UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

STRATEGIES TO IMPROVE THE EFFICIENCY OF GENOMIC SELECTION IN ANIMAL BREEDING PROGRAMS

Haroldo Henrique de Rezende Neves

Zootecnista

2013

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

STRATEGIES TO IMPROVE THE EFFICIENCY OF GENOMIC SELECTION IN ANIMAL BREEDING PROGRAMS

Haroldo Henrique de Rezende Neves

Advisor: Profa. Dra. Sandra Aidar de Queiroz

Co-Advisor: Dr. Roberto Carvalheiro

Thesis presented to
the Faculdade de Ciências Agrárias e
Veterinárias Unesp,
Câmpus de Jaboticabal,
in partial fulfilment of requirements for
the degree of Doutor em Genética e
Melhoramento Animal
(*Doctor in Animal Breeding and Genetics)*

2013

Neves, Haroldo Henrique de Rezende
N514s Strategies to improve the efficiency of genomic selection in animal breeding programs / Haroldo Henrique de Rezende Neves. — Jaboticabal, 2013
vi, 162 p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2013
Orientadora: Sandra Aidar de Queiroz
Co-orientador: Roberto Carvalheiro
Banca examinadora: Fabyano Fonseca e Silva, Fernanda Brito, Fernando Sebastián Baldi Rey, Rúsbel Raúl Aspilcueta Borquis
Bibliografia

1. Melhoramento genético. 2. Bovinos de corte. 3. Seleção genômica. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.2:636.082

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

unesp UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE JABOTICABAL

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS DE JABOTICABAL

CERTIFICADO DE APROVAÇÃO

TÍTULO: STRATEGIES TO IMPROVE THE EFFICIENCY OF GENOMIC SELECTION IN ANIMAL BREEDING PROGRAMS

AUTOR: HAROLDO HENRIQUE DE REZENDE NEVES
ORIENTADORA: Profa. Dra. SANDRA AIDAR DE QUEIROZ
CO-ORIENTADOR: Prof. Dr. ROBERTO CARVALHEIRO

Aprovado como parte das exigências para obtenção do Título de DOUTOR EM GENÉTICA E MELHORAMENTO ANIMAL , pela Comissão Examinadora:

Prof. Dr. ROBERTO CARVALHEIRO
Departamento de Zootecnia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

Prof. Dr. FERNANDO SEBASTIÁN BALDI REY
Departamento de Zootecnia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

Prof. Dr. RÚSBEL RAÚL ASPILCUETA BORQUIS
Pós-Doutorando / Departamento de Zootecnia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

Profa. Dra. FERNANDA VARNIERI BRITO
GenSys Consultores Associados / Porto Alegre/RS

Prof. Dr. FABYANO FONSECA E SILVA
Universidade Federal de Viçosa / Viçosa/MG

Data da realização: 12 de setembro de 2013.

CURRICULUM VITAE

HAROLDO HENRIQUE DE REZENDE NEVES – was born in Araxá, Minas Gerais, Brazil, on June 15th 1984, as the second child of Cecilia Maria de Rezende Neves and Antonino Neves de Resende. He started his undergraduate course in Zootecnia (Animal Science) at FCAV/Unesp, Jaboticabal, on February 2003. During the undergraduate studies he was supported by grants from "Programa de Educação Tutorial (PET)" and he was also a volunteer teacher at "Curso Pré-Vestibular Ativo". He received a Bachelor's Degree in Animal Science in 2007. After graduating, he started a M. Sc. within the postgraduate program on "Genética e Melhoramento Animal" (Genetics and Animal Breeding) at FCAV/Unesp, when he received financial support from CAPES and after from FAPESP. His M. Sc. thesis was focused on genetic heterogeneity of residual variance in beef cattle. In 2010, he received the degree of M. Sc. in "Genética e Melhoramento Animal", under advice of Profa. Sandra Aidar de Queiroz and Dr. Roberto Carvalheiro. In this same year he started a Doctorate in "Genética e Melhoramento Animal" at FCAV/Unesp, receiving financial support from FAPESP. In September 2013, he defended his Doctorate thesis, under advice of Profa. Sandra Aidar de Queiroz and Dr. Roberto Carvalheiro.

