any bias or accuracy loss in estimating GEBV of the preselected animals. The facts that these animals (the (great) great grandparents) were far from the preselected animals, and that the selection candidates' parental and grandparental generations had both genotypes and phenotypes may explain this. Lourenco et al. (2014) found that truncating phenotypic information to only two to three ancestral generations does not affect accuracy of predicting (G)EBV of young animals in dairy cattle and pig breeding programs. The findings of Lourenco et al. (2014) also mean that in our study, phenotypes of the selection candidates' (great) great grandparental generations did not contribute much in estimating GEBV of the selection candidates during the evaluation at preselection stage. Therefore, including genotypes and phenotypes of the selection candidates' parental and grandparental generations in our subsequent evaluation implies that the most relevant ancestral information used in our evaluation at preselection stage was considered.

Similarly, excluding genotypes of the preculled animals from the subsequent evaluation of their preselected sibs did not cause bias or accuracy loss in GEBV of the preselected animals. Because genotypes and phenotypes of the selection candidates' parents were already included in the subsequent evaluation, including genotypes and phenotypes of the preselected animals alone (without necessarily including genotypes of their preculled sibs) provided the ssGBLUP model in the subsequent evaluation with the remaining data it needed to estimate the positive average Mendelian sampling term of the preselected animals. In preventing bias and accuracy loss in subsequent ssGBLUP evaluation of the preselected animals, genotypes of preselected animals appear to be more important than phenotypes of the preselected animals. This can be seen by comparing scenarios 3 and 10. Although scenarios 3 and 10 have the same number of animals with both genotypes and phenotypes, results of scenario 3 (in which genotypes of preselected animals were excluded from the subsequent evaluation) were worse than those of scenario 10 (in which phenotypes of the preselected animals were excluded from the subsequent evaluation). The fact that genotypes of the preselected animals were included in the evaluation at preselection stage (and phenotypes of the preselected animals were not) may explain this. For preselected dairy sires, however, which usually have

performance of many daughters instead of a single own performance, the relative importance of own genomic and phenotypic information in preventing bias and accuracy loss due to preselection may be different from what we saw in this study. This is because performance of many daughters is more valuable, at least for accuracy of breeding values, than own performance (e.g. Mrode, 2014). In summary, to prevent preselection bias in subsequent ssGBLUP evaluation of genomically preselected animals, it is sufficient to supply the model with i) information representative of the reference data used in the evaluation at preselection stage and ii) genotypes and phenotypes of the preselected animals, which are the main source of information that informs ssGBLUP that the preselected animals are a better-than-average subset of offspring of their parents.

3.4.3 Comparison to observations in dairy cattle

In scenario 10, GEBV of the preselected animals are effectively the same as their GEBV at preselection stage, and the measures of bias and accuracy in these two evaluations were statistically similar (results not shown for the evaluation at preselection stage). In dairy cattle breeding programs, it is nowadays common to select all young sires in one stage as soon as they are genotyped (e.g. Mäntysaari et al., 2020), though some form of preselection based on parent average is often applied. If such GEBV are later compared to deregressed proofs or daughter yield deviations of such sires, we expect that some positive absolute bias, some accuracy loss, and a tendency towards deflation would be observed just as in our scenario 10. However, in practice, dairy cattle breeding companies observe negative absolute bias (overestimated genetic trend) and inflation when they make such comparisons (e.g. Mäntysaari et al., 2020). The reason for this is unclear, and should be investigated in future studies.

3.4.4 Role of genotypes of preculled animals in subsequent ssGBLUP evaluations

In ssGBLUP evaluations, pedigree relationships among genotyped and non-genotyped animals guide the implicit imputation of genotypes of non-genotyped animals (Christensen & Lund, 2010; Legarra et al., 2009; Misztal et al., 2009). Similarly, in subsequent ssGBLUP evaluations, when some or all parents of selection candidates are not genotyped, more accurate imputation of genotypes of the non-genotyped parents and other non-genotyped animals in the pedigree is achieved by including genotypes of all (both preselected and preculled) offspring of the non-genotyped parents than by including genotypes of their preselected offspring alone (Shabalina et al., 2017). Because in our study all parents of the selection candidates had genotypes and these genotypes were included in the subsequent evaluation,

genotypes of the preculled animals were no longer needed in the subsequent evaluation. However, just like the findings of Shabalina et al. (2017), our scenarios 8 and 9 show that including genotypes of preculled animals in subsequent ssGBLUP evaluations reduces bias and increases accuracy in estimating GEBV of their preselected sibs when their parents are not genotyped. In another ssGBLUP evaluation, Koivula et al. (2018) observed unexpected tendencies towards more dispersion bias and less accuracy in the GEBV of young selected bulls when genotypes of culled bulls were included compared to when they were excluded. Their results, however, were not statistically significant, and thus inconclusive. Our results show that when parents of selection candidates are genotyped, including genotypes of their preculled offspring in subsequent ssGBLUP evaluations neither improves nor deteriorates the quality of the evaluations. In current breeding programs for all livestock species, often not all dams are genotyped. If evaluations at preselection stage are done with ssGBLUP, dams that are not genotyped benefit from genotypes of all their offspring. Including genotypes of their preculled offspring in subsequent ssGBLUP evaluations ensures that the same levels of accuracy of imputing genotypes of such dams are achieved as in the evaluations at preselection stage. Therefore, in such situations, genotypes of preculled animals would be needed in subsequent ssGBLUP evaluations to estimate GEBV of preselected animals without preselection bias and accuracy loss.

3.5 Conclusion

To prevent preselection bias in subsequent ssGBLUP evaluation of genomically preselected animals, it is sufficient to supply the model with i) information representative of the reference data used in the evaluation at preselection stage and ii) genotypes and phenotypes of the preselected animals, which are the main source of information that informs ssGBLUP that the preselected animals are a better-than-average subset of offspring of their parents. When (some) parents of selection candidates are not genotyped, genotypes of preculled animals, together with genotypes of preselected animals, help in more accurately imputing genotypes of their ungenotyped parents in ssGBLUP evaluations at both preselection and subsequent evaluation stages. In such situations, genotypes of preculled animals are needed in subsequent ssGBLUP evaluations to estimate GEBV of their preselected sibs without preselection bias.

3.6 Acknowledgements

The authors thank Marco Bink, Katrijn Peeters, Jeroen Visscher and Abe Huisman

from Hendrix Genetics; Egbert Knol, Rob Bergsma and Egiel Hanenberg from Topigs Norsvin; John Henshall and Randy Borg from Cobb; and André Hidalgo, Chris Schrooten, and Gerben de Jong from CRV for their inputs towards the data simulation and implementation of preselection. We also thank the Dutch Ministry of Economic Affairs (TKI Agri & Food project 16022) and the Breed4Food consortium (with Cobb Europe, CRV, Hendrix Genetics and Topigs Norsvin in addition to Wageningen University and Research as partners) for their financial support to this project.

3.7 References

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93, 743–752.doi:10.3168/jds.2009-2730

Appel, L. J., Strandberg, E., Danell, B., & Lundeheim, N. (1998). Adjusting for Missing Data Due to Culling before Testing in Genetic Evaluations of Swine. *Journal of Animal Science*, 76, 1794–1802.doi:10.2527/1998.7671794x

Árnason, T., Albertsdóttir, E., Fikse, W. F., Eriksson, S., & Sigurdsson, Á. (2012). Estimation of genetic parameters and response to selection for a continuous trait subject to culling before testing. *Journal of Animal Breeding and Genetics*, 129, 50–59.doi:10.1111/j.1439-0388.2011.00941.x

Christensen, O. F. (2012). Compatibility of pedigree-based and marker-based relationship matrices for single-step genetic evaluation. *Genetics Selection Evolution*, 44:1. doi:10.1186/1297-9686-44-37

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, 42:2 doi:10.1186/1297-9686-42-2

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics* (4th ed.). London: Longman Group Ltd.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., & Thompson, R. (2009). ASReml user guide release 3.0. https://asreml.kb.vsni.co.uk/wp-content/uploads/sites/3/2018/02/ASReml-3-User-Guide.pdf

Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometris*, 31, 423–447.doi:10.2307/2529430

Hsu, W.-L., Garrick, D. J., & Fernando, R. L. (2017). The Accuracy and Bias of Single-Step Genomic Prediction for Populations Under Selection. *Genes|Genomes|Genetics*, 7, 2685–2694.doi:10.1534/g3.117.043596

Janhunen, M., Kause, A., Vehviläinen, H., Nousiainen, A., & Koskinen, H. (2014). Correcting within-family pre-selection in genetic evaluation of growth-A simulation study on rainbow trout. *Aquaculture*, 434, 220–226.doi:10.1016/j.aquaculture.2014.08.020

Jibrila, I., ten Napel, J., Vandenplas, J., Veerkamp, R. F., & Calus, M. P. L. (2020).

Investigating the impact of preselection on subsequent single-step genomic BLUP evaluation of preselected animals. *Genetics Selection Evolution*, 52:42.doi:10.1186/s12711-020-00562-6

Koivula, M., Strandén, I., Aamand, G. P., & Mäntysaari, E. A. (2018). Reducing bias in the dairy cattle single-step genomic evaluation by ignoring bulls without progeny. *Journal of Animal Breeding and Genetics*, 135, 107–115.doi:10.1111/jbg.12318

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, 92, 4656–4663.doi:10.3168/jds.2009-2061

Lourenco, D. A. L., Misztal, I., Tsuruta, S., Aguilar, I., Lawlor, T. J., Forni, S., & Weller, J. I. (2014). Are evaluations on young genotyped animals benefiting from the past generations? *Journal of Dairy Science*, 97, 3930–3942.doi:10.3168/jds.2013-7769

Mäntysaari, E. A., Koivula, M., & Strandén, I. (2020). Symposium review: Single-step genomic evaluations in dairy cattle. *Journal of Dairy Science*, 103, 5314–5326.doi:10.3168/jds.2019-17754

Masuda, Y., VanRaden, P. M., Misztal, I., & Lawlor, T. J. (2018). Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. *Journal of Dairy Science*, 101, 5194–5206.doi:10.3168/jds.2017-13310

Meyer, K., & Thompson, R. (1984). Bias in variance and covariance component estimators due to selection on a correlated trait. *Zeitschrift Für Tierzüchtung Und Züchtungsbiologie*, 101, 33–50.doi:10.1111/j.1439-0388.1984.tb00020.x

Misztal, I., Legarra, A., & Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*, 92, 4648–4655.doi:10.3168/jds.2009-2064

Mrode, R. (2014). *Linear Models for the Prediction of Animal Breeding Values* (3rd ed.). Boston: CAB International.

Patry, C., & Ducrocq, V. (2011). Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. *Journal of Dairy Science*, 94, 1011–1020.doi:10.3168/jds.2010-3804

Patry, Clotilde, & Ducrocq, V. (2011). Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. *Genetics Selection Evolution*, 43:30.doi:10.1186/1297-9686-43-30

Patry, Clotilde, Jorjani, H., & Ducrocq, V. (2013). Effects of a national genomic preselection on the international genetic evaluations. *Journal of Dairy Science*, 96, 3272–84.doi:10.3168/jds.2011-4987

Pollak, E. J., van der Werf, J., & Quaas, R. L. (1984). Selection Bias and Multiple Trait Evaluation. *Journal of Dairy Science,* 67, 1590-1595.doi:10.3168/jds.S0022-0302(84)81481-2

Powell, J. E., Visscher, P. M., & Goddard, M. E. (2010). Reconciling the analysis of IBD and IBS in complex trait studies. *Nature Reviews Genetics*, 11, 800–

805.doi:10.1038/nrg2865

Sargolzaei, M., & Schenkel, F. S. (2009). QMSim: A large-scale genome simulator for livestock. *Bioinformatics*, 25(5, 680–681.doi:10.1093/bioinformatics/btp045

Shabalina, T., Pimentel, E. C. G., Edel, C., Plieschke, L., Emmerling, R., & Götz, K.-U. (2017). Short communication: The role of genotypes from animals without phenotypes in single-step genomic evaluations. *Journal of Dairy Science*, 100, 8277–8281.doi:10.3168/jds.2017-12734

Sullivan, P. G. (2018). Mendelian Sampling variance tests with genomic preselection. *INTERBULL BULLETIN*, 54, 1-4. https://journal.interbull.org/index.php/ib/article/view/1463

Sullivan, P. G., Mäntysaari, E. A., Dejong, G., & Benhajali, H. (2019). Modifying MACE to accommodate genomic preselection effects. *INTERBULL BULLETIN*, 55, 77–80. https://journal.interbull.org/index.php/ib/article/view/1484

ten Napel, J., Vandenplas, J., Lidauer, M., Stranden, I., Taskinen, M., Mäntysaari, E., … Veerkamp, R. F. (2017). MiXBLUP: A user-friendly softwarevfor large genetic evaluation systems. http://www.mixblup.eu/documents/Manual MiXBLUP 2.1_June 2017_V2.pdf

Tsuruta, S., Lourenco, D. A. L., Masuda, Y., Misztal, I., & Lawlor, T. J. (2019). Controlling bias in genomic breeding values for young genotyped bulls. *Journal of Dairy Science*, 102, 9956–9970.doi:10.3168/jds.2019-16789

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91, 4414–4423.doi:10.3168/jds.2007-0980

Vitezica, Z. G., Aguilar, I., Misztal, I., & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, 93, 357–366.doi:10.1017/S001667231100022X

Xu, S., Martin, T. G., & Muid, W. M. (1995). Multistage Selection for Maximum Economic Return with an Application to Beef Cattle Breeding. *J. Anim. Sci.*, 73, 669–710.doi:10.2527/1995.733699x

Chapter 4

Impact of genomic preselection on subsequent genetic evaluations with ssGBLUP - using real data from pigs

Ibrahim Jibrila[1], Jeremie Vandenplas[1], Jan ten Napel[1], Rob Bergsma[2], Roel F Veerkamp[1] and Mario P L Calus[1]
[1] Wageningen University and Research Animal Breeding and Genomics, PO Box 338
6700 AH Wageningen, the Netherlands

[2] Topigs Norsvin Research Center B.V., Schoenaker 6, 6641 SZ Beuningen, the Netherlands

Accepted for publication in Genetics Selection Evolution

Abstract

Background

Empirically assessing the impact of preselection on subsequent genetic evaluations of preselected animals requires comparison of scenarios taking into account different approaches, including scenarios without preselection. However, preselection almost always takes place in animal breeding programs, so it is difficult to have a dataset without preselection. Hence most studies on preselection used simulated datasets, concluding that genomic estimated breeding values (GEBV) from subsequent single-step genomic best linear unbiased prediction (ssGBLUP) evaluations are unbiased. The aim of this study was to investigate the impact of genomic preselection (GPS) on accuracy and bias in subsequent ssGBLUP evaluations, using data from a commercial pig breeding program.

Methods

We used data on four pig production traits from one sire line and one dam line. The traits are average daily gain during performance testing, average daily gain throughout life, backfat thickness, and loin depth. As these traits had different weights in the breeding goals of the two lines, we analyzed the two lines separately. Per line, we had a reference GPS scenario which kept all available data, against which the next two scenarios were compared. We then implemented two other scenarios with additional layers of GPS by removing all animals without progeny either i) only in the validation generation, or ii) in all generations. We conducted subsequent ssGBLUP evaluations per GPS scenario, utilizing all the data remaining after implementing the GPS scenario. In computing accuracy and bias, we compared GEBV against progeny yield deviations of validation animals.

Results

Results for all traits in both lines showed marginal loss in accuracy due to the additional layers of GPS. Average accuracy across all GPS scenarios in both lines was 0.39, 0.47, 0.56, and 0.60 respectively for the four traits considered in this study. Bias was largely absent, and when present did not differ greatly among corresponding GPS scenarios.

Conclusion

As preselection generally has the same effect in animal breeding programs, we concluded that impact of preselection is generally minimal on accuracy and bias in subsequent ssGBLUP evaluations of selection candidates in pigs and in other animal breeding programs.

4.1 Background

In animal breeding, parents of the next generation are often selected in multiple stages, and the initial stages of this selection are called preselection [1–3]. Selection candidates that survive preselection are called preselected animals [1–3], and those that do not are called preculled animals [3,4]. Preselection aims to reduce costs and efforts spent on animals that are not interesting for the breeding program, and achieves this by avoiding phenotyping or further testing of preculled animals. Due to introduction of genomic prediction [5], preselection is now mostly based on genomic estimated breeding values (GEBV) of young animals even before they have records of any traits. This type of preselection is called genomic preselection (GPS; e.g. [1,2]). The popularity of GPS is because genotyping is becoming cheaper by the day, and the reasonable reliabilities of GEBV (e.g. [6–8]). As genomically preculled animals have neither progeny nor records for some or all breeding goal traits, they are generally not included in subsequent genetic evaluations (i.e. genetic evaluations that come after preselection). GPS therefore decreases the amount of information available for subsequent genetic evaluations of preselected animals. Properly assessing the impact of preselection on subsequent genetic evaluation of preselected animals requires comparison of scenarios taking into account different approaches, including a scenario without preselection. Because in animal breeding programs preselection almost always takes place, it is difficult, if not impossible, to have a scenario without preselection. This is why most studies available on preselection used simulated datasets (e.g. [1,3,9,10]) . Those studies have shown that when a subsequent genetic evaluation of preselected animals is done using pedigree-based best linear unbiased prediction (PBLUP), preselection results in accuracy loss and bias in the estimated breeding values (EBV) of preselected animals [1,3,9–12]. Some of these studies [9–12] further showed that the accuracy loss and bias caused by GPS can be avoided if the information on preculled animals that was utilized at preselection is included in subsequent PBLUP evaluations. On the other hand, our previous works [3,4] have shown that when the subsequent genetic evaluation is done with single-step genomic BLUP (ssGBLUP), genomic EBV (GEBV) of preselected animals are estimated without bias. We [4] further showed that to avoid GPS bias in subsequent ssGBLUP evaluation of preselected animals, genotypes of their preculled sibs are only needed if not all of their parents are genotyped.

In our previous works [3,4], being based on simulated datasets, preselection was the only possible source of bias in ssGBLUP evaluations. However, in real breeding programs, other sources of bias in ssGBLUP evaluations may exist and are potentially

difficult to control. Therefore, impact of preselection might be confounded by the impact of these other factors. These other possible sources of bias include, amongst others, inaccurate or incomplete pedigree [13], inaccurately estimated additive genetic (co)variances [13], and a reference population of selectively genotyped animals [14,15]. Although some ways to reduce the bias caused by these factors have been developed, the bias is usually not completely eliminated in evaluations using real data (e.g. [13–15]). This may explain the observation that in practice GEBV obtained from ssGBLUP evaluations are sometimes biased. The aim of this study was to investigate the impact of GPS on accuracy and bias in subsequent ssGBLUP evaluations, using data from a commercial pig breeding program in which preselection has taken place. To achieve this aim, we used the full dataset as control and retrospectively implemented additional layers of GPS. The additional layers of GPS were implemented by discarding animals that did not have progeny in the data. Since in the breeding program GEBV were used to select parents of next generations, discarding animals without progeny in the data can be considered as additional GPS. Then we compared results from subsequent ssGBLUP evaluations after these additional layers of GPS against results from ssGBLUP evaluation using the full available data. Our subsequent genetic evaluations only involved reevaluating preselected animals, either with or without preculled animals in the subsequent evaluations.

4.2 Methods

4.2.1 Data

We obtained pig production traits data on one sire line and one dam line from Topigs Norsvin. These data were collected between 1970 and 2020, and the traits were average daily gain during performance testing (ADGT), average daily gain throughout the lifetime (ADGL), backfat thickness, and loin depth. These traits are part of the breeding goals of each line. However, there was more emphasis on reproduction traits than on production traits in the dam line. Details on the amount of data utilized in this study are in Table 4.1. The data were recorded on originally preselected animals (i.e. the animals preselected by Topigs Norsvin), with the sire line being much more balanced than the dam line, in terms of proportions males and females with records per generation (ratio of males with records to females with records is about 50:50 in the sire line and about 20:80 in the dam line). We studied impact of genomic preselection (GPS) in the two lines separately, because the traits we studied had different weights in breeding goals of the two lines. Ancestors from the same line and year of birth with unknown parents were considered to be a separate base

population in the pedigree. Each base population was fitted as a genetic group to account for genetic trend and differences in origin and selection history [16].

4.2.2 Training and validation generations

Per line, we split all animals into two groups, according to a cut-off birth date. Animals born before or on the cut-off birth date were used as training population, and animals born after the cut-off birth date were used as validation population. The cut-off date (to split the data into training and validation populations) was 31st January, 2017 for the sire line, and 31st December, 2015 for the dam line. Then from the validation population, animals that met the following requirements were selected as validation animals: 1) none of their parents were in the validation population, and 2) the animals had phenotyped progeny. The first requirement ensured that our validation animals were from only one generation, and the second requirement enabled comparing GEBV of the validation animals against their progeny yield deviations (PYD) [17]. Since records on validation animals were included in some of our subsequent evaluation scenarios (as will be seen later), we chose to use PYD as proxy for true breeding values (TBV) because PYD are estimated from phenotypes that were not included in the subsequent genetic evaluations.

4.2.3 Genomic data and quality control

Our genomic data included genotypes of animals for about 21,000 SNP segregating in both lines, and distributed across the 18 autosomes in the pig genome. The SNP were genotyped using a custom SNP chip. We used Plink [18] for all quality control operations on our genomic data. Per GPS scenario (as described later) and per line, animals and SNPs with call rates less than 90% were removed, as well as SNPs that deviated from Hardy-Weinberg equilibrium (Hardy-Weinberg equilibrium exact test p value = 10^{-15}), or had a minor allele frequency below 0.005. Table 4.1 contains the summary of the pedigree, genomic and phenotypic information utilized in the subsequent genetic evaluations following each GPS scenario.

4.2.4 Computation of precorrected phenotypes

In our genetic evaluations, we used precorrected phenotypes (rather than raw phenotypes) as records. Animals of different lines were sometimes raised together, so they shared some fixed and non-genetic random effects. Because we studied impact of GPS within lines, it was necessary to correct phenotypes for all non-genetic effects before the data was divided into lines. Another motivation for using precorrected phenotypes is that after implementing our additional GPS scenarios (as described in detail in the next section), some classes of these non-genetic effects

could be left with only one or a few animals. Then correcting for these effects would be less accurate compared to correcting for them before implementing our additional GPS scenarios. To compute precorrected phenotypes ($\mathrm{y_c}$), we first ran the following multi-trait pedigree-based animal model as follows:

$$\mathbf{y_j} = \mathbf{X_j b_j} + \mathbf{W_j p_j} + \mathbf{Z_j u_j} + \mathbf{e_j}, \qquad (\mathbf{eq.\ 1}).$$

Where for every trait (j) $\mathbf{y_j}$ was the vector of phenotypes; $\mathbf{b_j}$ was the vector of fixed effects, with incidence matrix $\mathbf{X_j}$; $\mathbf{p_j}$ was the vector of non-genetic random effects, with incidence matrix $\mathbf{W_j}$; $\mathbf{u_j}$ was the vector of breeding values, with incidence matrix $\mathbf{Z_j}$; and $\mathbf{e_j}$ was the vector of residuals. The model assumed $\mathbf{u_j}$ and $\mathbf{e_j}$ to be normally distributed, each with mean of zero. For all traits (and across all animals), **u** and **e** had variance-covariance matrices **A** $\otimes$ **G** and **I** $\otimes$ **R**, respectively. Where **A** was the pedigree relationship matrix among animals, **I** was an identity matrix with dimensions equal to the number of animals with records, and **G** and **R** were respectively the trait by trait additive genetic and residual variance-covariance matrices. Then for every animal (i) with phenotype for trait j, we computed its precorrected phenotype (y_{cij}) as:

$$\mathrm{y_{cij}} = \hat{\mathrm{u}}_{\mathrm{ij}} + \hat{\mathrm{e}}_{\mathrm{ij}} \qquad \textbf{(eq. 2).}$$

The (co)variance components used for this analysis were estimated, before separating the data into lines, from a four-trait pedigree-based animal model in ASReml [19] using **eq. 1**. All computations of (G)EBV were performed using MiXBLUP [20]. We decided to use a pedigree-based model (instead of a single-step model) to estimate the variance components because previous studies [21,22] showed that in populations undergoing genomic selection (as in our data), pedigree-based models estimate variance components in the pedigree founders at least as good as single-step models.

4.2.5 Preselection

Per line, we implemented a reference scenario and two scenarios that added layers of GPS. The reference scenario - against which other scenarios could be compared - only included the original GPS implemented by Topigs Norsvin. Thus, the subsequent ssGBLUP evaluations following the reference scenario utilized the entire available data until the validation generation. The second scenario is called validation generation preselection (the VGP scenario). In this scenario, we only implemented

additional GPS in the validation generation, by discarding all animals in the validation generation that had no progeny in the data, but had genotypes and/or phenotypes. The third scenario is called multi-generation preselection (the MGP scenario), in which we discarded any animal in the validation and training generations without progeny in the data. Animals kept after each of the GPS scenarios are shown in Figure 4.1.

Table 4.1 Data utilized in subsequent ssGBLUP[a] evaluations following each preselection scenario, after quality control

Data in the subsequent ssGBLUP evaluation / Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference[b]	VGP[c]	MGP[d]
The sire line (number of validation animals per trait is ± 1383)						
Number of animals in the pedigree	81,875	60,950	12,777	81,875	60,950	12,777
Number of animals with record for at least one trait[e]	75,129	54,217	6,065	52,846	52,846	4,694
Number of animals with genotypes	33,506	23,315	5,131	33,506	23,315	5,131
Number of SNP genotyped	20,550	20,963	20,926	20,550	20,963	20,926
The dam line (number of validation animals per trait is ± 2051)						
Number of animals in the pedigree	160,426	124,031	33,485	160,426	124,031	33,485
Number of animals with record for at least one trait	139,403	103,018	12,514	100.710	100,710	10,206
Number of animals with genotypes	50,895	36,369	9,072	50,895	36,369	9,072
Number of SNP genotyped	19,199	19,256	20,647	19,199	19,256	20,647

[a] single-step genomic best linear unbiased prediction. [b] In the reference scenario, the subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection (VGP) scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection (MGP) scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] About 87% and 70% of the animals in sire and dam lines respectively have records for all the four traits utilized in this study, and even bigger numbers have records for any two and three traits. We decided to keep any animal with record for at least one of the traits (92% and 87% of the animals in the sire and dam lines respectively) because every animal in the analyses would benefit from records on relatives and records of correlated traits (see Table S4.1), in addition to its own record on the primary trait.

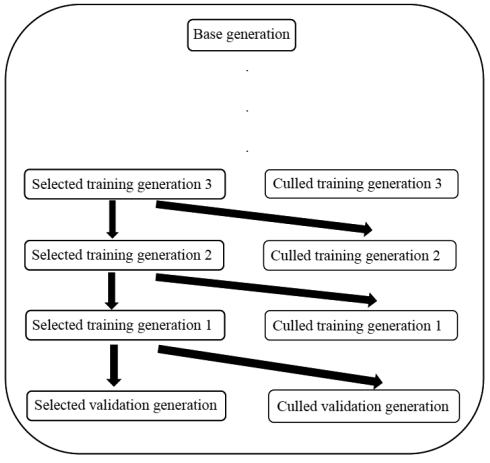


Figure 1a: Reference scenario

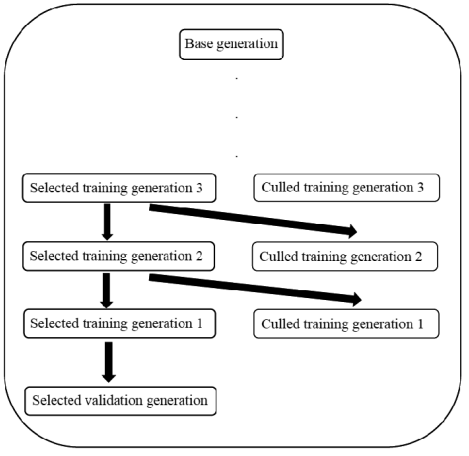


Figure 1b: Validation generation preselection (VGP) scenario

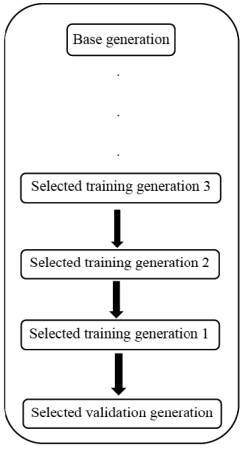


Figure 1c: Multi-generation preselection (MGP) scenario

Figure 4.1 Overview of groups of animals used in subsequent ssGBLUP evaluation of each of the considered GPS scenario

4.2.6 Subsequent genetic evaluations

Following every scenario of GPS, we implemented a subsequent ssGBLUP evaluation with all animals that survived the GPS. We call this evaluation subsequent because it came after the initial evaluation that provided the GEBV used in preselection. The ssGBLUP evaluations were conducted with and without records (i.e. own precorrected phenotypes) on the animals in the validation generation (see Table 4.1), to represent traits with records (e,g, production traits) and those without records (e,g. reproduction traits) available during subsequent evaluations. Progeny of validation animals were not included in the subsequent genetic evaluations. We estimated variance components for every preselection scenario, per line, using a pedigree-based multi-trait animal model in ASReml. We used these scenario-specific variance components in the subsequent genetic evaluations to ensure that the variance components used were appropriate for the precorrected phenotypes. At the subsequent genetic evaluations, the (multi-trait) model used for the estimations of both variance components and breeding values for every trait (j) was:

$$\mathbf{y_j} = \mathbf{1_j}\mu_j + \mathbf{Z_j u_j} + \mathbf{e_j}, \qquad \textbf{(eq. 3)},$$

where for every trait (j) $\mathbf{y_j}$ was the vector of precorrected phenotypes; $\mathbf{1_j}$ was an incidence vector of 1's, and $\mathbf{Z_j}$ was incidence matrix, linking precorrected phenotypes to overall mean and random animal effects, respectively; μ_j was the overall mean; $\mathbf{u_j}$ was the vector of breeding values; and $\mathbf{e_j}$ was the vector of residuals. The model assumed $\mathbf{u_j}$ and $\mathbf{e_j}$ to be normally distributed, each with mean of zero. For all traits (and across all animals), **u** and **e** had variance-covariance matrices **H** $\otimes$ **G** and **I** $\otimes$ **R**, respectively. Where **H** was the combined genomic and pedigree relationship matrix among animals as explained hereafter, **I** was an identity matrix with dimensions equal to the number of animals with records, and **G** and **R** were respectively the trait by trait additive genetic and residual variance-covariance matrices. We also repeated all subsequent genetic evaluations using PBLUP, to verify the impact of using genotypes on the observed results.

4.2.7 Implementation of single-step GBLUP

The inverse of the combined pedigree-genomic relationship ($\mathbf{H^{-1}}$) was obtained as follows [23,24]:

$$\mathbf{H^{-1}} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & (0.95\mathbf{G_t} + 0.05\mathbf{A}_{22})^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \qquad \textbf{(eq. 4)},$$

where $\mathbf{A}^{-1}$ was the inverse of the pedigree relationship matrix, and $\mathbf{A}_{22}$ was part of

the pedigree relationship matrix referring to genotyped animals. We considered inbreeding in setting up both $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$, as ignoring inbreeding in setting up $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$ has been reported to cause bias in GEBV [13]. The adjusted genomic relationship matrix $\mathbf{G_t}$ was computed as follows [14,25]:

$$\mathbf{G_t} = \left(1 - \bar{f}_p\right)\mathbf{G_r} + 2\bar{f}_p\mathbf{11}' \qquad \textbf{(eq. 5)},$$

where $\bar{f}_p$ was the average pedigree inbreeding coefficient across genotyped animals, $\mathbf{G_r}$ was the raw genomic relationship matrix computed following the first method of VanRaden [26], and $\mathbf{11}'$ was a matrix of 1s. As the animals with genotypes in this study were selectively genotyped, this transformation made sure that the impact of selective genotyping was taken care of and that **G** and $\mathbf{A_{22}}$ were on the same scale and therefore compatible [14,15]. To compute $\mathbf{G_r}$, we computed (current) allele frequencies using all available genomic data after quality control. We gave the weights of 0.95 to $\mathbf{G_t}$ and 0.05 to $\mathbf{A}_{22}$ to ensure that **G** was invertible [23,24].

4.2.8 Measures of accuracy and bias in the subsequent genetic evaluations

We used progeny yield deviation (PYD) [17] as a proxy for true breeding value (TBV), against which GEBV were compared when computing accuracy and bias. To compute PYD, we ran a multi-trait pedigree-based animal model per line in MiXBLUP, with precorrected phenotypes as records and an overall mean as the only fixed effect **(eq. 3)**. The (co)variance components used in this model were also estimated per line in ASReml, from precorrected phenotypes, using a multi-trait pedigree-based animal model that only included a mean fixed effect **(eq. 3)**. From the output of this analysis, we computed PYD for all validation sires and dams (i) for each trait (j) as:

$$\mathrm{PYD}_{ij} = \frac{\sum_{p=1}^{n} y_{cpj} - \frac{1}{2}a_{mj}}{n} \qquad \textbf{(eq. 6)},$$

where PYD_{ij} was the progeny yield deviation of a sire or dam i for trait j, y_{cpj} was the precorrected phenotype of a progeny p of the sire or dam i for trait j, a_{mj} was the breeding value of the mate of sire or dam i (for trait j) in producing offspring p, and n was the number of phenotyped progeny of sire or dam i. Estimation of PYD was done before removing progeny of validation animals from the data. Since progeny of validation animals were not included in subsequent genetic evaluations, comparing (G)EBV to PYD can be considered as a forward-in-time validation. We computed approximate reliability of PYD for each validation animal for each trait, and used this approximate reliability as the weighting factor to compute accuracy and bias, to

account for differences in number of progeny used to estimate PYD for different validation animals. The reliability of PYD was approximated as:

$$\frac{{}^{1}/_{4}\mathrm{nh}^{2}}{1+{}^{1}/_{4}(\mathrm{n}-1)\mathrm{h}^{2}} \qquad \textbf{(eq. 7)},$$

where *n* was the validation animal's number of half-sib progeny with records, and h^2 was the heritability of the trait [27]. For convenience, we assumed all progeny of a validation animal were half-sibs, though some of them were full-sibs. We also computed unweighted accuracy an bias, and accuracy and bias weighted by approximate PYD reliabilities obtained considering all progeny as full sibs. We did not observe statistically significant differences among the results, so we decided to only report the accuracy and bias weighted by approximate reliability computed considering all progeny as half sibs (as in eq. 7).

Validation accuracy was computed as weighted Pearson's correlation coefficient between PYD and GEBV of all validation animals, using the 'cor.test' function of the 'stats' package in R [28]. We computed the standard errors (SE) of the estimates from the confidence intervals (CI) produced by the 'cor.test' function. Validation accuracy is not numerically the same as the accuracy of predicting TBV, since PYD has some non-genetic component, in addition to TBV [17]. However, validation accuracy and accuracy of predicting TBV increase and decrease together [29], and this property of validation accuracy enables us to use it to make comparison among subsequent genetic evaluation scenarios.

We computed two types of bias. The first type is level bias, which is a measure of whether estimated genetic gain is equal to true genetic gain. Level bias was computed as the weighted mean difference between PYD and half of the (G)EBV across all validation animals, expressed in additive genetic standard deviation (SD) units of the trait. We used the 'weighted.mean' function of the 'stats' package in R [28] to compute estimates of the weighted mean differences, and used the 'weighted_se' function of the 'diagis' package in R [30] to compute SE of the estimates. A negative difference means that GEBV are on average overestimated, and therefore genetic gain is overestimated, and vice versa. Since PYD were computed from a dataset that included information on progeny of validation animals and (G)EBV were computed without information on progeny of validation animals, PYD and (G)EBV were on different scales. Therefore before computing differences between PYD and half of the (G)EBV of validation animals, we scaled PYD and (G)EBV

to be expressed against the same genetic base, consisting of the first three training generations. We did this in the following steps: from the model used in computing PYD, we computed average EBV across all animals in the first three training generations. We then subtracted half of this average EBV from PYD of each validation animal. Then for each subsequent genetic evaluation, we computed the average (G)EBV of all animals in the first three training generations. We then subtracted this average (G)EBV from (G)EBV of each validation animal.

The other type of bias we computed is dispersion bias, which was measured by the weighted regression coefficient of PYD on (G)EBV of all validation animals. We used the 'lm' function of the 'stats' package in R [28] to compute both the estimates and SE of the regression coefficients. If the regression coefficient is equal to the expected value, then there is no dispersion bias. Note that the expected value is 0.5, because PYD only includes half of the breeding value of a parent. A regression coefficient less than the expected value means that variance of (G)EBV is inflated, and vice versa.

4.3 Results

Table 4.2 shows normalized means and SD of precorrected phenotypes of the traits analyzed, following the implememted genomic preselection (GPS) scenarios. The results show that our additionally implemented GPS was effective, as the lines were selected (and preselected) for increased feed efficiency (i.e. higher feed intake and higher average daily gain), slightly decreased backfat thickness, and slightly increased loin depth. As the validation generation for both validation generation preselection (VGP) and multi-generation preselection (MGP) scenarios only contained the preselected animals, means and SD of precorrected phenotypes of the traits are the same for these two GPS scenarios when only considering the animals in the validation generation. When only the validation generation was considered (i.e. the middle part of Table 4.2), means of precorrected phenotypes of average daily gain during performance testing (ADGT) and average daily gain throughout life (ADGL) increased from reference scenario to VGP and MGP scenarios. At the same time, SD of precorrected phenotypes of these traits decreased from reference scenario to VGP and MGP scenarios. In other words, means of precorrected phenotypes of these traits were higher, and SD of the precorrected phenotypes were lower, among preselected animals in the validation generation than among all animals in the validation generation. For backfat thickness, both mean and SD slightly decreased from reference to VGP and MGP scenarios. The change was also only slight for loin depth, with the mean slightly increasing, and the SD slightly decreasing,

from reference to VGP and MGP scenarios. Higher effectiveness of MGP over VGP can be seen from the right part of Table 4.2 (i.e. when means and SD were computed across the entire data). For the positively (pre)selected traits (i.e ADGT, ADGL, and loin depth) means were higher for MGP than for VGP. For backfact, which was negatively (pre)selected, mean was lower for MGP than for VGP. As would be expected, SD were in all cases lower for MGP than for VGP.

Table 4.**2** Normalized[a] Means and SD (in brackets) of precorrected phenotypes of the traits utilized in this study, following each GPS scenario

Trait /Preselection scenario	Only within the validation generation		Across the entire data	
	Reference[b]	VGP[c]/MGP[d]	VGP	MGP
Sire line				
ADGT[e] (g/day)	0.12 (1.00)	0.51 (0.80)	-0.08 (1.00)	0.38 (0.84)
ADGL[f] (g/day)	0.03 (1.00)	0.41 (0.85)	-0.11 (1.00)	0.31 (0.86)
Backfat thickness (mm)	-0.27 (1.00)	-0.29 (0.95)	-0.01 (1.00)	-0.10 (0.96)
Loin depth (mm)	0.26 (1.00)	0.26 (0.97)	0.17 (1.00)	0.20 (0.97)
Dam line				
ADGT (g/day)	0.22 (1.00)	0.62 (0.90)	-0.06 (1.00)	0.40 (0.91)
ADGL (g/day)	0.11 (1.00)	0.53 (0.87)	-0.11 (1.00)	0.32 (0.90)
Backfat thickness (mm)	-0.13 (1.00)	0.05 (0.97)	-0.02 (1.00)	-0.06 (0.99)
Loin depth (mm)	0.22 (1.00)	0.13 (0.98)	0.20 (1.00)	0.24 (0.99)

[a] The values were normalized by dividing them by the standard deviations of their corresponding reference scenarios. [b] In the reference scenario, the subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection (VGP) scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection (MGP) scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] ADGT: Average daily gain during performance testing. [f] ADGL: ADG throughout life.

Results of the subsequent genetic evaluations conducted with ssGBLUP are presented in Tables 4.3 and 4.4, respectively for the sire line and the dam line. Results in Tables 4.5 and 4.6 are from subsequent genetic evaluations done with PBLUP, respectively for the sire line and the dam line. For every parameter in these tables (i.e. estimated heritability, validation accuracy, level bias, and dispersion bias), we showed the estimate and SE of the estimate. We always used a one-tailed two-sample t-test at 5% significance level to determine whether two estimates were different. We included the estimated heritabilities in our results because they help in explaining the results of accuracy and bias.

Estimated heritabilities for ADGT and ADGL did not differ between reference and VGP scenarios, but increased in MGP scenarios. This was observed in both lines. For backfat thickness, heritabilities did not differ across GPS scenarios, neither in the sire line nor in the dam line. For loin depth, the heritabilities did not differ across GPS scenarios in the sire line. In the dam line however, the heritability of loin depth was

higher in MGP scenarios than in reference and VGP scenarios. The above trends in heritabilities for all traits were observed whether records on animals in the validation generation were included or excluded in estimating the heritabilities. Without records on animals in the validation generation in the subsequent ssGBLUP evaluations, all results (including estimated heritabilities) for reference and VGP scenarios were the same. The increases in heritabilities observed with more preselection were generally due to decreases in residual variances with more preselection, while additive genetic variances generally did not differ across GPS scenarios (Tables S4.2 and S4.3).

4.3.1 Subsequent ssGBLUP evaluations with records on animals in the validation generation

Validation accuracies did not differ across GPS scenarios for all traits in both lines, except for ADGT in the dam line, where the accuracy was lower in MGP scenario than in reference scenario (Table 4.4). Tendencies (i.e. indications that may not be statistically significant) towards lower accuracies with more GPS were however observed for all traits in both lines (Tables 4.3 and 4.4). In both lines, level bias was absent in all scenarios for loin depth, and only in some scenarios for ADGT, ADGL and backfat thickness. Even when level bias was present, it was still only marginal. The highest value of level bias recorded was -0.17 additive genetic SD units, under the VGP scenario for ADGL in the sire line (Table 4. 3). Dispersion bias was absent (i.e. the regression coefficient of PYD on GEBV did not differ from its expected value of 0.5) for all traits in the sire line (Table 4.3), except in VGP and MGP scenarios for backfat thickness, where there was inflation (i.e. the regression coefficient was less than 0.5). On the other hand, dispersion bias was present for all traits in the dam line (Table 4.4), except in the reference scenario of ADGT and MGP scenario of loin depth. In the instances with dispersion bias in the dam line, the regression coefficients were greater than 0.5 (i.e. they were deflated) in reference and VGP scenarios of loin depth, and less than 0.5 in all other instances. However, the estimates of the regression coefficients did not differ across GPS scenarios within traits and lines (Tables 4.3 and 4.4), although they showed tendencies to decrease from reference to VGP to MGP scenarios.