"Se procurar bem você acaba encontrando,
não a explicação (**duvidosa**) da vida,
mas a poesia (**inexplicável**) da vida."

Carlos Drummond de Andrade

This thesis is dedicated in loving memory to my mother Cecilia, whose love, devotion, strength, faith and wisdom will always guide me.

Dedicated to my father Antonino, who has always encouraged me, providing good examples of hard work and of integrity, who taught me the value of farming. To my brother Aurélio, for his friendship, advice and support. To all my family, especially to my goddaughter Cecília.

ACKNOWLEDGEMENTS

I thank God for the miracle of life, for the strength to carry on.

I would like to express my gratitude to Profa. Sandra Aidar de Queiroz and to my friend Roberto Carvalheiro, for their guidance, friendship, support, advice, encouragement and for every contribution to my personal and professional development.

I would like to express my gratitude to all members of the internal and external examining comitees: Dr. Danísio Prado Munari, Dr. Henrique Nunes de Oliveira, Dr. Fernando Sebastián Baldi Rey, Dr. Rúsbel Raúl Aspilcueta Borquis, Dra. Fernanda Varnieri Brito and Dr. Fabyano Fonseca e Silva, for their kind attention and for their valuable suggestions and contributions to improve this thesis.

Financial support from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) is gratefully acknowledged. In addition, the Faculdade de Ciências Agrárias e Veterinárias - Unesp Jaboticabal (FCAV / Unesp) is acknowledged for all support provided to the development of this thesis as well as for every contribution to my professional development. The support and incentive of the postgraduate program on Genetics and Animal Breeding at FCAV⁄ Unesp is gratefully acknowledged.

I would like to express my gratitude to all professors of the postgraduate program on Genetics and Animal Breeding at FCAV/Unesp (Genética e Melhoramento Animal), for their decisive contribution to my academic formation. The postgraduate program in Computer Science and Computational Mathematics at ICMC-USP (São Carlos) is also gratefully acknowledged.

I would like to thank to all members of the Zebu Genome Consortium, for the constructive discussions and for their generous assistance. I thank to Prof. José Fernando Garcia, for the opportunity to participate in the group. I thank to Yuri, Solomon, Ana Maria, Leonardo Seno, Márcia, Adriana, as well as to many other colleagues, for the friendship, collaboration and all high-level discussions.

I would like to express my gratitude to all members of GenSys Consultores Associados for the collaboration in different moments of my graduate studies.

I wish to thank to all friends and colleagues I met during the graduate studies.

I thank to all staff of FCAV as well as to all professors I had, especially to Profa. Maria Imaculada, my advisor in PET Zootecnia during the undergraduate studies, and to Prof. João Ademir, who provided decisive encouragement for me to start studying animal breeding.

I thank to my friends from Mata Bixera for many remarkable moments throughout these years. I especially thank to Fernando (Pinico), Luiz (Kalango), Anderson (Batata), Otavio (Finkado), Sérgio (Grico), Henrique (Cadera), José Rodolfo (Bernento), Victor (Cervo), Oberdan (Lírou), Bruno (Mininão), Guilherme (Pedala), Mikhael (Teta), Victor (Sofredô), Thiago (Gozo), Daniel (Rombo), Eduardo (Biruleiby), Marco Túlio (K-pão), Gabriel (Xurros), Fábio (Muralha), David (Sarrafo), Aníbal (Babaka), Silas (Delegado), Ivaldo (De-dinho), Lucas (Ex-Seção), Leonardo (Mid Night), Rhaony (Zé-Mayer), Eduardo (Quem?), Ricardo (Limeira), Murilo (Zôtro), Gustavo (Dildo), Henrique (Taradão), Neil (Vápu) and Guilherme (Xangrilá). I thank my friends from the first years in Jaboticabal: Felipe (Tumor), Ronaldo (Bicudo), Gustavo (Bodão), Daniel (Boto), Alexandre (Bigato), Fabiano, Marco Antonio (Koringa) and Fernando (Futum).