Table 4.3 Performance of ssGBLUP[a] in the subsequent evaluations in the sire line (SE in brackets)

Measure/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference	VGP	MGP
Average daily gain during performance testing, size of validation population = 1382						
Estimated heritability[e]	0.24 (0.01)	0.25 (0.01)	0.33 (0.02)	0.24 (0.01)	0.24 (0.01)	0.35 (0.03)
Validation accuracy[f]	0.51 (0.02)	0.51 (0.02)	0.50 (0.02)	0.47 (0.02)	0.47 (0.02)	0.44 (0.02)
Level bias[g]	-0.09 (0.02)	-0.15 (0.02)	-0.01 (0.02)	-0.11(0.02)	-0.11(0.02)	-0.02 (0.02)
Dispersion bias[h]	0.48 (0.02)	0.49 (0.02)	0.48 (0.02)	0.48 (0.02)	0.48 (0.02)	0.46 (0.03)
Average daily gain throughout life, size of validation population = 1383						
Estimated heritability	0.26 (0.01)	0.28 (0.01)	0.33 (0.03)	0.27 (0.01)	0.27 (0.01)	0.35 (0.03)
Validation accuracy	0.57 (0.02)	0.56 (0.02)	0.55 (0.02)	0.52 (0.02)	0.52 (0.02)	0.48 (0.02)
Level bias	-0.10 (0.02)	-0.17 (0.02)	-0.06 (0.02)	-0.14 (0.02)	-0.14 (0.02)	-0.08 (0.02)
Dispersion bias	0.48 (0.02)	0.49 (0.02)	0.50 (0.02)	0.47 (0.02)	0.47 (0.02)	0.49 (0.02)
Backfat thickness, size of validation population = 1383						
Estimated heritability	0.58 (0.01)	0.58 (0.01)	0.58 (0.02)	0.58 (0.01)	0.58 (0.01)	0.60 (0.03)
Validation accuracy	0.69 (0.01)	0.68 (0.01)	0.67 (0.01)	0.63 (0.02)	0.63 (0.02)	0.56 (0.02)
Level bias	-0.02 (0.01)	-0.03 (0.01)	-0.03 (0.01)	-0.05 (0.01)	-0.05 (0.01)	-0.09 (0.01)
Dispersion bias	0.48 (0.01)	0.47 (0.01)	0.47 (0.01)	0.44 (0.01)	0.44 (0.01)	0.42 (0.02)
Loin depth, size of validation population = 1383						
Estimated heritability	0.55 (0.01)	0.55 (0.01)	0.55 (0.03)	0.55 (0.01)	0.55 (0.01)	0.57 (0.03)
Validation accuracy	0.68 (0.01)	0.67 (0.01)	0.65 (0.02)	0.62 (0.02)	0.62 (0.02)	0.54 (0.02)
Level bias	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.01 (0.01)
Dispersion bias	0.50 (0.01)	0.50 (0.01)	0.48 (0.02)	0.48 (0.02)	0.48 (0.02)	0.45 (0.02)

[a] single-step genomic best linear unbiased prediction. [b] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] The heritability was estimated from an equivalent pedigree-based animal model in ASReml. [f] Computed as weighted Pearson's correlation coefficient between progeny yield deviation and genomic estimated breeding value of all validation animals. [g] Computed as the weighted mean difference between progeny yield deviation and half of the genomic estimated breeding value across all validation animals, expressed in additive genetic SD units of the trait. [h] Measured by the weighted regression coefficient of progeny yield deviation on genomic estimated breeding value of all validation animals.

Table 4.4 Performance of ssGBLUP[a] in the subsequent evaluations in the dam line (SE in brackets)

Measure/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference	VGP	MGP
Average daily gain during performance testing, size of validation population = 2323						
Estimated heritability[e]	0.31 (0.01)	0.32 (0.01)	0.40 (0.02)	0.30 (0.01)	0.30 (0.01)	0.38 (0.02)
Validation accuracy[f]	0.35 (0.02)	0.31 (0.02)	0.29 (0.02)	0.28 (0.02)	0.28 (0.02)	0.23 (0.02)
Level bias[g]	-0.05 (0.02)	-0.14 (0.02)	0.04 (0.02)	0.03 (0.02)	0.03 (0.02)	0.14 (0.02)
Dispersion bias[h]	0.46 (0.03)	0.43 (0.03)	0.41 (0.03)	0.44 (0.03)	0.44 (0.03)	0.43 (0.04)
Average daily gain throughout life, size of validation population = 2405						
Estimated heritability	0.31 (0.01)	0.33 (0.01)	0.43 (0.02)	0.31 (0.01)	0.31 (0.01)	0.44 (0.02)
Validation accuracy	0.46 (0.02)	0.42 (0.02)	0.42 (0.02)	0.38 (0.02)	0.38 (0.02)	0.35 (0.02)
Level bias	-0.06 (0.01)	-0.16 (0.01)	-0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.08 (0.01)
Dispersion bias	0.45 (0.02)	0.42 (0.02)	0.42 (0.02)	0.43 (0.02)	0.43 (0.02)	0.43 (0.02)
Backfat thickness, size of validation population = 2312						
Estimated heritability	0.51 (0.01)	0.51 (0.01)	0.51 (0.02)	0.51 (0.01)	0.51 (0.01)	0.53 (0.02)
Validation accuracy	0.52 (0.01)	0.50 (0.02)	0.50 (0.02)	0.45 (0.02)	0.45 (0.02)	0.42 (0.02)
Level bias	0.02 (0.01)	-0.01 (0.01)	-0.03 (0.01)	0.02 (0.01)	0.02 (0.01)	-0.01 (0.01)
Dispersion bias	0.43 (0.01)	0.41 (0.01)	0.41 (0.01)	0.42 (0.02)	0.42 (0.02)	0.41 (0.02)
Loin depth, size of validation population = 1164						
Estimated heritability	0.50 (0.01)	0.50 (0.01)	0.55 (0.02)	0.49 (0.01)	0.49 (0.01)	0.53 (0.02)
Validation accuracy	0.62 (0.02)	0.60 (0.02)	0.59 (0.02)	0.55 (0.02)	0.56 (0.02)	0.49 (0.02)
Level bias	-0.02 (0.02)	-0.03 (0.02)	0.02 (0.02)	-0.04 (0.02)	-0.04 (0.02)	0.03 (0.02)
Dispersion bias	0.54 (0.02)	0.54 (0.02)	0.52 (0.02)	0.53 (0.02)	0.53 (0.02)	0.51 (0.03)

[a] single-step genomic best linear unbiased prediction. [b] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] The heritability was estimated from an equivalent pedigree-based animal model in ASReml. [f] Computed as weighted Pearson's correlation coefficient between progeny yield deviation and genomic estimated breeding value of all validation animals. [g] Computed as the weighted mean difference between progeny yield deviation and half of the genomic estimated breeding value across all validation animals, expressed in additive genetic SD units of the trait. [h] Measured by the weighted regression coefficient of progeny yield deviation on genomic estimated breeding value of all validation animals.

Table 4.5 Performance of PBLUP[a] in the subsequent evaluations in the sire line (SE in brackets)

Measure/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference	VGP	MGP
Average daily gain during performance testing, size of validation population = 1382						
Estimated heritability	0.24 (0.01)	0.25 (0.01)	0.33 (0.02)	0.24 (0.01)	0.24 (0.01)	0.35 (0.03)
Validation accuracy[e]	0.51 (0.02)	0.50 (0.02)	0.49 (0.02)	0.41 (0.02)	0.41 (0.02)	0.40 (0.02)
Level bias[f]	-0.04 (0.02)	-0.11 (0.02)	0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.01 (0.02)
Dispersion bias[g]	0.53 (0.02)	0.54 (0.03)	0.48 (0.02)	0.55 (0.03)	0.55 (0.03)	0.49 (0.03)
Average daily gain throughout life, size of validation population = 1383						
Estimated heritability	0.26 (0.01)	0.28 (0.01)	0.33 (0.03)	0.27 (0.01)	0.27 (0.01)	0.35 (0.03)
Validation accuracy	0.58 (0.02)	0.56 (0.02)	0.54 (0.02)	0.47 (0.02)	0.47 (0.02)	0.44 (0.02)
Level bias	-0.06 (0.02)	-0.14 (0.02)	-0.04 (0.02)	-0.05 (0.02)	-0.05 (0.02)	-0.05 (0.02)
Dispersion bias	0.55 (0.02)	0.55 (0.02)	0.51 (0.02)	0.56 (0.03)	0.56 (0.03)	0.54 (0.03)
Backfat thickness, size of validation population = 1383						
Estimated heritability	0.58 (0.01)	0.58 (0.01)	0.58 (0.02)	0.58 (0.01)	0.58 (0.01)	0.60 (0.03)
Validation accuracy	0.67 (0.01)	0.66 (0.02)	0.66 (0.02)	0.48 (0.02)	0.48 (0.02)	0.46 (0.02)
Level bias	-0.03 (0.01)	-0.03 (0.01)	-0.03 (0.01)	-0.09 (0.01)	-0.09 (0.01)	-0.10 (0.01)
Dispersion bias	0.50 (0.01)	0.50 (0.02)	0.50 (0.02)	0.46 (0.02)	0.46 (0.02)	0.43 (0.02)
Loin depth, size of validation population = 1383						
Estimated heritability	0.55 (0.01)	0.55 (0.01)	0.55 (0.03)	0.55 (0.01)	0.55 (0.01)	0.57 (0.03)
Validation accuracy	0.66 (0.02)	0.65 (0.02)	0.64 (0.02)	0.49 (0.02)	0.49 (0.02)	0.46 (0.02)
Level bias	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)
Dispersion bias	0.50 (0.02)	0.49 (0.02)	0.49 (0.02)	0.48 (0.02)	0.48 (0.02)	0.46 (0.02)

[a] Pedigree-based best linear unbiased prediction. [b] The subsequent PBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] Validation accuracy was computed as weighted Pearson's correlation coefficient between progeny yield deviation and estimated breeding value of all validation animals. [f] Level bias was computed as the weighted mean difference between progeny yield deviation and half of the estimated breeding value across all validation animals, expressed in additive genetic SD units of the trait. [g] Dispersion bias was measured by the weighted regression coefficient of progeny yield deviation on estimated breeding value of all validation animals.

Table 4.6 Performance of PBLUP[a] in the subsequent evaluations in the dam line (SE in brackets)

Measure/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference	VGP	MGP
Average daily gain during performance testing, size of validation population = 2323						
Estimated heritability	0.31 (0.01)	0.32 (0.01)	0.40 (0.02)	0.30 (0.01)	0.30 (0.01)	0.38 (0.02)
Validation accuracy[e]	0.35 (0.02)	0.30 (0.02)	0.30 (0.02)	0.24 (0.02)	0.24 (0.02)	0.21 (0.02)
Level bias[f]	-0.04 (0.02)	-0.16 (0.02)	0.01 (0.02)	0.08 (0.02)	0.08 (0.02)	0.13 (0.02)
Dispersion bias[g]	0.52 (0.03)	0.45 (0.03)	0.42 (0.03)	0.50 (0.04)	0.50 (0.04)	0.45 (0.04)
Average daily gain throughout life, size of validation population = 2405						
Estimated heritability	0.31 (0.01)	0.33 (0.01)	0.43 (0.02)	0.31 (0.01)	0.31 (0.01)	0.44 (0.02)
Validation accuracy	0.48 (0.01)	0.43 (0.02)	0.43 (0.02)	0.34 (0.02)	0.34 (0.02)	0.31 (0.02)
Level bias	-0.05 (0.01)	-0.18 (0.01)	-0.03 (0.01)	0.05 (0.02)	0.05 (0.02)	0.07 (0.01)
Dispersion bias	0.51 (0.02)	0.47 (0.02)	0.44 (0.02)	0.51 (0.03)	0.51 (0.03)	0.44 (0.03)
Backfat thickness, size of validation population = 2312						
Estimated heritability	0.51 (0.01)	0.51 (0.01)	0.51 (0.02)	0.51 (0.01)	0.51 (0.01)	0.53 (0.02)
Validation accuracy	0.52 (0.02)	0.50 (0.02)	0.50 (0.02)	0.37 (0.02)	0.37 (0.02)	0.36 (0.02)
Level bias	0.02 (0.01)	0.00 (0.01)	-0.03 (0.01)	0.04 (0.01)	0.04 (0.01)	0.00 (0.01)
Dispersion bias	0.45 (0.02)	0.43 (0.02)	0.42 (0.02)	0.41 (0.02)	0.41 (0.02)	0.39 (0.02)
Loin depth, size of validation population = 1164						
Estimated heritability	0.50 (0.01)	0.50 (0.01)	0.55 (0.02)	0.49 (0.01)	0.49 (0.01)	0.53 (0.02)
Validation accuracy	0.58 (0.02)	0.56 (0.02)	0.56 (0.02)	0.43 (0.02)	0.43 (0.02)	0.41 (0.02)
Level bias	0.00 (0.02)	-0.01 (0.02)	0.04 (0.02)	-0.02 (0.02)	-0.02 (0.02)	0.04 (0.02)
Dispersion bias	0.55 (0.02)	0.54 (0.02)	0.51 (0.02)	0.57 (0.03)	0.57 (0.03)	0.52 (0.03)

[a] Pedigree-based best linear unbiased prediction. [b] The subsequent PBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] Validation accuracy was computed as weighted Pearson's correlation coefficient between progeny yield deviation and estimated breeding value of all validation animals. [f] Level bias was computed as the weighted mean difference between progeny yield deviation and half of the estimated breeding value across all validation animals, expressed in additive genetic SD units of the trait. [g] Dispersion bias was measured by the weighted regression coefficient of progeny yield deviation on estimated breeding value of all validation animals.

4.3.2 Subsequent ssGBLUP evaluations without records on animals in the validation generation

For ADGT, validation accuracy did not differ across GPS scenarios in the sire line (Table 4.3), but in the dam line it was lower in MGP scenario than in reference/VGP scenarios (Table 4.4). For ADGL, validation accuracies did not differ across GPS scenarios in both lines, but only tended to decrease from reference/VGP to MGP scenarios (Tables 4.3 and 4.4). For backfat thickness, validation accuracy in the sire line decreased from reference/VGP scenarios to MGP scenario, but did not differ across GPS scenarios in the dam line. For loin depth, validation accuracies in both lines decreased from reference/VGP to MGP scenarios. Level bias was present in reference and VGP scenarios for ADGT in the sire line, but absent in MGP scenario. The reverse is the case in the dam line, where level bias was absent in reference and VGP scenarios for ADGT, and present in MGP scenario. For ADGL, level bias was present in all GPS scenarios in the sire line, and in the dam only present in the MGP scenario. For backfat thickness, level bias was present across all GPS scenarios in the sire line, and absent across all GPS scenarios in the dam line. For loin depth, level bias was absent across all GPS scenarios in both lines. Although level bias was present in many scenarios, it was still only marginal, with ±0.14 additive genetic SD units being its highest estimate (Tables 4.3 and 4.4). For ADGT and ADGL, dispersion bias was absent across GPS scenarios in the sire line, and present across GPS scenarios in the dam line. For backfat thickness, dispersion bias was present across all scenarios in both lines. For loin depth, dispersion bias was absent across all scenarios in both lines, except in the MGP scenario in the sire line, where there was inflation. Here too, just as when records on animals in the validation generation were included in the subsequent evaluations, the estimates of the regression coefficients did not differ across all GPS scenarios within traits, and this was observed for all traits in both lines.

4.3.2 Subsequent genetic evaluations with PBLUP

When records on animals in the validation generation were included in the subsequent evaluations, corresponding validation accuracies did not differ between PBLUP and ssGBLUP. This was observed for all traits in both lines. However, when records on animals in the validation generation were excluded from the subsequent evaluations in the sire line, validation accuracies of all traits were lower with PBLUP than with ssGBLUP. When records on animals in the validation generation were excluded from the subsequent evaluations in the dam line, validation accuracies did not differ between PBLUP and ssGBLUP for ADGT and ADGL, but were lower with PBLUP than with ssGBLUP for backfat thickness and loin depth. Just as validation

accuracies from subsequent ssGBLUP evaluations, validation accuracies from subsequent PBLUP evaluations in most instances did not differ across corresponding GPS scenarios. Just like when the subsequent evaluations were done with ssGBLUP, here too, level bias with PBLUP in most cases did not differ from its corresponding value with ssGBLUP, and in many instances it was not different from zero. The highest value of level bias when the subsequent evaluations were done with PBLUP was -0.18 additive genetic SD units (i.e. in the VGP scenario for ADGL in the dam line when records on animals in the validation included were included in the subsequent evaluation; Table 4.6). Regression coefficients of PYD on (G)EBV in most instances did not differ between PBLUP and ssGBLUP, or from their expected value of 0.5. However, with PBLUP, the regression coefficients were sometimes bigger than 0.5 (e.g. in reference and VGP scenarios for ADGL in the sire line, and in reference and VGP scenarios for loin depth in the dam line). The regression coefficients were also in some instances bigger with PBLUP than with ssGBLUP (e.g. in reference and VGP scenarios for ADGT in the sire line, and in reference and VGP scenarios for ADGL in both lines). The regression coefficients with PBLUP were lower in MGP scenarios than in reference scenarios for ADGT and ADGL in the dam line when records on animals in the validation generation was included in the subsequent evaluations (Table 4.6). This is unlike with ssGBLUP, where the regression coefficients in all instances did not differ across corresponding GPS scenarios (Tables 4.3 and 4.4).

4.4 Discussion

In this study, we investigated the impact of genomic preselection (GPS) on accuracy and bias in subsequent ssGBLUP evaluations of preselected animals. We used data from a commercial pig breeding program in which preselection has taken place, and retrospectively implemented additional layers of GPS. The data was on production traits of pigs from one sire line and one dam line. Per line, we implemented three GPS scenarios. We used precorrected phenotypes as records in the subsequent genetic evaluations, and progeny yield deviation (PYD) as the proxy for TBV. We did the subsequent genetic evaluations either with or without records on animals in the validation generation, and in all cases without progeny of validation animals. Validation accuracy decreased, or at least tended to, with more GPS. Dispersion bias was largely absent, and the regression coefficient of PYD on GEBV - the indicator of dispersion bias – in all instances did not differ among corresponding GPS scenarios. Level bias was also largely absent, and mean PYD minus Mean GEBV - the indicator of level bias - in most instances did not differ across GPS scenarios also. The above results were observed in both lines, for all traits, and whether records on animals in

the validation generation were included in or excluded from the subsequent ssGBLUP evaluations.

Empirically assessing the impact of preselection on subsequent genetic evaluations of preselected animals requires comparison of scenarios taking into account different approaches, including scenarios without preselection. Since some GPS had already taken place in the dataset we used for this study, it was not possible to have a scenario without preselection. We therefore needed to come up with another way of investigating whether ssGBLUP is able to estimate GEBV in the subsequent evaluation of preselected animals without preselection bias in our current dataset and by extension in real breeding programs. We hypothesized that if ssGBLUP in subsequent evaluations yields unbiased GEBV for preselected animals despite additional GPS in the current dataset (i.e. in our VGP and MGP scenarios), then it also yields unbiased GEBV for preselected animals in the subsequent evaluations with the current dataset (i.e. in our reference scenarios). This is why we implemented the VGP and MGP scenarios, to implement additional layers of GPS over the regular GPS already implemented by the commercial pig breeding program. While the equivalent of VGP and MGP do not happen in real breeding programs, implementing these GPS scenarios in this study enabled us to investigate the impact of GPS on subsequent genetic evaluations of preselected animals using real data, by including different amounts of pedigree, genomic and phenotypic information in the subsequent genetic evaluations. Our results have shown that ssGBLUP in subsequent evaluations of pigs is indeed able to estimate GEBV of preselected animals without preselection bias. We believe that the findings of the current study can be extended to other animal breeding programs. This is because preselection, irrespective of its type and intensity and in how many stages it is implemented, has similar effects across breeding programs (i.e. preselection ensures that only better-than-average animals are phenotyped for the traits measured at advanced stages of lives of animals). The findings of the current study are in line with what we showed in our previous studies using simulated datasets [3,4], that in subsequent evaluations ssGBLUP estimates unbiased GEBV for preselected animals. In [3], ssGBLUP in subsequent evaluations estimated GEBV of preselected animals with accuracy loss compared to scenarios without preselection. We attributed the accuracy loss in scenarios with preselection to less numbers of sibs with records compared to the scenarios without preselection. In this study too, accuracy decreased, or at least tended to, with more preselection. This decrease can also be attributed to less numbers of relatives with records compared to the in scenarios with less preselection, as can be seen in Table 4.1.

The preselection we implemented in this study as well as in [3] and [4] are forms of non-ignorable selection (e.g. [9,31–33]). In the absence of genomic information, all the information utilized at these preselection stages need to be included in subsequent evaluations in order to avoid preselection bias (e.g. [9,12,34,35]). In our previous studies [3,4], we showed that ssGBLUP in subsequent evaluations of preselected animals estimates GEBV of preselected animals without preselection bias even if genotypes of preculled animals are not included. In [4], we also showed that genotypes of preculled animals are only needed in the subsequent ssGBLUP evaluations of their preselected sibs if their parents are not genotyped. Still in [4], we suggested that ssGBLUP uses genotypes of preselected animals and their parents to estimate the on-average positive Mendelian sampling terms of the preselected animals, and this enables ssGBLUP in subsequent evaluations to estimate GEBV of preselected animals without preselection bias. In this study preselected animals and their parents were genotyped, and indeed we did not observe preselection bias in our subsequent ssGBLUP evaluations despite not including genotypes of preculled animals in the subsequent evaluations.

4.4.1 Comparison of results across preselection scenarios and between ssGBLUP and PBLUP

We have shown that for all traits, and especially for ADGT and ADGL, estimated heritabilities increased or at least tended to, with more GPS. These tendencies of the heritabilities of these traits to increase with more preselection were because changes in residual variances were bigger than their corresponding changes in additive genetic variances moving from reference to VGP to MGP scenarios (Tables S4.2 and S4.3). The likely explanation for this observation is that animals may have low phenotypes for non-genetic reasons such as injury, social stress or illness. This 'dilutes' the heritability. Typically, such poor-performing animals are not selected, and this results to selected groups of animals being more homogeneous in expressing their genetic potentials than unselected groups. This suggests heterogeneity of residual variances across herd-year classes for all traits, with the greatest heterogeneity in the reference scenarios. In both lines, we repeated the subsequent evaluations of the reference scenarios with records on animals in the validation generations included, with a model that corrects for heterogenous residual variances (results not shown; [20]). Since we found that estimates of accuracy and bias in these scenarios did not differ between our two models, we decided to continue with the simpler model (i.e. the model in eq. 3).

We have also shown that with both ssGBLUP and PBLUP, validation accuracy decreased or at least tended to, with more GPS. These tendencies can be explained by the fact that the amount of phenotypic information also reduced in that order (Table 4.1). The observed tendencies of heritabilities to increase with more preselection could have influenced, at least partly, the magnitudes of decreases in accuracies with decreases in amounts of phenotypic information due to preselection. This can contribute to explaining why decreases in validation accuracies with more GPS were in most instances not statistically significant. Although there were always tendencies that corresponding validation accuracies were higher with ssGBLUP than with PBLUP, we observed that the differences were in most instances not statistically significant. The fact that heritabilities were all relatively high (ranging from 0.24 to 0.58, Tables 4.3 to 4.6) explains, at least partly, the absence of significant differences between ssGBLUP and PBLUP evaluations when records on animals in the validation generation were included in the subsequent genetic evaluations. It is a common knowledge that the higher the heritability, the higher the importance of own performance information and the lesser the importance of genomic information in genetic evaluations (e.g. [17]).

We have shown that in this study, level bias was absent in most instances, and even when it was present it was only marginal. We have also shown that the measure of level bias (i.e. the difference between mean PYD and mean (G)EBV among validation animals) in most instances did not differ across corresponding GPS scenarios, regardless of whether ssGBLUP or PBLUP were used in the subsequent evaluations. In our previous study [3], we observed no level bias when ssGBLUP was used in subsequent genetic evaluations, irrespective of type or intensity of preselection. However, in [3], we found level bias to be increasing with intensity of preselection when we used PBLUP in subsequent genetic evaluations. Patry et al [1,10,11] also reported significant level bias when subsequent genetic evaluations of genomically preselected animals were done with PBLUP, except when some pseudo-phenotypic information on preculled animals was included in the subsequent PBLUP evaluations. Just like with level bias, we have shown that in this study dispersion bias was absent in most instances, regardless of whether ssGBLUP or PBLUP was used in the subsequent evaluations. In our previous study with a simulated dataset [3], we found that regression coefficients of TBV on (G)EBV - the indicator of dispersion bias - were bigger and closer to the expected value of 1 when ssGBLUP was used in the subsequent genetic evaluations compared to when PBLUP was used. In [3], we also found that the regression coefficient became smaller as preselection intensity increased when PBLUP was used, but did not differ when ssGBLUP was used,

irrespective of preselection intensity. Preselection and subsequent selection were multi-trait in this study, and single trait in the previous studies [1,3,10,11]. This means that the chance of having multiple litter mates left in the data after preselection and subsequent selection is higher in the current study than in the previous studies. With multiple litter mates with records in the data, MS terms can be estimated reasonably well even in the absence of genomic information. This is likely the reason we did not observe level or dispersion bias in subsequent PBLUP evaluations of preselected animals in this study.

In the absence of selection, the expectation of regression coefficient of PYD on (G)EBV is 0.5, because PYD only represents half of the breeding value of the parent. However, when validation animals are on average better than the average of their age group, the expectation of the regression coefficient decreases in single-trait subsequent evaluations, depending on how much the validation animals deviate from the average of their age group (e.g. [36,37]). In the data used in this study, ADGT and ADGL had heavier weights in the breeding goals of the two lines than backfat thickness and loin depth, so we expected that our GPS would have smaller impacts on the regression coefficients for backfat thickness and loin depth than for ADGT and ADGL. We however did not observe smaller regression coefficients or regression coefficients that were further away from 0.5 for ADGT and ADGL than for backfat thickness and loin depth, neither with ssGBLUP nor with PBLUP. As explained in the previous paragraph, the fact that we implemented multi-trait subsequent evaluations in this study, as opposed to the single-trait subsequent evaluations in [36] and [37], could explain the differences between these two groups of studies.

4.4.2 Comparison of results across the two lines

Even in the dam line where the original GPS was at least numerically more intense and ratio of males with records to females with records in any generation was about 20:80, we generally did not observe significant decrease in validation accuracy, or increase in level and dispersion biases across our GPS scenarios. We however found that corresponding validation accuracies for ADGT, ADGL and backfat thickness were always higher in the sire line than in the dam line, despite the corresponding estimated heritabilities for ADGT and ADGL in many instances being higher in the dam line than in the sire line. These higher accuracies in the sire line than in the dam line may be explained by the relatively higher phenotyping and genotyping rates in the sire line than in the dam line (Table 4.1), meaning that validation animals in the sire line had more relatives with phenotypes and/or genotypes than validation animals in the dam lines.

4.4.3 Genotypes of preculled animals did not affect the subsequent ssGBLUP evaluations

In the subsequent ssGBLUP evaluations without records on animals in the validation generation, results from corresponding reference and VGP scenarios are exactly the same, at least up to two decimal places (Tables 4.3 and 4.4). However, in terms of data content, reference scenarios contained genotypes of the animals preculled in the corresponding VGP scenarios, in addition to all the data contained in the corresponding VGP scenarios (Table 4.1). The fact that results from these two scenarios are the same means that genotypes of the preculled animals did not affect the reference scenarios. In this study, most (about 95%) of the validation animals and their parents had genotypes. This is in line with the conclusion from our previous study [4], that genotypes of preculled animals are only useful in subsequent ssGBLUP evaluations of their preselected sibs when their parents are not genotyped.

4.4.4 Potential additional sources of bias in ssGBLUP from our data

In practical datasets as used in this study, it is difficult to completely rule out some mistakes in pedigree recording and in genotyping. At our genomic data quality control stage, genotypes of a few thousand animals were discarded because the animals did not meet the genomic data quality standard (of being genotyped for at least 90% of the SNP). Genotyping mistakes could still not be completely ruled out in the genomic data that passed quality control. In Tables 4.3 to 4.6, we saw that for some traits, heritabilities were different across the implemented GPS scenarios, although the animals in the base generation were the same. This implies that different subsets of the same data gave rise to different estimated (co)variance components in the base generation, and that it is likely that after some of the GPS scenarios were implemented, the estimated (co)variance components were different from their true values, at least for some of the traits. While these are all potential additional sources of bias in ssGBLUP evaluations, they are difficult to avoid in practice [13]. However, in general, we have shown that these potential additional sources of bias did not cause significant bias in our ssGBLUP evaluations, as both level and dispersion biases were in most cases absent, and even when present they were only marginal and mostly not different across corresponding GPS scenarios.

4.5 Conclusions

When subsequent genetic evaluations of preselected animals are done with ssGBLUP, genomic preselection in single or multiple generations only slightly decreases realized accuracy, and hardly causes level or dispersion bias. This

conclusion is expected to hold regardless of whether records on animals in the validation generation are included or excluded in the subsequent evaluations, and regardless of the weight of the trait in the breeding goal. Although these conclusions were derived using data from a pig breeding program, we believe that they can be generalized to other animal breeding programs, because preselection presumably has the same effect in any animal breeding program.

4.6 Acknowledgements

The authors thank Marco Bink and Katrijn Peeters from Hendrix Genetics, John Henshall from Cobb Europe, and Chris Schrooten and Gerben de Jong from CRV, for their inputs towards designing this study.

4.7 References

1. Patry C, Ducrocq V. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci. 2011;94:1011–20.
2. Masuda Y, VanRaden PM, Misztal I, Lawlor TJ. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in us holsteins. J Dairy Sci. 2018;101:5194–206.
3. Jibrila I, ten Napel J, Vandenplas J, Veerkamp RF, Calus MPL. Investigating the impact of preselection on subsequent single-step genomic blup evaluation of preselected animals. Genet Sel Evol. 2020;52:42.
4. Jibrila I, Vandenplas J, ten Napel J, Veerkamp RF, Calus MPL. Avoiding preselection bias in subsequent single-step genomic blup evaluations of genomically preselected animals. J Anim Breed Genet. 2021;138:432–441.
5. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker map. Genetics. 2001;157:1819–29.
6. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci. 2009;92:433–43.
7. Hayes BJ, Visscher PM, Goddard ME. Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res (Camb). 2009;91:47–60.
8. Su G, Guldbrandtsen B, Gregersen VR, Lund MS. Preliminary investigation on reliability of genomic estimated breeding values in the danish holstein population. J Dairy Sci. 2010;93:1175–83.
9. Henderson CR. Best linear unbiased estimation and prediction under a selection model. Biometris. 1975;31:423–47.
10. Patry C, Ducrocq V. Accounting for genomic pre-selection in national blup evaluations in dairy cattle. Genet Sel Evol. 2011;43:30.
11. Patry C, Jorjani H, Ducrocq V. Effects of a national genomic preselection on the international genetic evaluations. J Dairy Sci. 2013;96:3272–84.
12. Pollak EJ, van der Werf J, Quaas RL. Selection bias and multiple trait evaluation. J Dairy Sci. 1984;67:1590–5.

13. Tsuruta S, Lourenco DAL, Masuda Y, Misztal I, Lawlor TJ. Controlling bias in genomic breeding values for young genotyped bulls. J Dairy Sci. 2019;102:9956–70.
14. Vitezica ZG, Aguilar I, Misztal I, Legarra A. Bias in genomic predictions for populations under selection. Genet Res (Camb). 2011;93:357–66.
15. Hsu W-L, Garrick DJ, Fernando RL. The accuracy and bias of single-step genomic prediction for populations under selection. Genes|Genomes|Genetics. 2017;7:2685–94.
16. Westell RA, Quaas RL, Van Vleck LD. Genetic groups in an animal model. J Dairy Sci. 1988;71:1310–8.
17. Mrode RA. Linear models for the prediction of animal breeding values. 3rd ed. 2014.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
19. Gilmour AR, Gogel BJ, Cullis BR, Thompson R. ASReml user guide release 3.0. VSN Int. Ltd. 2009. https://asreml.kb.vsni.co.uk/wp-content/uploads/sites/3/2018/02/ASReml-3-User-Guide.pdf
20. ten Napel J, Vandenplas J, Lidauer M, Stranden I, Taskinen M, Mäntysaari E, et al. MiXBLUP: A user-friendly software for large genetic evaluation systems. 2020. https://www.mixblup.eu/documents/RvT_06335_ASG_WLR_MixBlup%20Manual_LR-spread.pdf
21. Gao H, Madsen P, Aamand GP, Thomasen JR, Sørensen AC. Bias in estimates of variance components in populations undergoing genomic selection : a simulation study. BMC Genomics. 2019;20:1–8.
22. Aldridge MN, Vandenplas J, Bergsma R, Calus MPL. Variance estimates are similar using pedigree or genomic relationships with or without the use of metafounders or the algorithm for proven and young animals. J Anim Sci. 2020;98.
23. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of holstein final score. J Dairy Sci. 2010;93:743–52.
24. Christensen OF, Lund MS. Genomic prediction when some animals are not genotyped. Genet Sel Evol. 2010;42:2.
25. Powell JE, Visscher PM, Goddard ME. Reconciling the analysis of ibd and ibs in complex trait studies. Nat Rev Genet. 2010;11:800–5.
26. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.
27. Cameron ND. Selection indices and prediction of genetic merit in animal breeding. 1997.
28. R Core Team. R: a language and environment for statistical computing. 2019.
29. Duenk P, Calus MPL, Wientjes YCJ, Breen VP, Henshall JM, Hawken R, et al. Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles. Genet Sel Evol. 2019;51:38.
30. Helske J. Diagis: Diagnostic plot and multivariate summary statistics of weighted

samples from importance sampling. R package version 0.1.5. 2020.
31. Schaeffer LR. Estimation of variance components under a selection model. J Dairy Sci. 1987;70:661–71.
32. Im S, Fernando R, Gianola D. Likelihood inferences in animal breeding under selection: a missing-data theory view point. Genet Sel Evol. 1989;21:399–414.
33. Gianola D, Fernando RL, Foulley JL. Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. Genome. 1989;31:768–77.
34. Appel LJ, Strandberg E, Danell B, Lundeheim N. Adjusting for missing data due to culling before testing in genetic evaluations of swine. J Anim Sci. 1998;76:1794–802.
35. Janhunen M, Kause A, Vehviläinen H, Nousiainen A, Koskinen H. Correcting within-family pre-selection in genetic evaluation of growth-a simulation study on rainbow trout. Aquaculture. 2014;434:220–6.
36. Mäntysaari EA, Liu Z, VanRaden P. Interbull validation test for genomic evaluations. Interbull Bull. 2010;17. https://journal.interbull.org/index.php/ib/article/view/1134
37. Mäntysaari EA, Koivula M. GEBV validation test revisited. Interbull Bull. 2012;45. https://journal.interbull.org/index.php/ib/article/view/1244

Supplementary information

Table S4.1 Estimated heritabilities (diagonal), genetic correlations (below diagonal) and phenotic correlations (above diagonals) for the traits utilized in this study, using the full data from the sire line. Standard errors are in brackets.

Traits	ADGT	ADGL	Backfat thickness	Loin depth
ADGT	**0.24 (0.01)**	0.91 (0.00)	0.20 (0.01)	-0.17 (0.01)
ADGL	0.92 (0.00)	**0.26 (0.01)**	0.20 (0.01)	-0.17 (0.01)
Backfat thickness	0.27 (0.02)	0.32 (0.02)	**0.58 (0.01)**	-0.04 (0.01)
Loin depth	-0.29 (0.02)	-0.30 (0.02)	-0.11 (0.02)	**0.55 (0.01)**

ADGT: Average daily gain during performance testing. ADGL: Average daily gain throughout life.

Table S4.2 Estimated additive genetic and residual variances (standard errors in bracket) in the sire line

Variance/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation	
	Reference[a]	VGP[b]	MGP[c]	Reference/VGP	MGP
Average daily gain during performance testing					
Additive genetic	2369 (101)	2486 (120)	2139 (189)	2357 (118)	2333 (228)
Residual	7548 (67)	7526 (80)	4293 (135)	7668 (80)	4254 (160)
Average daily gain throughout life					
Additive genetic	1048 (43)	1102 (51)	929 (82)	1057 (50)	961 (95)
Residual	2939 (27)	2878 (32)	1821 (58)	2922 (32)	1771 (66)
Backfat thickness					
Additive genetic	1.41 (0.04)	1.46 (0.05)	1.29 (0.08)	1.47 (0.05)	1.38 (0.10)
Residual	1.03 (0.02)	1.06 (0.02)	0.95 (0.04)	1.07 (0.02)	0.94 (0.05)
Loin depth					
Additive genetic	7.96 (0.22)	7.90 (0.25)	7.45 (0.49)	7.96 (0.26)	7.80 (0.58)
Residual	6.44 (0.11)	6.40 (0.13)	6.06 (0.27)	6.39 (0.13)	5.90 (0.32)

[a] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [b] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [c] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded.

Table S4.3 Estimated additive genetic and residual variances (standard errors in bracket) in the dam line

Variance/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation	
	Reference[a]	VGP[b]	MGP[c]	Reference/VGP	MGP
Average daily gain during performance testing					
Additive genetic	2542 (72)	2575 (83)	2605 (156)	2374 (80)	2392 (163)
Residual	5659 (44)	5343 (50)	3968 (102)	5430 (50)	3860 (109)
Average daily gain throughout life					
Additive genetic	844 (24)	885 (28)	928 (53)	835 (27)	933 (59)
Residual	1871 (14)	1796 (17)	1246 (34)	1825 (17)	1211 (37)
Backfat thickness					
Additive genetic	1.63 (0.03)	1.68 (0.04)	1.62 (0.08)	1.70 (0.04)	1.73 (0.10)
Residual	1.57 (0.02)	1.63 (0.02)	1.55 (0.05)	1.63 (0.02)	1.56 (0.06)
Loin depth					
Additive genetic	5.90 (0.14)	5.84 (0.16)	6.24 (0.33)	5.78 (0.16)	6.02 (0.34)
Residual	5.96 (0.07)	5.93 (0.08)	5.20 (0.19)	5.96 (0.08)	5.25 (0.20)

[a] In the reference scenario, the subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [b] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [c] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded.

This page intentionally left blank

Chapter 5

Impact of genomic preselection on subsequent ssGBLUP evaluation of preselected animals for scarcely recorded traits

Ibrahim Jibrila[1], Jan ten Napel[1], Jeremie Vandenplas[1], Rob Bergsma[2], Roel F Veerkamp[1] and Mario P L Calus[1]
[1] Wageningen University and Research Animal Breeding and Genomics, PO Box 338
6700 AH Wageningen, the Netherlands

[2] Topigs Norsvin Research Center B.V., Schoenaker 6, 6641 SZ Beuningen, the Netherlands

Under review for publication

Abstract

Background

We have previously shown that single-step genomic best linear unbiased prediction (ssGBLUP) estimates breeding values of genomically preselected animals without preselection bias, for widely-recorded traits. However, no study has investigated whether this also holds for scarcely-recorded traits, which generally have lower prediction accuracy than widely-recorded traits, mainly due to having smaller numbers of relatives with records. This study investigated the impact of genomic preselection (GPS) on accuracy and bias in ssGBLUP evaluation of genomically preselected animals for scarcely-recorded traits.

Methods

We used data on production traits from a commercial pig breeding program. We used feed intake as scarcely recorded target trait, and average daily gain, backfat thickness, and loin depth as widely-recorded predictor traits. The data contained the routine GPS implemented by commercial animal breeding programs, and we retrospectively implemented two scenarios with additional layers of GPS by discarding pedigree, genotypes and phenotypes of animals without progeny. In the next ssGBLUP evaluation following GPS, we used records only from the target trait, only from the predictor traits, or both.

Results

Accuracy of feed intake did not differ statistically across GPS scenarios, although it tended to decrease with more intense GPS. The accuracy had average values of 0.37, 0.44, and 0.45 across all GPS scenarios when only records of the target trait, of the predictor traits, and of both the target and the predictor traits were respectively utilized in the ssGBLUP evaluation. Bias was in most instances absent, and only marginal or did not differ statistically across GPS scenarios when present. The above results for accuracy and bias were observed whether records of the scarcely-recorded target trait, of the predictor traits, or both were utilized in the ssGBLUP evaluation.

Conclusion

Just like in evaluation of animals for widely-recorded traits, ssGBLUP in evaluation of animals for scarcely-recorded traits is able to estimate breeding values of preselected animals without preselection bias.

5.1 Background

The importance of recording phenotypes in animal breeding cannot be overemphasized. Some traits are measured routinely on the majority of animals in a breeding population. We refer to such traits as widely-recorded traits. Other traits, however, are difficult or expensive to measure, and are therefore only measured on a small proportion of animals in each generation of a breeding population. We refer to such traits as scarcely-recorded. Examples of scarcely-recorded traits include individual feed intake in all livestock species, and carcass quality traits in meat animals. The small numbers of animals with records for scarcely-recorded traits means that reference populations for genomic evaluation of animals for such traits are small as well, and this may result in genomic estimated breeding values (GEBV) with low accuracies (e.g. [1–3]). It has been shown that multi-trait evaluation of animals for scarcely-recorded traits together with predictor traits (widely-recorded traits that are moderately to highly genetically correlated with scarcely-recorded traits) gives more accurate genetic evaluation of scarcely-recorded traits compared to single-trait evaluation (e.g. [1–5]). Phenotypes of animals for predictor traits mainly help in genetic evaluation of scarcely-recorded traits by enabling more accurate estimation of Mendelian sampling (MS) terms of selection candidates, thereby increasing accuracy (e.g. [1–7]).

Selection of parents of the next generation involves multiple stages, such as a genomic preselection (GPS) of young selection candidates and a subsequent selection when the preselected candidates have records. Genetic evaluation models implicitly assume that the datasets analyzed are unselected or are random subsets of the unselected datasets. In reality however, the datasets analyzed at subsequent selection stages are neither unselected nor random samples of the unselected datasets, since preselection usually skews the distribution of the datasets (e.g. [8–11]). In our previous study [12], we investigated the impact of GPS on accuracy and bias in subsequent single-step genomic best linear unbiased prediction (ssGBLUP) evaluation of preselected animals, using data from a commercial pig breeding program. We were able to show that ssGBLUP in subsequent evaluation estimates GEBV of genomically preselected animals without preselection-related bias and accuracy loss. In this previous study [12], we studied impact of GPS in subsequent evaluation of animals for widely-recorded traits. The traits were average daily gain, backfat thickness, and loin depth, which are some of the most important widely recorded production traits in a typical pig sire line (e.g. [13]).

We hypothesize that the ability of ssGBLUP to prevent accuracy loss and bias due to preselection applies to both widely-recorded and scarcely-recorded traits. This is because we expect that the ability of ssGBLUP to estimate unbiased GEBV of preselected animals lies in its use of genotypes of the preselected animals and their parents to estimate the MS terms of preselected animals. The aim of this study was to validate this hypothesis, i.e. to investigate the impact of GPS on accuracy and bias in subsequent ssGBLUP evaluation of preselected animals for scarcely-recorded traits. Like in [12], we used a full dataset derived from a commercial pig breeding program as reference, and retrospectively implemented additional layers of GPS. Since in every generation GEBV was used to select the parents of the next generation, we implemented our additional layers of GPS by discarding pedigree, genotypes and phenotypes of the animals that did not have progeny in the data. We compared accuracy and bias of subsequent ssGBLUP evaluation using the remaining data after these additional layers of GPS against those obtained for the full available data.

5.2 Methods

5.2.1 Data

We obtained pig production traits data on a sire line from Topigs Norsvin, collected between 1970 and 2020. The data were on scarcely-recorded and widely-recorded production traits. The data included two scarcely-recorded feed intake traits - feed intake from the start to the end of performance testing (FISE), and feed intake from the middle to the end of performance testing (FIME). Animals could only have records for one of the two feed intake traits, and the number of animals that had records was slightly higher for FISE than for FIME. The widely-recorded traits were average daily gain during performance testing (ADGT), average daily gain throughout the lifetime (ADGL), backfat thickness, and loin depth. These production traits were included in the breeding goal of this line, which was the basis for (pre)selection. Details on the amount of data utilized in this study are in Table 5.1. The data were recorded on originally preselected animals (i.e. the animals routinely preselected by Topigs Norsvin). In the pedigree, animals with one or both parents missing were assigned to genetic groups (e.g. [14]), according to line and year of birth of each animal. We used the same data in our previous study [12], except that there we did not use FISE and FIME, and we additionally used a dataset from a dam line.

Table 5.1 Data utilized in subsequent ssGBLUP[a] evaluation following each preselection scenario, after quality control

Number of animals with	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[b]	VGP[c]	MGP[d]	Reference	VGP	MGP
Pedigree entry	81,875	60,950	12,777	81,875	60,950	12,777
FISE[e] records	12,136	8,648	248	8,610	8,610	210
FIME[e] records	10,607	8,389	240	8,257	8,257	108
ADGT[f] records	71,859	51,811	5,939	50,463	50,463	4591
ADGL[f] records	74,893	54,053	6,064	52,683	52,683	4,694
Backfat thickness records	74,411	53,674	6,058	52,304	52,304	4,688
Loin depth records	73,544	52,803	5,943	51,433	51,433	4,573
Records for ≥ 1 production trait	75,129	54,217	6,065	52,846	52,846	4,694
Genotypes	33,506	23,315	5,131	33,506	23,315	5,131

[a] Single-step genomic best linear unbiased prediction. [b] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] Feed intake from the start (FISE) or the middle (FIME) to the end of performance testing. [f] Average daily gain during performance testing (ADGT) or throughout life (ADGL).

5.2.2 Training and validation generations

We grouped all animals into two groups based on their birth dates. Animals born before or on 31st January, 2017 were used as training population. Animals born after 31st January, 2017 that met the following requirements were selected as validation animals: 1) their parents were born before or on 31st January, 2017, and 2) they had phenotyped progeny. The first requirement ensured that our validation animals were from only one generation, and the second requirement enabled comparing GEBV of the validation animals against their progeny yield deviation (PYD) [15]. Since records on validation animals were included in some of our subsequent evaluation scenarios (as will be seen later), we chose to use PYD as our proxy for true breeding value (TBV) because PYD is estimated from phenotypes that were not included in the subsequent genetic evaluation.