I thank to the friends I made during the undergraduate studies in Animal Science (especially those from the XLVII Turma de Zootecnia da FCAV, ZOO 2003), for the support and friendship throughout these years. Thanks to André (Mama-ki), Daniel (Furunku), Daniel (Refugo), Fernando (Mala), Giancarlo (Caxão), Gustavo (Bodão), Gustavo (Vurto), Henrique (Tansso), José Rodolfo (Bernento), Rodrigo (Pinto), Rodrigo (Xoko) and many others.

I would like to express my gratitude to Keyla, for her support, care and patience.

CONTENTS

STRATEGIES TO IMPROVE THE EFFICIENCY OF GENOMIC SELECTION IN ANIMAL BREEDING PROGRAMS

ABSTRACT

This thesis comprises four different studies carried out to evaluate alternative strategies aimed to improve the efficiency of genomic selection (GS) in animal breeding programs. The performance of different statistical methods used to predict the genetic merit of animals based on genomic information was assess in the first study. Five different traits of a real mice dataset were analyzed and methods with large conceptual differences reached very similar predictive abilities in some situations, while a clear re-ranking of methods was observed in function of the trait analyzed.

In the second study, the adequacy of using different response variables (pseudo-phenotypes) to estimate marker effects was investigated through simulation of multi-step genomic evaluation in a large beef cattle population. There was evidence that deregressed proofs comprised a more suitable response variable for both model training and model validation, when compared to estimated breeding values and progeny-yield deviations.

Long-term consequences of application of GS in a beef cattle population undergoing selection were assessed though simulation. Large benefits were envisaged for GS over traditional selection for scenarios mimicking selection for meat quality and female reproduction. There was evidence that larger advantage can be expected for GS compared to BLUP when the selected trait is under less polygenic background and that attributing more weight to favorable alleles of low-frequency can contribute to reduce inbreeding rates and loss of favorable alleles in GS.

Different selective genotyping approaches were compared to update the reference population used in GS, considering a simulated beef cattle population undergoing selection. There was no clear advantage for a particular genotyping strategy under investigation, although, strategies including genotyping of superior males resulted in larger accumulated genetic gain under single-step genomic evaluation.

The results of this thesis highlight the potential of using GS to improve female reproduction and meat quality traits in beef cattle, also showing that genomic information can be a valuable tool to monitor inbreeding incidence and thus enhance strategies to maintain genetic diversity in the long-term.

Keywords: genetic architecture, inbreeding, quantitative trait loci

ESTRATÉGIAS PARA AUMENTO DE EFICIÊNCIA DA SELEÇÃO GENÔMICA EM PROGRAMAS DE MELHORAMENTO GENÉTICO ANIMAL

RESUMO

Esta tese compreende quatro diferentes estudos conduzidos a fim de avaliar estratégias alternativas para aumentar a eficiência de seleção genômica (GS) em programas de melhoramento animal. Um primeiro estudo foi desenvolvido com a finalidade de avaliar a performance preditiva de diferentes métodos estatísticos com base na informação de painéis de marcadores densamente distribuídos ao longo do genoma. Cinco diferentes características de uma população real de camundongos foram analisadas. Verificou-se que métodos com grandes diferenças conceituais apresentaram performance preditiva similar em algumas situações, também havendo variação na performance relativa dos métodos em função da característica analisada.

O uso de diferentes variáveis resposta (pseudo-fenótipos) para estimação de efeitos de marcadores foi avaliado num segundo estudo, por meio da simulação de uma grande população de bovinos de corte, para a qual predições genômicas foram obtidas usando um procedimento de múltiplas etapas. Houve evidência de que provas desregredidas (dEBV) são mais apropriadas do que valores genéticos preditos (EBV) e médias ajustadas de desempenho da progênie (PYD), tanto para o treinamento de modelos quanto para a validação de predições genômicas.