5.2.3 Genomic data and quality control

Our genomic data included genotypes of animals for about 21,000 segregating single nucleotide polymorphisms (SNP), distributed across the 18 autosomes in the pig genome. The SNP were genotyped using a custom SNP chip. We used Plink [16] for all quality control operations on our genomic data. Per GPS scenario (as described later), animals and SNP with call rates less than 90% were removed, as well as SNP that deviated from Hardy-Weinberg equilibrium (Hardy-Weinberg equilibrium exact test p value = 10^{-15}), or had a minor allele frequency below 0.005. The very low Hardy-Weinberg equilibrium exact test p value was used because the genomic data came from a selected population, and therefore some level of deviation from Hardy-

Weinberg equilibrium was expected. Across the GPS scenarios, the number of remaining SNP ranged from 20,550 to 20,963.

5.2.4 Computation of precorrected phenotypes

We used precorrected phenotypes as records in our genetic evaluations. To compute precorrected phenotypes (y_c), we first ran the following six-trait pedigree-based animal model:

$$\mathbf{y_j} = \mathbf{X_j b_j} + \mathbf{W_j p_j} + \mathbf{Z_j u_j} + \mathbf{e_j} \qquad (\mathbf{eq.\ 1}).$$

Where for every trait (j) $\mathbf{y_j}$ was the vector of phenotypes; $\mathbf{b_j}$ was the vector of fixed effects, with incidence matrix $\mathbf{X_j}$; $\mathbf{p_j}$ was the vector of non-genetic random effects, with incidence matrix $\mathbf{W_j}$; $\mathbf{u_j}$ was the vector of breeding values, with incidence matrix $\mathbf{Z_j}$; and $\mathbf{e_j}$ was the vector of residuals. The model assumed $\mathbf{u_j}$ and $\mathbf{e_j}$ to be normally distributed and uncorrelated, each with mean of zero. For all traits and across all animals, **u** and **e** had variance-covariance matrices **A** $\otimes$ **G** and **I** $\otimes$ **R**, respectively. Where **A** was the pedigree relationship matrix among animals, **I** was an identity matrix with dimensions equal to the number of animals with records, and **G** and **R** were respectively the trait by trait additive genetic and residual variance-covariance matrices. Then for every animal (i) with phenotype for trait j, we computed its precorrected phenotype (y_{cij}) as:

$$y_{cij} = \hat{u}_{ij} + \hat{e}_{ij} \qquad \textbf{(eq. 2).}$$

5.2.5 Preselection

We implemented a reference scenario and two scenarios with additional layers of GPS. The reference scenario only included the routine GPS implemented by Topigs Norsvin. Thus, the subsequent ssGBLUP evaluation following the reference scenario utilized the entire set of available data until the validation generation. The second scenario is called validation generation preselection (VGP) scenario, where we only implemented additional GPS in the validation generation, by discarding the pedigree, genotypes and phenotypes of all animals in the validation generation that had no progeny in the data. The third scenario is called multi-generation preselection (MGP) scenario, where we discarded pedigree, genotypes and phenotypes of any animal in the validation or training generations with no progeny in the data. An overview of all finally remaining animals for each of the GPS scenarios is shown in Figure 5.1.

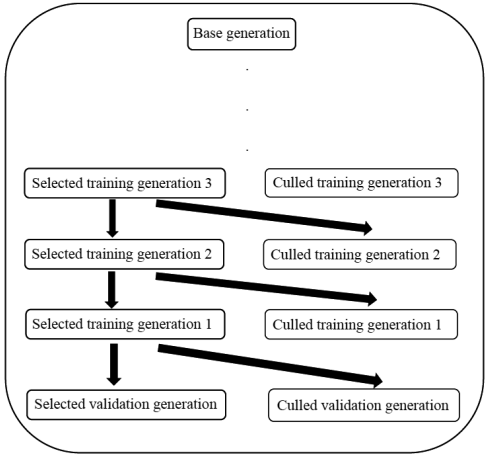


Figure 1a: Reference scenario

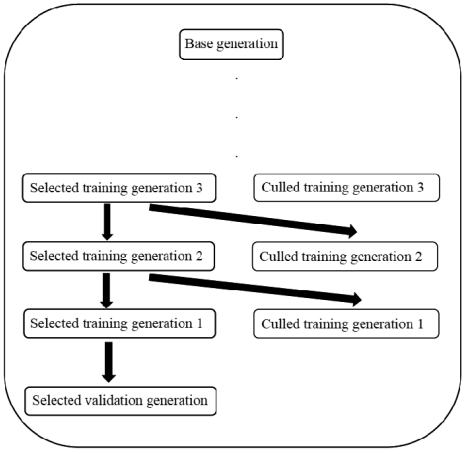


Figure 1b: Validation generation preselection (VGP) scenario

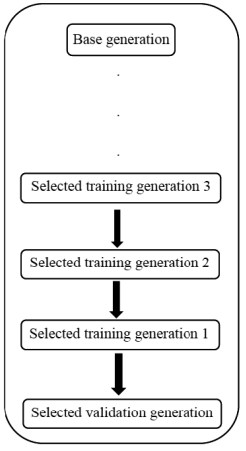


Figure 1c: Multi-generation preselection (MGP) scenario

Figure 5.1 Overview of groups of animals used in subsequent ssGBLUP evaluation of each of the considered GPS scenario

5.2.6 Subsequent ssGBLUP evaluation

5.2.6.1 Scenarios

Following each GPS scenario, we implemented a subsequent ssGBLUP evaluation with all animals that survived the GPS. We call this evaluation subsequent because it came after preselection, and preselection usually involves evaluating animals and then keeping some and removing some from the breeding program. Progeny of validation animals were not included in the subsequent ssGBLUP evaluation. The subsequent ssGBLUP evaluation was conducted with and without records on the animals in the validation generation (see Table 5.1), to represent traits with records (e.g., production traits) and those without records (e.g., reproduction traits) available on selection candidates at the time of subsequent evaluation. We assumed three situations in conducting the subsequent evaluation, where records were available for: i) only the scarcely recorded target traits, ii) only the predictor traits, and iii) both the scarcely-recorded target traits and the predictor traits. For the situation where the only records available were of the scarcely recorded target traits, we still analyzed FISE and FIME together, as this mimicked reality. We however decided to only report results of subsequent evaluation for FISE, as FIME had only 70 validation animals, while FISE had 944 validation animals.

5.2.6.2 Model

The multi-trait model used for estimation of GEBV for every trait (j) was:

$$\mathbf{y_j} = \mathbf{1_j}\mu_j + \mathbf{Z_j}\mathbf{u_j} + \mathbf{e_j} \qquad \textbf{(eq. 3)},$$

where for every trait (j), $\mathbf{y_j}$ was the vector of precorrected phenotypes; $\mathbf{1_j}$ was an incidence vector of 1's, and $\mathbf{Z_j}$ was an incidence matrix, linking precorrected phenotypes to overall mean and random animal effects, respectively; μ_j was the overall mean; $\mathbf{u_j}$ and $\mathbf{e_j}$ and their assumptions are the same as for **eq. 1**. However, the **A** in the variance-covariance matrix of **u** in **eq. 1** was replaced by **H** in **eq. 3,** and **H** was the combined genomic-pedigree relationship matrix among animals.

5.2.6.3 Setting up ssGBLUP

The inverse of the combined pedigree-genomic relationship ($\mathbf{H^{-1}}$) was obtained as follows [18,19]:

$$\mathbf{H^{-1}} = \mathbf{A^{-1}} + \begin{bmatrix} 0 & 0 \\ 0 & (0.95\mathbf{G_t} + 0.05\mathbf{A}_{22})^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \qquad \textbf{(eq. 4)},$$

where $\mathbf{A}^{-1}$ was the inverse of the pedigree relationship matrix, and $\mathbf{A}_{22}$ was part of

the pedigree relationship matrix referring to genotyped animals. We considered inbreeding in setting up both $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$, as ignoring inbreeding in setting up $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}$ has been reported to cause bias in GEBV [20]. The adjusted genomic relationship matrix $\mathbf{G_t}$ was computed as follows [21,22]:

$$\mathbf{G_t} = \left(1 - \bar{f}_p\right)\mathbf{G_r} + 2\bar{f}_p\mathbf{11}' \qquad \textbf{(eq. 5)},$$

where $\bar{f}_p$ was the average pedigree inbreeding coefficient across genotyped animals, $\mathbf{G_r}$ was the raw genomic relationship matrix computed following the first method of VanRaden [23], and $\mathbf{11}'$ was a matrix of 1s. As the animals with genotypes in this study were selectively genotyped, this transformation made sure that the impact of selective genotyping was taken care of and that **G** and $\mathbf{A_{22}}$ were on the same scale and therefore compatible [22,24]. To compute $\mathbf{G_r}$, we computed current allele frequencies using all available genomic data after quality control. We gave the weights of 0.95 to $\mathbf{G_t}$ and 0.05 to $\mathbf{A}_{22}$ to ensure that **G** was invertible [18,19].

For every scenario, the data used in the subsequent evaluation (as in Table 5.1) were used in a pedigree version of **eq. 3** in ASReml [17] to estimate scenario-specific variance components. We used these scenario-specific variance components in the subsequent genetic evaluation, to ensure that the variance components used were appropriate for the precorrected phenotypes. All estimations of breeding values were done using MiXBLUP [25].

5.2.6.4 Measures of accuracy and bias

We used progeny yield deviation (PYD) [15] as a proxy for true breeding value (TBV), against with which GEBV were compared when computing accuracy and bias. To compute PYD, we ran a multi-trait pedigree-based animal model per line in MiXBLUP, with precorrected phenotypes as records and an overall mean as the only fixed effect **(eq. 3)**. From the output of this analysis, we computed PYD for each validation sire and dam (i) for each trait (j) as:

$$\mathrm{PYD}_{ij} = \frac{\sum_{p=1}^{n} y_{cpj} - \frac{1}{2}a_{mj}}{n} \qquad \textbf{(eq. 6)},$$

where PYD_{ij} was the progeny yield deviation of a sire or dam i for trait j, y_{cpj} was the precorrected phenotype of a progeny p of the sire or dam i for trait j, a_{mj} was the breeding value of the mate of sire or dam i (for trait j) in producing offspring p, and n was the number of phenotyped progeny of sire or dam i. Estimation of PYD was

done before removing progeny of validation animals from the data. We computed approximate reliability of PYD for each validation animal for each trait, and used this approximate reliability as the weighting factor to compute accuracy and bias, to account for differences in number of progeny used to estimate PYD for different validation animals. The reliability of PYD was approximated as:

$$\frac{{}^{1}/_{4}\mathrm{n}\mathrm{h}^{2}}{1+{}^{1}/_{4}(\mathrm{n}-1)\mathrm{h}^{2}} \qquad \textbf{(eq. 7)},$$

where n was the validation animal's number of half-sib progeny with records, and h^2 was the heritability of the trait [26]. For convenience, we assumed all progeny of a validation animal were half-sibs, though some of them were full-sibs. We also computed unweighted accuracy an bias, and accuracy and bias weighted by approximate PYD reliabilities obtained considering all progeny as full sibs. We did not observe significant differences among the results, so we decided to only report the accuracy and bias weighted by approximate reliability computed considering all progeny as half sibs (as in eq. 7).

Validation accuracy was computed as weighted Pearson's correlation coefficient between PYD and GEBV of all validation animals, using the 'cor.test' function of the 'stats' package in R [27]. We computed the standard errors (SE) of the estimates from the confidence intervals (CI) produced by the 'cor.test' function. We computed two types of bias. The first type is level bias, which is a measure of whether genetic gain is over- or under-estimated. Level bias was computed as the weighted mean difference between PYD and half of the GEBV across all validation animals, expressed in additive genetic standard deviation (SD) units of the trait. We used the 'weighted.mean' function of the 'stats' package [27] to compute estimates of the weighted mean differences, and used the 'weighted_se' function of the 'diagis' package in R [28] to compute SE of the estimates. A negative difference means that GEBV are on average overestimated, and therefore genetic gain is overestimated, and vice versa. Since PYD were computed from a dataset that included information on progeny of validation animals and GEBV were computed without information on progeny of validation animals, PYD and GEBV were expressed against different bases. Therefore, before computing differences between PYD and half of the GEBV of validation animals, we made sure that PYD and GEBV were on the same scale. We did this in the following steps: from the model used in computing PYD, we computed average EBV across all animals in the first three training generations. We then subtracted half of this average EBV from PYD of each validation animal. Then from

each subsequent ssGBLUP evaluation, we computed the average GEBV of all animals in the first three training generations. We then subtracted this average GEBV from GEBV of each validation animal. The second type of bias we computed is dispersion bias. Dispersion bias was measured by the weighted regression coefficient of PYD on GEBV of all validation animals. We used the 'lm' function of the 'stats' package in R [27] to compute both the estimates and SE of the regression coefficients. If the regression coefficient is equal to the expected value, then there is no dispersion bias. Note that the expected value of regression coefficient of PYD on GEBV is 0.5, because PYD of an animal only includes half of its breeding value. A regression coefficient less than the expected value means that variance of GEBV is inflated, and vice versa. We always used a one-tailed two-sample t-test at 5% significance level to determine whether two estimates (of accuracy, bias and heritability) were different.

5.3 Results

5.3.1 Effectiveness of the additional GPS scenarios

Table 5.2 shows normalized means and SD of precorrected phenotypes of the traits analyzed in this study, following the implemented GPS scenarios. The means and SD were normalized by dividing them by the SD of their corresponding reference scenarios. The results show that our additionally implemented GPS was effective, as the lines were (pre)selected for, among others, increased feed efficiency (i.e. slightly higher feed intake, and much higher average daily gain), slightly decreased backfat, and slightly increased loin depth. As the validation generation for both VGP and MGP scenarios only contained the preselected animals, means and SD of the precorrected phenotypes of the traits are the same for these two preselection scenarios when only considering the animals in the validation generation. When only the validation generation was considered (i.e. the middle part of Table 5.2), the average precorrected phenotypes of FISE, FIME ADGT, and ADGL increased going from the reference scenario to the VGP/MGP scenario. At the same time, SD of the precorrected phenotypes of these traits decreased from reference to VGP/MGP scenarios. For backfat thickness, both the mean and the SD only slightly decreased from reference to VGP and MGP scenarios. The change was also only slight for loin depth, with the mean remaining the same at least up to two decimal places, and the SD slightly decreasing, from reference to VGP and MGP scenarios. Higher effectiveness of MGP over VGP can be seen from the right part of Table 5.2 (i.e. when the means and SD were computed across the entire data). For the positively (pre)selected traits (i.e FISE, FIME, ADGT, ADGL, and loin depth) means were higher

for MGP than for VGP. And for backfact thickness, which was negatively (pre)selected, the mean was lower for MGP than for VGP.

Table 5.2 Normalized[a] Means and standard deviations (in brackets) of precorrected phenotypes of the traits utilized in this study, following each genomic preselection scenario

Trait/Preselection scenario	Validation generation		Full data	
	Reference[b]	VGP[c] & MGP[d]	VGP	MGP
FISE[e] (g/day)	-0.08 (1)	0.14 (0.83)	-0.12 (1.00)	-0.01 (0.91)
FIME[e] (g/day)	-0.30 (1)	-0.11 (0.82)	-0.29 (1.00)	-0.10 (0.86)
ADGT[f] (g/day)	0.12 (1)	0.51 (0.80)	-0.08 (1.00)	0.38 (0.84)
ADGL[f] (g/day)	0.03 (1)	0.41 (0.85)	-0.11 (1.00)	0.31 (0.86)
Backfat thickness (mm)	-0.27 (1)	-0.29 (0.95)	-0.01 (1.00)	-0.10 (0.96)
Loin depth (mm)	0.26 (1)	0.26 (0.97)	0.17 (1.00)	0.20 (0.97)

[a] The values were normalized by dividing them by the standard deviations of their corresponding reference scenarios. [b] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [c] Validation generation preselection scenario, in which all animals in the validation generation without progeny in the data were discarded. [d] Multi-generation preselection (MGP) scenario, in which all animals in the validation and training generations without progeny in the data were discarded. [e] Feed intake from the start (FISE) or the middle (FIME) to the end of performance testing. [f] Average daily gain during performance testing (ADGT) or throughout life (ADGL).

5.3.2 Heritabilities and correlations among the traits

Table 5.3 shows estimated heritabilities (diagonal), genetic correlations (below diagonal) and phenotic correlations (above diagonals) for the traits analyzed in this study, using the full data. All the traits have moderate to high heritabilities - from 0.24 for ADGT to 0.58 for backfat thickness. The two feed intake traits have close-to-unity genetic and phenotypic correlations with each other (0.97 and 0.85, respectively). They also have moderate to high genetic correlations and low to moderate phenotypic correlations with the other traits (absolute values of genetic correlations ranged from 0.33 to 0.80, and absolute values of phenotypic correlations ranged from 0.16 to 0.78). The variance-covariance components used in computing these heritabilities and correlations were estimated using a PBLUP model, according to **equ. 3**.

Table 5.3 Estimated heritabilities (diagonal), genetic correlations (below diagonal) and phenotypic correlations (above diagonal) for the traits utilized in this study, using the full data

Traits	FISE[a]	FIME[a]	ADGT[b]	ADGL[b]	Backfat	Loin depth
FISE	**0.37 (0.02)**	0.85 (0.01)	0.70 (0.00)	0.72 (0.00)	0.43 (0.01)	-0.16 (0.01)
FIME	0.97 (0.02)	**0.35 (0.02)**	0.76 (0.00)	0.78 (0.00)	0.43 (0.01)	-0.19 (0.01)
ADGT	0.79 (0.02)	0.78 (0.02)	**0.24 (0.01)**	0.91 (0.00)	0.20 (0.01)	-0.17 (0.01)
ADGL	0.80 (0.02)	0.77 (0.02)	0.92 (0.00)	**0.26 (0.01)**	0.20 (0.01)	-0.17 (0.01)
Backfat	0.60 (0.02)	0.54 (0.03)	0.27 (0.02)	0.32 (0.02)	**0.58 (0.01)**	-0.04 (0.01)
Loin depth	-0.33 (0.03)	-0.35 (0.03)	-0.29 (0.02)	-0.30 (0.02)	-0.11 (0.02)	**0.55 (0.01)**

[a] Feed intake from the start (FISE) or the middle (FIME) to the end of performance testing. [b] Average daily gain during performance testing (ADGT) or throughout life (ADGL).

Table 5.4 Accuracy and bias in subsequent ssGBLUP evaluation of FISE

Measure/Preselection scenario	With records on animals in the validation generation			Without records on animals in the validation generation		
	Reference[a]	VGP[b]	MGP[c]	Reference	VGP	MGP
Estimated heritability	0.37 (0.02)	0.35 (0.02)	0.24 (0.09)	0.36 (0.02)	0.36 (0.02)	0.32 (0.10)
When only records of the scarcely-recorded target traits were utilized[d]						
Validation accuracy	0.45 (0.03)	0.40 (0.03)	0.32 (0.03)	0.39 (0.03)	0.39 (0.03)	0.24 (0.03)
Level bias	-0.07 (0.02)	-0.12 (0.02)	-0.17 (0.03)	-0.11 (0.02)	-0.11 (0.02)	-0.14 (0.03)
Dispersion bias	0.42 (0.03)	0.42 (0.03)	0.82 (0.08)	0.39 (0.03)	0.40 (0.03)	0.56 (0.07)
When only records of the predictor traits were utilized[e]						
Validation accuracy	0.46 (0.03)	0.45 (0.03)	0.45 (0.03)	0.42 (0.03)	0.43 (0.03)	0.40 (0.03)
Level bias	-0.18 (0.02)	-0.21 (0.02)	-0.16 (0.03)	-0.19 (0.02)	-0.19 (0.02)	-0.14 (0.03)
Dispersion bias	0.48 (0.03)	0.51 (0.03)	0.71 (0.05)	0.47 (0.03)	0.47 (0.03)	0.57 (0.04)
When records of both the target and the predictor traits were utilized						
Validation accuracy	0.49 (0.02)	0.47 (0.03)	0.45 (0.03)	0.44 (0.03)	0.44 (0.03)	0.40 (0.03)
Level bias	-0.06 (0.02)	-0.11 (0.02)	-0.17 (0.03)	-0.09 (0.02)	-0.09 (0.02)	-0.15 (0.03)
Dispersion bias	0.45 (0.03)	0.47 (0.03)	0.69 (0.04)	0.44 (0.03)	0.44 (0.03)	0.55 (0.04)

[a] The subsequent ssGBLUP evaluation utilized the entire available data until the validation generation. [b] Validation generation preselection scenario, in which we discarded all animals in the validation generation with no progeny in the data. [c] Multi-generation preselection scenario, in which we discarded all animals in the validation or training generations with no progeny in the data. [d] The scarcely-recorded target traits are feed intake from the start (FISE) or the middle (FIME) to the end of performance testing. [e] The predictor traits are average daily gain during performance testing or throughout life, back fat thickness, and loin depth.

5.3.3 Accuracy and bias

Accuracy and bias of subsequent ssGBLUP evaluation of FISE are presented in Table 5.4. We included estimated heritabilities in Table 5.4 because they help in explaining the results of accuracy and bias. The estimated heritability had a tendency (i.e. an inclination that may not be statistically significant) to decrease with more preselection, whether records on animals in the validation generation were included or excluded in estimating the heritabilities.

5.3.3.1 Subsequent ssGBLUP evaluation only utilizing records of the target trait

Validation accuracy did not differ across reference and VGP scenarios, but decreased in the MGP scenario (Table 5.4). We observed this whether records on animals in the validation generation were included or excluded in the subsequent evaluation. Level bias was present, and increased with more preselection when records on animals in the validation were included in the subsequent evaluation. However, when records on animals in the validation generation were excluded from the subsequent evaluation, level bias did not differ across GPS scenarios. Level bias was however only marginal in all cases, as its biggest estimate across all GPS scenarios was only 0.17 additive genetic SD units. Dispersion bias (inflation in this case) was present in reference and VGP scenarios, and did not differ between the two scenarios, whether records on animals in the validation generation were included or excluded in the subsequent evaluation. Deflation was observed (i.e. the regression coefficient was bigger than the expected value of 0.5) with MGP when records on animals in the validation generation were included in the subsequent evaluation. When we repeated the subsequent evaluation for this MGP scenario using the (co)variance components estimated under the reference scenario (results not shown), the deflation disappeared. For the MGP scenario when records on animals in the validation generation were excluded from the subsequent evaluation, there was no dispersion bias, although a tendency towards deflation was observed.

5.3.3.2 Subsequent ssGBLUP evaluation only utilizing records of the predictor traits

Validation accuracy did not differ across GPS scenarios (Table 5.4). Level bias was always present, but did not differ across GPS scenarios. Similar to the subsequent ssGBLUP evaluation based on records from the target traits, level bias was always only marginal, with the highest estimate being -0.21 additive genetic SD units. There was no dispersion bias in most instances, whether records of the animals in the validation generation were included or excluded in the subsequent evaluation. The only exception is the MGP scenario with records on animals in the validation generation included in the subsequent evaluation, where there was deflation.

Similar to the subsequent ssGBLUP evaluation based on records from the target traits, the deflation disappeared when we repeated the subsequent ssGBLUP evaluation for this MGP scenario using the (co)variance components estimated under the reference scenario.

5.3.3.3 Subsequent ssGBLUP evaluation utilizing records of all traits

Similar to the subsequent ssGBLUP evaluation based on records from the predictor traits only, validation accuracy did not differ across GPS scenarios (Table 5.4). Similar to the subsequent ssGBLUP evaluation based on records from the target traits only, level bias was present, and increased with more GPS. However here too level bias was always only marginal, with the highest estimate being -0.17 additive genetic SD units. Dispersion bias was absent in most scenarios, whether records of the animals in the validation generation were included or excluded in the subsequent evaluation. The only exception is the MGP scenario with records on animals in the validation generation included in the subsequent evaluation, where there was deflation. Just like in the previous subsections, the deflation disappeared when we repeated the subsequent ssGBLUP evaluation for this MGP scenario using the (co)variance components estimated under the reference scenario.

5.4 Discussion

We studied the impact of GPS on accuracy and bias in subsequent ssGBLUP evaluation of preselected animals, for scarcely-recorded traits. We used data from a commercial pig breeding program in which routine preselection was already implemented, and retrospectively implemented additional layers of GPS by excluding animals with no progeny in the complete dataset from the subsequent evaluation. The data was on production traits in a sire line, with feed intake as scarcely-recorded target trait, and average daily gain, backfat thickness, and loin depth as widely-recorded predictor traits. In the subsequent evaluation, we assumed that records were available for only the scarcely-recorded target trait, only the predictor traits, or both. We performed the subsequent ssGBLUP evaluation either including or excluding records on animals in the validation generation, and in all cases without progeny of validation animals. We observed that validation accuracy generally only tended to decrease with more GPS. We also observed that although level and dispersion biases were sometimes present, the former was generally only marginal, and the latter did not differ across GPS scenarios.

In Jibrila et al. [12] we explained that our reference GPS scenario already contains the routine preselection implemented in commercial animal breeding programs. We also explained that the VGP and MGP scenarios implemented in our study do not happen in reality, and are used here to investigate the ability of ssGBLUP to estimate GEBV of preselected animals without preselection bias in subsequent evaluation in real breeding programs. The idea is that if ssGBLUP in subsequent evaluation of the scenarios with additional GPS is able to estimate GEBV of preselected animals as unbiased as in the subsequent evaluation of our reference scenario, then it is also able to estimate GEBV of preselected animals without preselection bias in subsequent evaluation in real breeding programs. Our results have shown that just like for widely-recorded traits [12], ssGBLUP in subsequent evaluation of animals for scarcely-recorded is able to estimate GEBV of preselected animals without preselection bias. Nevertheless, bias was observed in some cases. The increase in dispersion bias we sometimes saw moving from reference and VGP scenarios to MGP scenario was due to biased (co)variance components, most likely resulting from the small amount of data which came from heavily preselected animals in our MGP scenario. Note that there were only about 250 animals with records of the scarcely-recorded target traits (Table 5.1). However, when we repeated the subsequent ssGBLUP evaluation using the (co)variance components of the reference scenario, the dispersion bias in the MGP scenarios disappeared. In summary, ssGBLUP in subsequent evaluation of animals for scarcely-recorded is able to estimate GEBV of preselected animals without preselection bias, provided that the (co)variance components used are unbiased.

For level bias, the consistently small negative level bias observed throughout Table 5.4 most likely can be attributed to biased progeny yield deviation (PYD). In this study, we computed PYD using a PBLUP model. The animals that were recorded on feed intake were intensely preselected based on an index mainly including traits genetically correlated to feed intake. As a result, the EBV of the preselected animals for feed intake are expected to be underestimated even in multi-trait PBLUP evaluation like the one we performed to estimate PYD. Because level bias was computed as the weighted mean difference between PYD and half of the GEBV across all validation animals, a consistent underestimation of the PYD would explain the consistent (small) negative level bias. Finally, the accuracy loss we observed in this study with more preselection, though statistically significant, is as expected, as with more preselection there are less relatives with records [10,29].

5.4.1 Impact of preselection tended to be larger when utilizing records of the target trait

Although results were in all cases not statistically different between corresponding reference and VGP scenarios, tendencies of accuracy to decrease and of dispersion bias to increase from reference to VGP to MGP scenarios were bigger when records of FISE were utilized compared to when only records of the predictor traits were utilized. Since records of the target traits are by default scarce, any further reduction in the amount of these records is likely to cause a bigger impact than a corresponding reduction in records of predictor traits. With fewer records that are also more intensely preselected, validation accuracy reduces, and there is more shrinkage to the mean, making the GEBV less variable, thereby leading to larger dispersion bias. This underlines the importance of obtaining sufficient numbers of records for scarcely recorded traits.

5.4.2 Impact of predictor traits

Accuracy of predicting FISE tended to be higher when only records on predictor traits were utilized than when only records on FISE itself were utilized. This shows that most of the information provided by the relatively few records of FISE and FIME had already been provided by the relatively more abundant records of the predictor traits. Looking at Table 5.1, the predictor traits had 3 to 15 times more records than the two feed intake traits combined, depending on the preselection scenario. Philipsson et al., Pszczola et al., and Manzanilla-Pech et al., [2,3,5] all reported higher prediction accuracies for scarcely-recorded traits when only records of predictor traits were used compared to when only records of the scarcely-recorded traits themselves were used. It has been shown that for a predictor trait to increase prediction accuracy of a scarcely-recorded trait, the two traits have to be moderately to highly genetically correlated (i.e. ≥ 0.3, e.g. [1,2,4]), and the two traits have to be more genetically than phenotypically correlated (e.g. [7,30]). Looking at Table 5.3, all our predictor traits had reasonably high correlations with FISE, and the genetic correlations are higher than their corresponding phenotypic correlations. It has also been shown that the contribution of a predictor trait to increase in prediction accuracy of a scarcely-recorded trait increases with increase in difference in heritabilities of the two traits, with the predictor trait having the higher heritability (e.g. [3,4,6,7]). All our predictor traits had moderate to high heritabilities, and backfat thickness and loin depth had higher heritabilities than FISE (Table 5.3). As long as genetic and phenotypic correlations among traits are reliably estimated [15], accuracy of prediction of a scarcely-recorded trait increases with more predictor traits included in a multi-trait evaluation (e.g.[2,3]). All the factors discussed in this

paragraph likely contributed to some extent to the increase in accuracy of predicting FISE moving from only utilizing records of the scarcely-recorded FISE and FIME to only utilizing records of the predictor traits.

In this study we had predictor traits that all had moderate to high heritabilities and moderate to high genetic correlations with the target trait. In situations where heritabilities of the predictor traits and genetic correlations between predictor and target traits are smaller, we expect smaller prediction accuracies. However, even in such situations, we expect ssGBLUP in subsequent evaluation of animals for the scarcely-recorded target traits to be able to estimate GEBV of preselected animals without preselection bias. This is because we expect that ssGBLUP estimates unbiased GEBV of preselected animals mainly because it is able to estimate the Mendelian sampling terms of preselected animals from genotypes of the preselected animals and their parents, which are not affected by heritability and genetic correlations among traits.

5.5 Conclusions

As long as the (co)variance components are unbiased, ssGBLUP in subsequent evaluation of animals for scarcely-recorded traits is able to estimate GEBV of preselected animals without preselection bias. We observed this whether records on animals in the validation generation were included or excluded in the subsequent evaluation, and whether the subsequent evaluations were done utilizing records only from the scarcely-recorded target traits, only from the predictor traits, or from both the target and the predictor traits. The presented approach with additional preselection implemented allows to evaluate the impact of preselection using real data from an ongoing breeding program. Existing preselection in the data may affect derived proxies for true breeding value, but this can be detected based on analysis of the full data.

5.6 Acknowledgements

The authors thank Marco Bink and Katrijn Peeters from Hendrix Genetics, John Henshall from Cobb Europe, and Chris Schrooten and Gerben de Jong from CRV, for their inputs towards the design of this study.

5.7 References

1. Calus MPL, Veerkamp RF. Accuracy of multi-trait genomic selection using different methods. Genet Sel Evol. 2011;43:26.

2. Pszczola M, Veerkamp RF, De Haas Y, Wall E, Strabel T, Calus MPL. Effect of predictor traits on accuracy of genomic breeding values for feed intake based on a limited cow reference population. Animal. 2013;7:1759–68.
3. Manzanilla-Pech CIV, Gordo D, Difford GF, Løvendahl P, Lassen J. Multitrait genomic prediction of methane emissions in danish holstein cattle. J Dairy Sci. 2020;103:9195–206.
4. Ducrocq V. Multiple trait prediction: principles and problems. Proc World Congr Genet Appl to Livest Prod. 1994. p. 455–62.
5. Philipsson J, Ral G, Berglund B. Somatic cell count as a selection criterion for mastitis resistance in dairy cattle. Livest Prod Sci. 1995;41:19–24.
6. Jia Y, Jannink J-L. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics. 2012;192:1513–22.
7. Thompson R, Meyer K. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest Prod Sci. 1986;15:299–313.
8. Henderson CR. Best linear unbiased estimation and prediction under a selection model. Biometris. 1975;31:423–47.
9. Pollak EJ, van der Werf J, Quaas RL. Selection bias and multiple trait evaluation. J Dairy Sci. 1984;67:1590-1595.
10. Patry C, Ducrocq V. Evidence of biases in genetic evaluation due to genomic preselection in dairy cattle. J Dairy Sci. 2011;94:1011–20.
11. Sullivan PG. Mendelian sampling variance tests with genomic preselection. INTERBULL Bull. 2018;53.
12. Jibrila I, Vandenplas J, ten Napel J, Bergsma R, Veerkamp RF, Calus MPL. Impact of genomic preselection on subsequent genetic evaluation with ssgblup - using real data from pigs. https://doi.org/10.1101/2021.06.18.449002.
13. Dekkers JCM, Mathur PK, Knol EF. Genetic improvement of the pig. In: Rothschild MF, Ruvinsky A, editors. The genetics of the pig. 2nd ed. 2011. p. 390–425.
14. Westell RA, Quaas RL, Van Vleck LD. Genetic groups in an animal model. J Dairy Sci. 1988;71:1310–8.
15. Mrode RA. Linear models for the prediction of animal breeding values. 3rd ed. 2014.
16. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
17. Gilmour AR, Gogel BJ, Cullis BR, Thompson R. ASReml user guide release 3.0. VSN Int. Ltd. 2009. p. 398. https://asreml.kb.vsni.co.uk/wp-content/uploads/sites/3/2018/02/ASReml-3-User-Guide.pdf
18. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of holstein final score. J Dairy Sci. 2010;93:743–52.
19. Christensen OF, Lund MS. Genomic prediction when some animals are not genotyped. Genet Sel Evol. 2010;42:2.
20. Tsuruta S, Lourenco DAL, Masuda Y, Misztal I, Lawlor TJ. Controlling bias in

genomic breeding values for young genotyped bulls. J Dairy Sci. 2019;102:9956–70.
21. Powell JE, Visscher PM, Goddard ME. Reconciling the analysis of ibd and ibs in complex trait studies. Nat Rev Genet. 2010;11:800–5.
22. Vitezica ZG, Aguilar I, Misztal I, Legarra A. Bias in genomic predictions for populations under selection. Genet Res (Camb). 2011;93:357–66.
23. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.
24. Hsu W-L, Garrick DJ, Fernando RL. The accuracy and bias of single-step genomic prediction for populations under selection. Genes|Genomes|Genetics. 2017;7:2685–94.
25. ten Napel J, Vandenplas J, Lidauer M, Stranden I, Taskinen M, Mäntysaari E, et al. MiXBLUP: a user-friendly software for large genetic evaluation systems. 2020. p. 62. https://www.mixblup.eu/download.html.
26. Cameron ND. Selection indices and prediction of genetic merit in animal breeding. 1997.
27. R Core Team. R: a language and environment for statistical computing. 2019.
28. Helske J. Diagis: diagnostic plot and multivariate summary statistics of weighted samples from importance sampling. r package version 0.1.5. 2020.
29. Jibrila I, ten Napel J, Vandenplas J, Veerkamp RF, Calus MPL. Investigating the impact of preselection on subsequent single-step genomic blup evaluation of preselected animals. Genet Sel Evol. 2020;52:42.
30. Schaeffer LR. Sire and cow evaluation under multiple trait models. J Dairy Sci. 1984;67:1567–80.

Chapter 6

General Discussion

6.0 Introduction

This thesis aimed at investigating the impact of preselection in genomic evaluations, with ssGBLUP as the standard genomic evaluation model. Throughout this thesis, preselection was performed based on certain criteria, before the selection candidates had records for most breeding goal traits. Then preselected animals were raised further and phenotyped for more breeding goal traits, after which a subsequent evaluation was done. We call this evaluation 'subsequent' because it comes after the genetic evaluation usually performed before preselection. The thesis had two research questions, and the first one was to what extent does preselection affect accuracy and unbiasedness of GEBV of preselected animals in subsequent genomic evaluations? Chapters 2, 4 and 5 investigated this question. Chapter 2 was based on simulated datasets, and Chapters 4 and 5 were based on datasets from a commercial pig breeding program. While Chapter 4 focused on widely-recorded traits, Chapter 5 focused on scarcely-recorded traits. Results from these three chapters show that in subsequent genetic evaluation, ssGBLUP is able to estimate GEBV of preselected animals without bias and only with minimal accuracy loss due to preselection.

The other research question was whether genotypes of preculled animals (i.e. animals removed from the breeding program at preselection stage) are needed in subsequent genomic evaluations of their preselected sibs. This question was the main research question in Chapter 3, using a simulated dataset, and was also investigated in Chapters 4 and 5, using real datasets. Results from these chapters show that genotypes of preculled animals are only needed in subsequent ssGBLUP evaluation of their preselected sibs if their parents were not genotyped. In this chapter (Chapter 6), I draw inferences on some related topics not directly covered in the research chapters of this thesis, and make some recommendations on genotyping and evaluation strategies for commercial animal breeding organizations.

6.1 How different genetic evaluation models account for preselection

In multi-stage selection settings, breeding values or phenotypes available early in lives of animals are used to preselect animals. These types of information based on which animals are preselected usually have some genetic correlations with, and serve as indicators for, traits on which animals are evaluated at subsequent selection stages (e.g. Henderson, 1975; Jensen & Mao, 1991; Árnason et al., 2012; Patry & Ducrocq, 2011a). This means that usually, preselected animals on-average have positive Mendelian sampling (MS) terms for subsequently recorded traits, depending

on accuracy of the preselection (e.g. Sullivan, 2018; Tyrisevä et al., 2018; Sullivan et al., 2019).

If subsequent PBLUP evaluation of preselected animals is done without the information utilized at preselection stage, EBV of preselected animals are estimated with preselection bias and accuracy loss (e.g. Henderson, 1975; Pollak et al. 1984; Patry & Ducrocq, 2011a). However, if subsequent PBLUP evaluation is done in a multi-trait manner, including all the information used at preselection stage in addition to records that become available after preselection, preselection bias and accuracy loss are minimized or even completely eliminated (e.g. Henderson, 1975; Pollak et al., 1984; Appel et al., 1998; Patry & Ducrocq, 2011b; Janhunen et al., 2014). PBLUP in multi-trait subsequent evaluation uses the above pieces of information to estimate more accurate and less biased MS terms of preselected animals for subsequently recorded traits, compared to in single-trait evaluation of subsequently-recorded traits.

There are instances in which single-trait subsequent PBLUP evaluation can be unbiased. This is related to the concept of ignorability of selection, i.e. situations where previous selection can be ignored without introducing bias (e.g. Henderson, 1975; Schaeffer, 1987; Im et al.,1989; Gianola et al., 1989). One of the situations in which (pre)selection is ignorable is when it is done using criteria that do not change the average MS terms of the (pre)selected animals for traits of interest in subsequent evaluation (e.g. Henderson, 1975; Schaeffer, 1987; Im et al.,1989; Gianola et al., 1989). This occurs when preselection is based on for example traits that are not genetically correlated with traits of interest in subsequent evaluation, or when preselection is random as shown in Chapter 2 of this thesis. In such situations, subsequent PBLUP evaluation can be unbiased even if the information utilized at preselection stage is not utilized in the subsequent evaluation. In such cases, the average MS terms of preselected animals and of preculled animals for traits of interest in subsequent evaluation are not expected to be different from zero by preselection. Because the preselected animals are a good representation of their sibs before preselection, single-trait PBLUP subsequent evaluation of the preselected animals for subsequently recorded traits, just like its multi-trait counterpart, has all the information it needs to be able to estimate EBV of preselected animals without preselection bias.

On the other hand, even when selection is not ignorable like the GPS implemented in this thesis, ssGBLUP and other genomic evaluation models in subsequent

evaluation can estimate GEBV of preselected animals without preselection bias. This applies regardless of whether the subsequent evaluation is multi-trait or single-trait. The ability of ssGBLUP and other genomic evaluation models to prevent preselection bias has to do with the fact that these models use genomic instead of pedigree relationships among animals. Within families, genomic relationships are usually more accurate than pedigree relationships, such that for example full sibs can have different genomic relationships among themselves. If preselection has been effective, preselected sibs within a family on average have higher true MS terms for the traits to be evaluated in subsequent evaluation than the entire family and of course than the preculled sibs. If preselection has been effective, I expect that within a family, preselected sibs on average have higher genomic relationships among themselves than among all their sibs (e.g. VanRaden, 2008; Hayes et al., 2009; Gondro et al., 2013). The different genomic relationships among preselected family members are reflected in the coefficients of the inverse of the left hand side (LHS) of the mixed model equation (MME). This reflection ensures that preselected animals get the positive MS terms they truly have, regardless of whether the subsequent genomic evaluation is done in single-trait or multi-trait manner. With multi-trait subsequent genomic evaluation, the estimation of MS terms of preselected animals only becomes more accurate and less biased. This is as opposed to subsequent PBLUP evaluation, where genetic relationships among for example preselected full sibs are the same, and therefore coefficients of the inverse of the LHS are the same for preselected full sibs. In subsequent PBLUP evaluation, the only way preselected animals can have appropriate MS terms is if the subsequent evaluation is done utilizing all the information used at preselection stage.

To illustrate the abilities of different BLUP models to estimate the on-average positive MS terms of preselected animals in subsequent evaluation and thereby minimize accuracy loss and bias due to preselection, I give an example with a full sib family from the data I used in Chapters 4 and 5 of this thesis. The full sib family is made up of a sire, a dam, and four full sibs. Table 6.1 shows the pedigree and precorrected phenotypes of these animals for two traits, named Trait 1 and Trait 2. Assume that i) animals are selected for increased levels of the two traits, ii) animals have records for, and are evaluated on Trait 1 earlier than Trait 2, iii) only those animals preselected on Trait 1 are recorded for Trait 2, and iv) two of the four full sibs are preselected based on Trait 1. Then animals 3 and 4 will be preselected, and animals 5 and 6 will be preculled, based on Trait 1. When Trait 2 is subsequently recorded on the two preselected full sibs, the subsequent evaluation of these two preselected full sibs can be done either in a single-trait manner or in a two-trait

manner together with Trait 1. For comparison, I also considered a scenario without preselection, where all four full sibs had records of both traits.

Table 6.1 Pedigree and precorrected phenotypes of an example full sib family[a]

Animal	Sire	Dam	Trait 1	Trait 2 in the presence of preselection	Trait 2 in the absence of preselection
1	0	0	0.24	99.54	99.54
2	0	0	-1.75	46.91	46.91
3	1	2	-0.21	73.82	73.82
4	1	2	0.15	76.77	76.77
5	1	2	-0.71	-	35.82
6	1	2	-1.83	-	21.77

[a] The genetic and residual variance-covariance matrices of the two traits are $\begin{bmatrix} 1.41 & 12.39 \\ 12.39 & 1048 \end{bmatrix}$ and $\begin{bmatrix} 1.03 & 7.41 \\ 7.41 & 2939 \end{bmatrix}$, respectively.

For different types of genetic evaluation for Trait 2, average MS terms of the preselected full sibs (i.e. animals 3 and 4 in Table 6.1) and the preculled full sibs (i.e. animals 5 and 6 in Table 6.1) are shown in Table 6.2. Note that the results in Table 6.2 are from evaluations I conducted only using the animals in Table 6.1. I also only used genotypes of these six animals in ssGBLUP evaluations in Table 6.2, and the genotypes were extracted from the genotype file I used in Chapters 4 and 5. The MS term of an animal is the difference between the breeding value of the animal and the average breeding value of its parents. It can be seen from Table 6.2 that moving from single-trait PBLUP to two-trait PBLUP, to single-trait ssGBLUP, to two-trait ssGBLUP, higher MS terms were estimated for the preselected full sibs. This illustrates that MS terms of preselected animals are closer to their true values with two-trait ssGBLUP than with single-trait ssGBLUP, than with two-trait PBLUP, than with single-trait PBLUP. In comparison to average MS terms of the preselected full sibs, it can be seen from Table 6.2 that the average MS terms of the preculled full sibs are lower moving from single-trait PBLUP to two-trait PBLUP, and from single-trait ssGBLUP to two-trait ssGBLUP. This further illustrates that with more information, genetic evaluation models are better able to differentiate between preselected and preculled animals, thereby better preventing bias and accuracy loss associated with preselection.