No terceiro estudo, procurou-se avaliar consequências em longo-prazo da aplicação de GS numa população de bovinos de corte sob seleção. Verificou-se grande benefício da aplicação de GS em cenários simulando seleção para características de qualidade de carne e reprodução de fêmeas. Houve evidência de que pode-se esperar maior benefício para GS, quando comparada à seleção por BLUP, no caso de características oligogênicas. Também foi possível inferir que em aplicações de GS, o uso de um critério de seleção em que se atribui maior peso a alelos favoráveis de menor frequência poderia proporcionar menor incidência de endogamia e redução na perda de alelos favoráveis.

No quarto estudo, diferentes estratégias de genotipagem seletiva foram comparadas, numa situação que contemplou a atualização periódica das equações

de predição utilizadas para aplicação de GS numa população simulada de bovinos sob seleção. Não foi possível identificar claro benefício decorrente da aplicação de nenhuma estratégia em particular, embora as estratégias que incluíram genotipagem de machos superiores tenham proporcionado maior progresso genético nos cenários em que predições genômicas foram obtidas por meio de um procedimento de avaliação implementado numa única etapa.

Como resultado geral, evidenciou-se o potencial de aplicação de GS para acelerar o progresso genético em características de reprodução de fêmeas e qualidade de carne em bovinos de corte, bem como a utilidade da informação genômica para monitoramento de endogamia e definição de estratégias mais eficientes para manutenção da diversidade genética em longo-prazo.

Palavras-chave: arquitetura genética, endogamia, locos de característica quantitativa

CHAPTER 1 - General considerations

Introduction

Improvements in production levels and product quality are needed in livestock systems to meet the growing world demand for animal-source foods. Besides this increasing demand, the productive sector must deal with constraints related to competition for land, greenhouse gas emissions and also due to hardening legislation in the fields of environment and animal welfare (FAO, 2011). In this context, animal breeding has played and will continue to play an important role to improve the efficiency of such production systems, especially in terms of competitiveness, safety, sustainability and biodiversity conservation (Harlizius et al., 2004).

The main objective of animal breeding programs is to improve the performance of the next generations, through identification and reproduction of the animals with better genetic pool to efficiently produce in a specific environment (herein, superior animals). In the last decades, animal breeders succeeded in achieving this goal, mostly through the application of statistical tools grounded in quantitative genetics theory, what could be called as 'classical animal breeding'. In this case, the traditional prediction of the genetic merit of individuals (estimated breeding values, EBV) is obtained based on information of pedigree and phenotypes (own records and measures on relatives).

With the advent of dense molecular marker panels, the implementation and design of breeding programs, especially in dairy cattle, had changed dramatically as a consequence of incorporating this new information to identify superior animals earlier and more precisely. Pioneer simulation studies drew attention of animal breeders to the possibility of making accurate predictions of the genetic merit of individuals by using genotypic information from dense marker panels, a process known as genomic selection (GS) (Nejati-Javaremi et al., 1997; Meuwissen et al., 2001). Other influential work in this field was made by Schaeffer (2006), who highlighted the potential benefits of using genomic selection in breeding programs, especially in terms of increasing (doubling) rates of genetic progress and reduction costs (up to 92%) of proving dairy bulls, compared to conventional progeny testing.

The sequencing of bovine genome (Bovine Genome Consortium, 2009) and the availability of dense panels of SNP markers, allowed GS to migrate from simulation to real-word. First applications of GS started in dairy cattle (e.g. Harris et al., 2008; Van Raden et al., 2009). Although individually less informative than multi-allelic markers, SNPs became the markers of choice for GS applications, because they are abundant throughout the genome, have low mutation rate and enable screening samples at relatively low cost. Since then, several studies have focused on the use of genomic selection to enhance both animal and plant breeding.