It is well known that the closer MS terms are to their true values, the more accurate and less biased EBV and GEBV are. This means that in subsequent evaluation of preselected animals, prediction accuracy increases and bias decreases moving from single-trait PBLUP to two-trait PBLUP, to single-trait ssGBLUP, to two-trait ssGBLUP. In bigger genetic evaluations like the ones that take place in real breeding programs, and the ones conducted in Chapters 2 to 5 of this thesis, usually every animal has

more relatives with records, and MS terms are estimated more accurately than in the small example in Tables 6.1 and 6.2. Therefore, the difference in average MS terms of preselected animals moving across different types of genetic evaluation is not likely to be as high as we observed in this small example. This is especially the case for the differences in MS terms from ssGBLUP evaluations with compared to without preselection. Without preselection, records of the preculled sibs were included in genetic evaluations. When included in ssGBLUP, records of the preculled animals make GEBV of the parents smaller, and GEBV of the preselected sibs larger than with preselection. Both smaller GEBV of parents and larger GEBV of the preselected full sibs make MS terms of the preselected full sibs bigger.

Table 6.2 Average Mendelian sampling (MS) terms of the example full sib family estimated in different genetic evaluations for Trait 2, expressed in additive genetic SD of Trait 2

Preselection (present = yes, absent = no)	Evaluation	Number of traits included in the evaluation	Average MS term of the preselected full sibs (animals 3 and 4)	Average MS term of the preculled full sibs (animals 5 and 6)
Yes	PBLUP	1	0.01	0[a]
No	PBLUP	1	0.07	-0.15
Yes	PBLUP	2	0.07	-0.06
No	PBLUP	2	0.15	-0.33
Yes	ssGBLUP	1	0.24	0.09[b]
No	ssGBLUP	1	1.32	-1.18
Yes	ssGBLUP	2	0.66	-0.39
No	ssGBLUP	2	1.38	-1.20

[a] Records of the preculled full sibs were not used, but the preculled full sibs were left in the pedigree file, so they got the average EBV of their parents, hence the deviation from the parental average is zero. [b] Records of the preculled full sibs were not used, but the preculled full sibs were left in the pedigree and genotypic files, so they got their individual GEBV from their genomic relationships with their phenotyped parents and preselected full sibs.

6.2 How heritability affects results and conclusions of this thesis

In Chapter 2, I used simulated datasets to investigate how heritability modifies impact of preselection on subsequent ssGBLUP evaluation of preselected animals. I found that across all the heritabilities I considered (0.5, 0.3 and 0.1), ssGBLUP was able to estimate GEBV of preselected animals without level or dispersion bias, and only with minimal accuracy loss. So I concluded that ssGBLUP in subsequent evaluation is able to estimate GEBV of preselected animals without bias, regardless of the heritability of the trait in question. In Chapters 4 and 5, I used real data on pig production traits, with heritabilities ranging from 0.24 to 0.58, to investigate the impact of genomic preselection on accuracy and bias in subsequent ssGBLUP evaluation of preselected animals. For all traits, I only observed marginal accuracy loss and marginal to absent level and dispersion biases which generally did not increase with more intense preselection. I did not see different trends for traits with

different heritabilities, and this reaffirms the earlier conclusion using simulated datasets that ssGBLUP in subsequent evaluations is able to estimate GEBV of preselected animals without preselection bias, regardless of the heritability of the trait in question.

There are however traits with much lower heritabilities than I considered in this thesis. For example, reproduction traits are generally known to have heritabilities below 0.1 (e.g. Berry et al., 2014; Wolc et al., 2009; Ye et al., 2018). Although this thesis did not directly cover traits with heritability below 0.1, I expect genomic models in subsequent evaluation of animals for these traits to still be able to estimate GEBV of preselected animals with minimal accuracy loss and bias, just the way they do for traits with higher heritabilities. This is because, as established in Subsection 6.1, genomic models account for preselection by using genomic relationships among animals to accurately estimate the MS terms of preselected animals. Genomic relationships among animals are not affected by heritability.

6.3 How intensity of preselection affects results and conclusions of this thesis

In Chapters 2, 4, and 5, I studied the impact of preselection intensity on the ability of ssGBLUP in subsequent evaluation to estimate GEBV of preselected animals without bias and only with minimal accuracy loss. In these chapters, I implemented intensities of preselection up to and beyond what is obtainable in practice. For example, preselection intensities in real animal breeding programs are around what I implemented as high preselection in Chapter 2 (i.e. 10% of male and 15% of female selection candidates preselected). The very high preselection scenario I implemented in Chapter 2 (i.e. 5% of male and 12.5% of female selection candidates preselected) was just to investigate what happens when preselection intensities are higher than what is implemented in practice. In fact, the very high preselection scenario resulted in all the females required to produce the next generation being selected already at preselection stage, so there was no room to cull more females at the subsequent selection stage. Similarly, in Chapters 4 and 5, the reference scenario already contained the routine preselection implemented in reality, and the validation generation and multi-generation preselection scenarios only implemented additional layers of preselection in single and in multiple generations, respectively. In all the above situations, I observed that ssGBLUP in subsequent evaluations was able to estimate GEBV of preselected animals with minimal accuracy loss and with no bias compared to situations without preselection.

As preselection intensity increases, number of sibs with records reduces for most preselected animals. This in principle means that there is less information for estimating GEBV, and therefore there is more shrinkage to the mean and GEBV of preselected animals would be less accurate, underestimated and inflated (e.g. Mäntysaari et al., 2010; Patry & Ducrocq, 2011a; Mäntysaari & Koivula, 2012). Although in this thesis I observed that accuracy decreased as preselection intensity increased, I generally did not observe consistent trends in level and dispersion biases as preselection intensity increased. This likely has to do with what I explained in subsection 6.1, that ssGBLUP corrects for preselection by using genomic information to correctly estimate the MS terms of preselected animals. Furthermore, in Chapter 2 I showed that although subsequent selection accuracy decreased from scenarios without preselection to scenarios with preselection, genetic gain did not significantly decrease in scenarios with GPS – the most important type of preselection. This has to do with the fact that the accuracy that influences genetic gain is the overall selection accuracy and not only subsequent/final selection accuracy. And with GPS, I showed (in Chapter 2) that accuracies at both preselection and subsequent evaluation stages were reasonably high (e.g. with heritability of 0.1, GPS accuracy was 0.71, and subsequent selection accuracy was 0.48), meaning that overall selection accuracy with GPS is comparable to selection accuracy without preselection. In conclusion, I expect that for GPS to cause significant bias and significantly lower genetic gain in subsequent single-step genomic evaluation compared to a situation without preselection, the preselection intensity needs to be much higher than what is currently implemented in real animal breeding programs.

In principle, higher preselection intensities result in less available information to be used to estimate (co)variance components in the future, and as reference data in genomic evaluation of future generations. In Chapters 4 and 5, the multi-generation preselection (MGP) scenario represented a situation in which in all generations only animals with offspring were kept in the data. In Chapter 4 where widely-recorded traits were considered, (co)variance components were generally estimated in the MGP scenario as accurately and unbiasedly as they were estimated in the reference scenario where all the available data were utilized. Accuracy and bias were also generally not worse in the MGP scenario compared to in the reference scenario, especially when selection candidates had records. In Chapter 5 where a scarcely-recorded trait was considered, (co)variance components were generally estimated less accurately and with bias in the MGP scenario compared to in the reference scenario. Accuracy was also lower in the MGP scenario than in the reference scenario, especially when only records of the scarcely-recorded trait were utilized in

the subsequent evaluation. However, I explained in Chapter 5 that such poor estimates of (co)variance components and GEBV in the MGP scenario were mainly because they came from a small amount of data on heavily preselected animals (less than 500 animals had records for the scarcely-recorded trait), and this usually does not happen in reality.

Schaeffer (2018) suggested that preselecting ≤ 10% of animals based on a correlated trait will lead to substantial bias in a second correlated trait even if the subsequent evaluation is done in a two-trait manner. Although Schaeffer (2018) was referring to subsequent PBLUP evaluations, the suggestion may be applicable to subsequent ssGBLUP evaluations as well. I did not study preselection on a correlated trait in the research chapters of this thesis. Instead, I mainly studied preselection based on early GEBV of selection candidates for traits analyzed in subsequent evaluations, i.e. GPS. Since with GPS both preselection and subsequent evaluation are on the same trait(s), and the main information used in preselection (i.e. genotypes of selection candidates) are also utilized in subsequent ssGBLUP evaluation, preselection bias is not expected even when ≤ 10% of the selection candidates were preselected. This was shown in Chapters 2 and 3 of this thesis. However, in the small example in Subsection 6.1, preselection was based on a correlated trait - Trait 1, though preselection intensity was only 50%. It can be seen in Table 6.2 that in subsequent evaluation of Trait 2, single-trait ssGBLUP did better than two-trait PBLUP, and two-trait ssGBLUP did better than single-trait ssGBLUP in estimating positive MS terms for preselected animals, and by implication reducing preselection bias. I also explained why in large genetic evaluations, differences in the sizes of MS terms for preselected animals estimated by the different models may not be as pronounced as they are in the small example in Tables 6.1 and 6.2. Altogether, in large genetic evaluations like the ones conducted in commercial animal breeding programs, I expect ssGBLUP in subsequent evaluation of animals preselected on a correlated trait to prevent most of the preselection bias.

6.4 Recommendations to animal breeding programs

While discussing the findings of my thesis with fellow animal breeders, I was often asked the following questions: 1. Since you have shown that information from preculled animals is not needed in subsequent ssGBLUP evaluation of their preselected sibs, should poor animals still be genotyped? 2. If the answer to question 1 is no, what is the implication of only genotyping top animals on the quality of the genomic reference population? These are indeed interesting questions, and answers

to them constitute the main recommendation to animal breeding programs from the results of this thesis.

To answer the first question, I give a bit of background as follows. Previous studies have shown benefits of genotyping poor animals in genomic evaluation models like SNPBLUP and GBLUP (e.g. Boligon et al., 2012; Jiménez-Montero et al., 2012). However, in the single-step era, Howard et al., (2018) has shown that selectively genotyping a percentage of top selection candidates does not result in lower genetic gain or more bias compared to randomly genotyping the same percentage of selection candidates. Chapters 2 to 5 of this thesis have also shown that genotyping only the top animals does not result in significant accuracy loss and bias. The facts that single-step genomic models utilize phenotypes of both poor/ungenotyped and top/genotyped animals, and that the models have been optimized to take care of selective genotyping (e.g. Vitezica, et al., 2011; Hsu et al., 2017) explain why these models are less affected by selective genotyping compared to genomic models like GBLUP and SNPBLUP. To come back to answering the first question, I will say animals that have the highest chance to make it to subsequent evaluation should be genotyped. After that, if there is still some room to genotype more animals, then other animals can be genotyped. This is because selection candidates that survive to subsequent evaluation are the ones to be evaluated during the subsequent evaluation, and the most relevant genotypes and phenotypes in genetic evaluation of animals are the genotypes and phenotypes of the animals to be evaluated.

As for the second question, I expect only genotyping top animals to have no significant effect on the quality of the genomic reference population. This is because the animals that contribute the most to the quality of the reference population are those animals with both genotypes and phenotypes. Even if poor animals are genotyped, they usually end up not being phenotyped, because they are usually preculled. Additionally, the argument offered by those who think poor animals should be genotyped is for poor alleles to be represented in the reference population. Poor alleles can be represented in the reference population even without having poor animals therein. This is because most breeding goal traits are quantitative, and governed by large numbers of genes with small effects. This means that top animals also carry some poor alleles, only that they have more favorable combinations of alleles than poor animals. In Chapters 4 and 5 of this thesis, I implemented a GPS scenario in which only top animals, i.e. animals with progeny in the data, were included in subsequent evaluation of preselected animals (i.e. the multi-generation preselection scenario). Generally, I did not observe a significant

decrease in accuracy or (increase in) bias compared to in the reference scenario, where genotypes and phenotypes of culled animals were also included. In Chapters 3 and 4, I showed that genotypes of preculled animals, which are poor animals, are only needed in the subsequent ssGBLUP evaluation of their preselected sibs if their parents are not genotyped. The contents of the last two sentences are indications that top animals also possess poor alleles, as excluding poor animals from ssGBLUP evaluations did not cause a significant decrease in quality of the evaluations.

The final and perhaps the most obvious recommendation for animal breeding programs from the results of this thesis is for the breeding programs to continue or start to use single-step genomic evaluation models in evaluating their animals. This thesis has shown single-step models to be very robust in handling incomplete/preselected data, in addition to the higher accuracy and lower bias from single-step models compared to from other genetic evaluation models, as widely reported in the literature (e.g. Vitezica et al., 2011; Misztal et al., 2013; Legarra et al., 2014).

6.5 Concluding remarks and recommendation for future research

In this thesis, I have shown that single-step genomic evaluation models are able to estimate GEBV of preselected animals without preselection bias, using both simulated and real datasets. Nonetheless, there are reports that some bias is observed in single-step genomic evaluation of preselected animals in commercial breeding programs. In Chapters 4 and 5 of this thesis, where I used data from a commercial pig breeding program, I also observed bias in some instances, only that the bias did not get worse with more intense preselection. So, I concluded that the bias must have come from something else. In Chapter 4, I noted that there are a number of other sources of bias in single-step genomic evaluations, such as inaccurate pedigree and (co)variance components, and that the bias sometimes observed could have come from these other sources. I know that research has been conducted and is still being conducted on how to make single-step genomic evaluation models free from all kinds of bias (e.g. Misztal et al., 2020). However, we are still not there. So, I recommend that research in this direction should continue.

References

Appel, L. J., Strandberg, E., Danell, B., & Lundeheim, N. (1998). Adjusting for Missing Data Due to Culling before Testing in Genetic Evaluations of Swine. *Journal of Animal Science*, 76(7), 1794–1802. https://doi.org/10.2527/1998.7671794x

Árnason, T., Albertsdóttir, E., Fikse, W. F., Eriksson, S., & Sigurdsson, Á. (2012).

Estimation of genetic parameters and response to selection for a continuous trait subject to culling before testing. *Journal of Animal Breeding and Genetics*, 129(1), 50–59. https://doi.org/10.1111/j.1439-0388.2011.00941.x

Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal*, 8(SUPPL. 1), 105–121. https://doi.org/10.1017/S1751731114000743

Boligon, A. A., Long, N., Albuquerque, L. G., Weigel, K. A., Gianola, D., & osa, G. J. M. (2012). Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection. *Journal of Animal Science*, 90(13), 4716–4722. https://doi.org/10.2527/jas.2012-4857

Gianola, D., Fernando, R. L., & Foulley, J. L. (1989). Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome*, 31, 768–777.

Gondro, C., van der Werf, J., & Hayes, B. (Eds.). (2013). *Genome-Wide Association Studies and Genomic Prediction*. Springer. https://doi.org/10.1007/978-1-62703-447-0

Hayes, B. J., Visscher, P. M., & Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research*, 91(1), 47–60. https://doi.org/10.1017/S0016672308009981

Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometris*, 31(2), 423–447. https://doi.org/10.2307/2529430

Howard, J. T., Rathje, T. A., Bruns, C. E., Wilson-Wells, D. F., Kachman, S. D., & Spangler, M. L. (2018). The impact of selective genotyping on the response to selection using single-step genomic best linear unbiased prediction. *Journal of Animal Science*, 96(11), 4532–4542. https://doi.org/10.1093/jas/sky330

Hsu, W.-L., Garrick, D. J., & Fernando, R. L. (2017). The Accuracy and Bias of Single-Step Genomic Prediction for Populations Under Selection. *Genes|Genomes|Genetics*, 7(August), 2685–2694. https://doi.org/10.1534/g3.117.043596

Im, S., Fernando, R., & Gianola, D. (1989). Likelihood inferences in animal breeding under selection: a missing-data theory view point. *Genetics Selection Evolution*, 21(4), 399–414. https://doi.org/10.1186/1297-9686-21-4-399

Janhunen, M., Kause, A., Vehviläinen, H., Nousiainen, A., & Koskinen, H. (2014). Correcting within-family pre-selection in genetic evaluation of growth-A simulation study on rainbow trout. *Aquaculture*, 434, 220–226. https://doi.org/10.1016/j.aquaculture.2014.08.020

Jensen, J., & Mao, I. L. (1991). Estimation of Genetic Parameters Using Sampled Data from Populations Undergoing Selection. *Journal of Dairy Science*, 74(10), 3544–3551. https://doi.org/10.3168/jds.S0022-0302(91)78546-9

Jiménez-Montero, J. A., González-Recio, O., & Alenda, R. (2012). Genotyping strategies for genomic selection in small dairy cattle populations. *Animal*, 6(8), 1216–1224. https://doi.org/10.1017/S1751731112000341

Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, 166(1), 54–65.

https://doi.org/10.1016/j.livsci.2014.04.029

Mäntysaari, E A, & Koivula, M. (2012). GEBV validation test revisited. *Interbull Bulletin*, 45.

Mäntysaari, Esa A, Liu, Z., & VanRaden, P. (2010). Interbull validation test for genomic evaluations. *Interbull Bulletin*, 41. https://journal.interbull.org/index.php/ib/article/viewFile/1134/1125

Misztal, I., Aggrey, S. E., & Muir, W. M. (2013). Experiences with a single-step genome evaluation. *Poultry Science*, 92(9), 2530–2534. https://doi.org/10.3382/ps.2012-02739

Misztal, I., Lourenco, D., & Legarra, A. (2020). Current status of genomic evaluation. *Journal of Animal Science*, (January), 1–14. https://doi.org/10.1093/jas/skaa101

Patry, C., & Ducrocq, V. (2011a). Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. *Journal of Dairy Science*, 94(2), 1011–1020. https://doi.org/10.3168/jds.2010-3804

Patry, Clotilde, & Ducrocq, V. (2011b). Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. *Genetics Selection Evolution*, 43(30). https://doi.org/10.1186/1297-9686-43-30

Pollak, E. J., van der Werf, J., & Quaas, R. L. (1984a). Selection bias and multiple trait evaluation. *Journal of Dairy Science*, 67, 1590–1595. https://doi.org/10.3168/jds.S0022-0302(84)81481-2

Pollak, E. J., van der Werf, J., & Quaas, R. L. (1984b). Selection Bias and Multiple Trait Evaluation. *Journal of Dairy Science*. https://doi.org/10.3168/jds.S0022-0302(84)81481-2

Schaeffer, L. R. (1987). Estimation of Variance Components Under a Selection Model. *Journal of Dairy Science*, 70(3), 661–671. https://doi.org/10.3168/jds.S0022-0302(87)80056-5

Schaeffer, L. R. (2018). Necessary changes to improve animal models. *Journal of Animal Breeding and Genetics*, 135(February), 124–131. https://doi.org/10.1111/jbg.12321

Sullivan, P. G. (2018). Mendelian Sampling variance tests with genomic preselection. *INTERBULL BULLETIN*, 53.

Sullivan, P. G., Mäntysaari, E. A., Dejong, G., & Benhajali, H. (2019). Modifying MACE to accommodate genomic preselection effects. *Interbull Bulletin*, 55.

Tyrisevä, A.-M., Mäntysaari, E. A., Jakobsen, J., Aamand, G. P., Dürr, J., Fikse, W. F., & Lidauer, M. H. (2018). Detection of evaluation bias caused by genomic preselection. *Journal of Dairy Science*, 101(4), 3155–3163. https://doi.org/10.3168/jds.2017-13527

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Vitezica, Z. G., Aguilar, I., Misztal, I., & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, 93(5), 357–366. https://doi.org/10.1017/S001667231100022X

Wolc, A., White, I. M., Olori, V. E., & Hill, W. G. (2009). Inheritance of fertility in broiler

chickens. *Genetics Selection Evolution*, 41(47). https://doi.org/10.1186/1297-9686-41-47

Ye, J., Tan, C., Hu, X., Wang, A., & Wu, Z. (2018). Genetic parameters for reproductive traits at different parities in Large White Pigs. *Journal of Animal Science*, 96(4), 1215–1220. https://doi.org/10.1093/jas/sky066

Summary

Summary

The development of genomic evaluation models over the last two decades has resulted in more accurate estimation of breeding values of animals, compared to when only pedigree-based genetic evaluation models were used. In large animal breeding programs, selection of parents of the next generation usually takes place in multiple stages. The initial stages of selecting parents of the next generation are collectively called preselection. Preselection takes place when selection candidates are young, sometimes even before they have records for any breeding goal trait. As the selection candidates grow older, they generally get records for more breeding goal traits, and they are re-evaluated in subsequent evaluations to select the final set of parents of the next generation. Impact of preselection on accuracy and bias of subsequent pedigree-based evaluation is known, but this is not the case for subsequent genomic evaluation. Role of genotypes from preculled animals, i.e. animals removed from the breeding program at preselection stage, in subsequent genomic evaluation of their preselected sibs is also poorly understood. In this thesis, I investigated the impact of preselection on accuracy and bias in subsequent genomic evaluation of preselected animals, using single-step genomic best linear unbiased prediction (ssGBLUP) as the representative genomic evaluation model.

In **Chapter 2**, I used simulated datasets to investigate, for different heritabilities, the impact of types and intensities of preselection on accuracy and bias in ssGBLUP evaluation of preselected animals. A trait was simulated with heritabilities of 0.1, 0.3, and 0.5, and the types of preselection implemented were genomic, parental average, and random preselection. I implemented three intensities of preselection, ranging from no preselection to preselecting 5% of male and 12.5% of female selection candidates. Subsequent ssGBLUP evaluation of preselected animals was always performed excluding genotypes of preculled animals. I showed that preselection, regardless of its type and intensity, and heritability, results in accuracy loss in subsequent ssGBLUP evaluation of preselected animals, compared to a scenario without preselection. I also explained that the accuracy loss is mainly due to loss of relatives with records, and/or reduction in heritability. I also showed that ssGBLUP estimates genomic estimated breeding values (GEBV) of preselected animals without preselection bias, regardless of type and intensity of preselection, and heritability. The results of this chapter also showed that if ssGBLUP is used in subsequent evaluation of genomically preselected animals, realized genetic gain is only slightly lower compared to a scenario without preselection.

In **Chapter 3**, using part of the simulated data used in Chapter 2, I investigated the roles of genotypes and phenotypes from various groups of animals in preventing preselection bias in subsequent ssGBLUP evaluation of preselected animals. In other words, I established the minimum information required in subsequent ssGBLUP evaluation of preselected animals to estimate GEBV of genomically preselected animals without preselection bias. I showed that to prevent preselection bias it is sufficient to supply the model with i) data of the reference population used in the evaluation at preselection stage and ii) genotypes and phenotypes of the preselected animals. I also showed that genotypes of preculled animals are only needed in subsequent ssGBLUP evaluation of their genomically preselected sibs if some of their parents are not genotyped and included in the reference data.

Although in Chapter 2 I showed that ssGBLUP in subsequent evaluation estimates GEBV of preselected animals without preselection bias, there are unpublished reports that bias is observed in subsequent ssGBLUP evaluation of preselected animals in commercial breeding programs. So, in **Chapters 4 and 5**, I used datasets from a commercial pig breeding program to verify whether what I found using simulated datasets holds in reality as well. In **Chapter 4**, I investigated the impact of genomic preselection (GPS) on accuracy and bias in subsequent ssGBLUP evaluation of preselected animals, for widely-recorded traits – traits that are routinely measured on the majority of animals in a breeding population. The traits were average daily gain, backfat thickness, and loin depth. I used the full data provided by the commercial pig breeding program as control, and retrospectively implemented additional layers of GPS. After subsequent evaluation, I compared accuracy and bias of subsequent ssGBLUP evaluation after these additional layers of GPS with accuracy and bias of ssGBLUP evaluation of the data as I received it from the commercial breeding program. Results for all traits showed only marginal loss in accuracy due to the additional layers of GPS. Bias was largely absent, and when present did not increase with more intense preselection. These results show that even in real animal breeding programs, ssGBLUP in subsequent genetic evaluation estimates GEBV of preselected animals without preselection bias. I suggested that, since the bias that was sometimes observed in subsequent ssGBLUP evaluation did not increase with more intense preselection, it was most-likely caused by something else.

It is generally known that prediction accuracy is higher, and probability of bias is lower, in genetic evaluation of animals for widely-recorded traits than for scarcely-recorded traits - traits only measured on a small proportion of animals in each generation. To verify whether ssGBLUP in subsequent evaluation is able to estimate

GEBV of preselected animals without preselection bias for all categories of traits, in **Chapter 5** I repeated what I did in Chapter 4, but now using scarcely-recorded traits. The scarcely-recorded trait I used was feed intake, and it was also the target trait in this chapter. The widely-recorded traits I used in Chapter 4 were genetically correlated with feed intake, so they could be used as predictors of feed intake. So in Chapter 5, I performed the subsequent ssGBLUP evaluation of preselected animals using records of the scarcely-recorded target trait, records of widely-recorded predictor traits, or records of both the scarcely-recorded target trait and widely-recorded predictor traits. Just like in Chapter 4, only marginal loss in accuracy due to the additional layers of GPS was observed. Bias was also largely absent, and when present, did not increase with more intense preselection. The above results were observed whether records of the scarcely-recorded target trait, of the widely-recorded predictor traits, or of both the scarcely-recorded target trait and the widely-recorded predictor traits were used in the subsequent ssGBLUP evaluation. These results show that even for scarcely-recorded traits, ssGBLUP in subsequent genetic evaluation estimates GEBV of preselected animals without preselection bias.

Finally, in **Chapter 6**, I explained and illustrated the mechanism that enables ssGBLUP and other genomic models in subsequent evaluation of preselected animals to minimize accuracy loss and bias associated with preselection, even when the information used in the subsequent genomic evaluation does not include all the information used at preselection stage. This mechanism is the fact that ssGBLUP uses genomic information to estimate the on-average positive Mendelian sampling terms of preselected animals. I also made inferences on subjects not directly covered in thesis. Specifically, I discussed that even for traits with very low heritabilities such as reproduction traits, ssGBLUP in subsequent evaluation is expected to estimate GEBV of preselected animals without preselection bias. I also discussed why ssGBLUP in subsequent evaluation should be able to estimate GEBV of preselected animals without preselection bias even if the preselection intensity is higher than what is currently implemented in commercial animal breeding programs. Finally, I recommended that commercial animal breeding programs should genotype as many young selection candidates as economically possible, so that preselection can be as accurate as possible. This will in the end ensure that loss of genetic gain as a result of preselection is minimized.

Appendices

Curriculum vitae

About the author

Ibrahim Jibrila was born August 15, 1986 in Kasarawa, Sokoto, Nigeria. He grew up with cattle, sheep and chickens in his backyard, which were raised in small scale by his mother. Ibrahim studied General Agriculture with specialization in Animal Science at Usmanu Danfodiyo University, Sokoto from 2003 to 2008. His undergraduate research project was on 'Effect of tannin concentration and tanning duration on physical properties of leather produced from Sokoto Red goat skin'. He then performed his one-year mandatory Nigerian National Youth Service in Ibadan, Nigeria between 2009 and 2010, where he taught Agricultural Science to high school students. From January 2012 to January 2018, he was a research and teaching assistant first at Ahmadu Bello University, Zaria and later at Usmanu Danfodiyo University, Sokoto, both in Nigeria.

Between 2014 and 2017, Ibrahim studied MSc Animal Sciences with specialization in Animal Breeding and Genetics at Wageningen University. His MSc theses were on 'Predicting milk Phosphorus content using infrared and genomic data', and on 'Genome-wide association study for direct and indirect genetic effects on plumage condition in laying hens'. This MSc programme was sponsored by the Dutch Government through the Netherlands Fellowship Programme.

In February 2018, Ibrahim started as a PhD candidate at the Animal Breeding and Genomics group of Wageningen University & Research. He worked within the Breed4Food project 'Utilizing DNA information', subproject 'improve genomic prediction models', topic 'impact of preselection in genomic prediction'. The main outcomes of his PhD candidature are presented in this thesis. Since February 2022, Ibrahim works a researcher with the cattle improvement organisation 'Cooperatieve Rundvee Verbetering (CRV)', Arnhem, the Netherlands.

Publications

Journal articles

Ibrahim Jibrila, Jeremie Vandenplas, Jan ten Napel, Rob Bergsma, Roel F. Veerkamp, and Mario P.L. Calus, (2022). Impact of genomic preselection on subsequent genetic evaluations with ssGBLUP - using real data from pigs. Accepted for publication in Genetics Selection Evolution. https://doi.org/10.1101/2021.06.18.449002.

Jibrila I, Vandenplas J, ten Napel J, Veerkamp RF, Calus MPL, (2021). Avoiding preselection bias in subsequent single-step genomic BLUP evaluations of genomically preselected animals. Journal of Animal Breeding and Genetics, 138:432–441. https://doi.org/10.1111/jbg.12533.

Ibrahim Jibrila, Jan ten Napel, Jeremie Vandenplas, Roel F. Veerkamp and Mario P.L. Calus, (2020). Investigating the impact of preselection on subsequent single-step genomic BLUP evaluation of preselected animals. Genetics Selection Evolution, 52:42. https://doi.org/10.1186/s12711-020-00562-6.

Jibir M., Isa A. M., Adamu I., Garba S. and **Jibrila I.,** (2013). Value Perception of Shell Egg Characteristics as Determined by Size and Colour Journal of Animal Science Advances, 3(9): 457-461. ISSN: 2251-7219.

Jibir, M., **Jibrila I.**, Garba S., Isa A. M and Sonfada M. L., (2013). Fat deposition and distribution in autochthonous goats of the semi-arid North-Western Nigeria. Journal of Animal Science Advances, 3(9): 481-484. ISSN: 2251-7219.

Jibir M., Isa A.M., Garba S., **Jibrila I.** And Omojola A.B., (2013). Linear Body Measurements and Slaughter Characteristics of Meat Goats in the Semi-Arid Zone of North-Western Nigeria. Journal of Animal Science Advances, 3(6): 297-303. ISSN: 2251-7219.

M. Jibir, **I. Jibrila**, S. Garba, A. M. Isa A. M and A. B. Omojola, (2012). Carcass and Lean Quality Characterization of the Autochthonous Goats in the Semi-arid Zone of North- western Nigeria. International Journal of Meat Science, 2 (2): 34-39. ISSN 2071-7113.

Contributions to scientific conferences

I. Jibrila, J. Vandenplas, J. ten Napel, R. Bergsma, R.F. Veerkamp and M.P.L. Calus, (2021). Impact of preselection on subsequent ssGBLUP evaluations of pigs. Book of Abstracts of the 72nd Annual Meeting of the European Federation of Animal Science, 30th August to 3rd September 2021, Davos, Switzerland. P203.

Adamu Mani Isa, Yanyan Sun, **Ibrahim Jibrila**, Aliyu Sa'adu, and Jilan Chen, (2021). In-silico transcription factor enrichment analysis suggests that CCAAT/enhancer binding proteins in the follicles of hybrid hens may influence heterosis for clutch size. Proc. 46th Conf., Nig. Soc. for Anim. Prod. 16 - 19 March, 2021. Federal University Dutsin-Ma, Katsina State. P 708-711. ISSN: 1596-5570.

I. Jibrila, J. Vandenplas, J. Ten Napel, R.F. Veerkamp and M.P.L. Calus, (2020). What information does ssGBLUP need to give correct evaluations in the presence of preselection? Book of Abstracts of the 71st Annual Meeting of the European Federation of Animal Science, 1st – 4th December, 2020, Virtual. P582.

Ibrahim Jibrila, Jeremie Vandenplas, Jan ten Napel, Roel F. Veerkamp and Mario P. L. Calus, (2020). Impact of preselection varies across genetic evaluation models. Wageningen Institute of Animals Science Annual Conference, 13th – 14th February, 2020, Lunteren, the Netherlands. https://edepot.wur.nl/518508.

I. Jibrila, J. Vandenplas, J. ten Napel, R.F. Veerkamp and M.P.L. Calus, (2019). Impact of preselection on genetic evaluation of selection candidates using single step GBLUP**.** Book of Abstracts of the 70th Annual Meeting of the European Federation of Animal Science, 26th - 30th August, 2019, Ghent, Belgium. P449.

H. Bovenhuis**, I. Jibrila** and J. Dijkstra, (2018). Predicting milk Phosphorus content based on genotypic and milk infrared data. Proceedings of the World Congress on Genetics Applied to Livestock Production, 5th – 9th February, 2018, Auckland, New Zealand. Volume Electronic Poster Session - Biology - Feed Intake and Efficiency 11, 534.

H. Bovenhuis, **I. Jibrila** and J. Dijkstra, (2018). Improving Phosphorus efficiency using infrared predicted milk phosphorus content. Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science, 27-31 August 2018, Dubrovnik, Croatia. P286.

Jibir, M., Alli-Balogun, J. K., **Jibrila, I.**, Garba, S., Ribah, M. I. and Isa, A. M., (2014). Preference Studies on Rabbit Meat Products: A Consumer Approach. Proc. 39th Conf., Nig. Soc. for Anim. Prod. 16 - 19 March, 2014. Babcock Univ. Ilishan-Remo, Ogun State, Nigeria. P 502-504. ISSN: 1596-5570.

Popular publications

Ibrahim Jibrila, Jeremie Vandenplas, Jan ten Napel, Roel Veerkamp & Mario Calus, (2021). Not all genotypes are needed in final evaluations of preselected animals. https://www.wur.nl/en/newsarticle/not-all-genotypes-are-needed-in-final-evaluations-of-preselected-animals.htm.

Ibrahim Jibrila, Jan ten Napel, Jeremie Vandenplas, Roel Veerkamp & Mario Calus, (2020). New models evaluate animals correctly even with incomplete data. https://www.wur.nl/en/newsarticle/new-models-evaluate-animals-correctly-even-with-incomplete-data.htm.

Education & training during the PhD Program, under the auspices of the Graduate School Wageningen Institute of Animal Sciences (WIAS)

A. The Basic Package (3 ECTS*)	year
Attending the WIAS Introduction Day	2018
Attending the Wageningen Graduate Schools (WGS) course 'Scientific Integrity & Ethics in Animal Sciences'	2018
Attending the WIAS course on 'Essential Skills'	2018
B. Disciplinary Competences (15 ECTS)	year
Writing WIAS Research Proposal	2018
Attending the Nordic Forestry, Veterinary and Agricultural University Network (NOVA) course 'Linear Models in Animal Breeding'	2018
Attending the NOVA course 'Application of Genome Wide SNPs in Single Step Genomic Analysis'	2019
Participating in the study group of the book 'Introduction to Quantitative Genetics by Falconer & Mackay'	2018 to 2019
Attending the WIAS course on Genotype by Environment Interaction, Resilience and Uniformity	2020
Attending Quantitative Genetics Discussion Group meetings	2018 to 2021
C. Professional Competences (10.7 ECTS)	year
Attending the course 'WGS PhD Workshop Carousel'	2018
Organising Quantitative Genetics Discussion Group meetings	2018 to 2021
Attending the WUR Facilities & Services course 'Basic Linux for HPC'	2018
Attending the WUR Facilities & Services course 'HPC Basic'	2019
Attending the Wageningen in'to Languages Course 'Efficient Writing Strategies'	2019
Attending the Wageningen in'to Languages Course 'Presenting with Impact'	2019
Attending the WGS course 'Stress Identification and Management'	2019
Attending the WGS course 'Effective Behaviour in Professional Surroundings'	2019
Membership of WIAS Associated PhD Students' Council, and membership of the WIAS Education Committee	2018 to 2020
Attending the WIAS course 'Career Orientation'	2021
Attending the WIAS course 'The Final Touch'	2021

D. Societal Relevance (1.5 ECTS)	year
Attending the WIAS course on Societal Impact of Research	2019
E. Presentation Skills (4 ECTS)	year
Impact of genomic preselection on genomic evaluation of animals; WIAS Science Day, Lunteren, the Netherlands; Oral presentation	2019
Impact of preselection on genetic evaluations using single-step GBLUP; EAAP annual meeting, Ghent, Belgium; Oral presentation	2019
Impact of preselection varies across genetic evaluation models; WIAS Annual Conference, Lunteren, the Netherlands; Oral presentation	2020
What information does ssGBLUP need to give correct evaluations in the presence of preselection?; EAAP annual meeting, Online; Oral presentation	2020
Impact of preselection on subsequent ssGBLUP evaluations of pigs; EAAP annual meeting, Davos, Switzerland; Oral presentation	2021
F. Teaching competences (4 ECTS)	year
Assisting in teaching the course 'Animal Breeding and Genetics'	2018 & 2019
Reviewing WIAS Research Master Cluster proposals	2019
Supervising MSc Thesis	2020
Education and Training Total ECTS = 38.20	

**One ECTS credit equals a study load of approximately 28 hours*

Acknowledgements

My PhD journey was an embodiment of the popular African saying that “it takes a village to raise a child”. Many people contributed towards the success of this journey, and I cannot acknowledge everyone individually. I will however do my best to be as complete and concise as possible.

I will start with my supervisory team. **Mario**, thank you for having confidence in me throughout this journey. You encouraged me to do things I wouldn’t have done on my own, and I really appreciate that. I enjoyed and already miss our regular and impromptu discussions. I learnt a lot from them, and they developed my confidence tremendously. You always patiently listened to me until I landed, even when I struggled to make my points. **Roel**, thank you for always assessing my work from an outsider perspective, and for helping me to see the bigger picture. Thank you also for all the managerial assistance you offered to me! **Jan**, thank you for all the help with MiXBLUP. On a number of occasions you improved some options mainly because of me. I appreciate your detailed feedback on my work, including suggestions for improvement on my English grammar and on my writing style. **Jeremie**, you have been part of my supervisory team throughout the journey. You are not one of my (co-)promotors only because according to WUR policy, it is not possible to have more than three (co-)promotors from the same group. Thank you for all your feedback of my work, and for your help with codes.

Although I have acknowledged our Breed4Food industrial partners in all my papers and in the Colophon of this PhD project, I want to mention some highlights here. Thank you all for your inputs in designing my studies, and for your regular feedbacks on my work during our quarterly project meetings. This PhD project wouldn’t have been possible without your contributions. **Rob**, thank you for facilitating my use of data from **Topigs Norsvin** in the fourth and fifth chapters of this thesis. Thank you also for your feedbacks on the contents of these chapters. **Gerben** and **Chris**, thank you for your efforts to make data from **CRV** available for me to use. Unfortunately we had to change our plan, due to time limitation.

Henk, thank you for your friendliness, advices, and recommendations. **Piter**, thank you for the nice discussions I had with you, especially while interpreting the results of the second chapter of this thesis. **Michael**, I really appreciate your help and friendliness throughout this journey. For example, you helped me with codes at the beginning of the journey, and later with presentation skills. You also gave me regular

feedbacks during our quarterly Breed4Food project meetings. Also thanks for the regular Teams 'catch-up' moments during the lockdown. **Biaty**, I am glad I had you around, as I could easily share my struggles with you, and you could easily understand them. Thanks for your brotherly advices to me on PhD-related issues and on other aspects of life. **Chiara**, thanks for our great walks and discussions, and for being there for me through thick and thin. You are indeed a friend-turned-sister! **Shuwen**, thanks for our neighbour-neighbour chats and our after-lunch walks on campus. They gave me quick refreshes after hours of staring at the screen. **Harmen**, **Pascal** and **Yvonne**, thanks a lot for being sources of inspiration to me. **Renzo**, we often encountered similar issues during our PhD trajectories. The two of us working on similar topics ensured that. Thanks for all the discussion sessions we regularly had. **Xiaofei**, **Carolina**, **Rayner, Fatma**, **Anouk** and **other members of our sports team**, thanks for the fun and sporty weekends and evenings we've been having together. I always look forward to these moments, as they help me revitalize after long days/weeks. I hope we continue to have such moments for as long as possible. **Lisette**, **Wilma** and **Fadma**, thanks for all the administrative assistance you provided me. **Lim**, thanks for the chats we regularly had at the coffee corner. On a number of occasions we shared our challenges with each other and felt better by reminding ourselves that everyone is dealing with one challenge or another. **Larissa**, although you joined the group towards the end of my PhD journey, your friendly personality has been uplifting! Thanks for regularly checking on me! **Meenu**, although you left for Zurich after our MSc together, you remain a great friend, and it is always a pleasure to catch up with you. **Robert Mukiibi**, we were supposed to start our MSc programs in Wageningen together. Unfortunately, I could only come to Wageningen after you had finished your program and left. Although we haven't met each other physically, you have been a great friend, regularly checking on me throughout my MSc and PhD programs. I will never forget how you helped me at my weakest point. A friend in need is a friend indeed!

Back at Usmanu Danfodiyo University, Sokoto. **Prof. H.M. Tukur, Prof. S.A. Maigandi, Prof. W.A. Hassan, and Prof. A. Abubakar,** thank you very much for all the support you offered to me when I was trying to get a scholarship to come to Wageningen. **Dr Adamu Mani**, you have been a great colleague, friend and brother. Thanks for your support throughout the journey! **My Nigerian social media friends, especially my schoolmates and students in Sokoto**, thanks a lot for always motivating me to keep going. Trust me it wasn't easy, and on several occasions I felt like giving up. But whenever I remembered that many of you were looking up to me, I got reasons to keep pushing harder. The kind words I get from you whenever I share

something on social media mean a lot to me! **The Nigerian community in Wageningen**, u na weldone o! No dulling for una side. I gbadun all our gists on naija politics, and all the time we jolly together. Una make Wageningen be like Naija for me with all the Naija vibes from una. I hail o! I also have people in Wageningen who are like family to me. **Miranda**, **or my Wageningen Auntie as I call you**, thank you for all the help you offered to me during my MSc program as I went through health challenges. Thanks also for regularly checking on me during my PhD journey. **Peter** and **Alie**, thank you for all the help you render to me and my family. You really make us feel at home. **Nuhu,** thanks for being a reliable friend! I always look forward to our evening walks and chats.

Malou, you made organising QDG meetings a fun activity. I think your very polite way of talking to people contributed a lot to the successes we recorded during our time as 'the QDG team'. Like I told you on several occasions, my interactions with you have made me way politer. **Marieke**, thanks for being a confidence-giving familiar face around at courses and conferences both within and outside the Netherlands. **Malou** and **Marieke,** thanks for being the great friends that you are to me, and for being my reliable sources of information on PhD-related activities that you did before me. Thanks also for proofreading some parts of my thesis and for agreeing to be my paranymphs. It will be an honour to be on the podium with you by my side!

Finally, **my immediate family members**. I cannot thank you enough for your constant prayers and other forms of good wishes towards me. **Usman**, **or Yaa Alhaaj as I call you**, thank you for always giving me your listening ears when I share my struggles with you, and for understanding me and offering me advices. You gave me one of your kidneys despite your advanced age. You simply gave me a second chance to live, and words cannot express how appreciative I am of that gesture! **Mom**, you are my hero! You did all kinds of odd jobs to send my siblings and me to school, despite the fact that you did not go to any western school yourself. Although I don't always do what you want me to do, the love and respect I have for you are unquantifiable. Thank you for not succumbing to societal pressures when raising us as a single mom after the death of our dad. You sacrificed everything for us! **Faiza**, you are my cheerleader. Thank you for being there for me, and I am sorry I was not as available as you wanted me to be. You wanted us to be going to different cities in the weekends. Unfortunately, I was often too tired or too busy to do that. Your moral support for me and patience with me throughout this journey are priceless! **Sarah**, the way you welcome me home makes me eager to come home whenever I'm out. As soon as you hear my footsteps, you shout "daddy" and come to the door to wait for me to arrive. As I come in, you hug me and help me in removing my jacket and shoes. Then you remind me to go and wash my hands in the bathroom. You also always tie my backpack to one the legs of my home desk. I don't know why you do

that, and this is one of the questions I will ask you when you are able to express yourself better. I'm sorry I'm often not able to play with you as much you would want. Sometimes you even come to me and pull my hand towards you, inviting me to come and play with you. But I have to resist. I know now you don't understand my excuses, but I'm sure you will when you grow older.

Colophon

The research described in this thesis was financially supported by the Dutch Ministry of Economic Affairs (TKI Agri & Food project 16022) and the Breed4Food partners Cobb Europe, CRV, Hendrix Genetics and Topigs Norsvin.

Financial support from Wageningen University, the G. Schieter Foundation and Biomoney for printing this thesis is gratefully acknowledged.

Cover designed by Ibrahim Jibrila and Mercedes Benjaminse

Thesis printed by DigiForce | Proefschriftmaken.nl