Genomic selection

The reasoning behind GS is that high-density panels, providing enough genome coverage, would allow markers to capture variability associated to most of the loci (herein QTL) influencing traits of economic relevance. The sum of the effects of such markers would be an accurate predictor of the genetic merit of selection candidates (Meuwissen et al., 2001), given that these effects are estimated simultaneously and that the panel is dense enough. Herein, the prediction of genetic merit obtained through GS will be referred to as genomic estimated breeding value (GEBV).

Markers and QTL must be in sufficient linkage disequilibrium (LD) (i.e. close/associated enough) to ensure that markers will consistently predict QTL effects across generations and/or populations (Hayes et al., 2009; Calus, 2010). As QTL position is generally not known, the average LD between adjacent markers is referred to as the key parameter to evaluate the precision with which GEBV can be predicted, in a way that the lower the extent of LD in a population, the denser must be the marker panel for GS to work (de Ross et al., 2008).

The implementation of genomic selection can be divided in the following steps:
-*Model training:* statistical models are fitted using phenotypic and genotypic information from animals of a 'reference population' (or 'training set') to estimate markers effects for a trait of economic importance. The ideal response variable to estimate marker effects would be the true genetic merit of the individuals (Garrick et al., 2009). Since this information is not available in real-world, the reference populations often involve genotyped animals with alternative types of information

including single or repeated measures of individual phenotypic performance, progeny records and even estimated breeding values (EBV) from genetic evaluations.

In some sense, using EBVs as response variable could be a convenient choice, because this measure is expected to be a more reliable predictor of the genetic merit of individuals, if considerable amount of information is available for EBV calculation, when compared to the use of single phenotypic measurements. On the other hand, Garrick et al. (2009) highlighted some potential problems of this approach and proposed using deregressed EBVs (dEBV) to estimate marker effects. While some studies had pointed out benefit in using dEBV as response variable (e.g. Ostersen et al. 2011), other studies did not confirm such advantage (Guo et al., 2010). A more careful analysis of such discrepancies suggests that more benefit is expected for using dEBVs when the amount of information used to predict EBV is lower and more heterogeneous among reference animals.

Given the costs involved in genotyping animals for high-density panels, strategies are needed to allow more cost-effective scenarios for application of GS, of which one example is selective genotyping. Some studies have identified strategies to choose a more informative set of animals to compose the reference population that would provide greater predictive ability (e.g. genotyping animals with extreme yield deviation values, Boligon et al., 2012).

-*Validation:* before GS is routinely applied, its feasibility needs to be evaluated. For this, the reliability of the estimated marker effects (and thus of the genomic predictions) is verified in an independent set of animals. A general scheme to carry out such validation is forward prediction (e.g. VanRaden et al., 2009; Habier et al., 2011). In this case, data is split such that the information on older animals are included in the reference population, while the GEBV of younger animals ('testing set' or 'validation set') is predicted based on their genotypes. As these younger animals are also progeny-tested, and thus have reliable EBVs, the correlation between GEBV and EBV in the testing set is considered as a good proxy for the reliability of genomic predictions.

According to Daetwyler et al. (2013), when there is lack of pedigree or when highly accurate EBVs are not available to apply forward prediction, replicated cross-validation approaches can be used to evaluate prediction accuracy, so that strategies

similar to K-fold cross-validation (e.g. Legarra et al., 2008; Saatchi et al., 2012) are often employed in this situation.

According to Calus (2010), there is no consensus on the ideal design of reference and validation populations, what is driven by economical or practical constraints in some situations. However, some studies have compared different strategies for such design, suggesting that multi-generational reference populations would allow lower decrease in prediction accuracy due to marker-QTL associations being broken by recombination along the generations (e.g. Muir, 2007) as well as proposing alternatives that would reduce the need of re-estimation of marker effects every generation (Solberg et al., 2009). While such issues are important for species in which generation intervals are smaller and/or phenotyping is too costly, in dairy cattle, the routine genetic evaluations already include periodic update of reference populations, in a way that all phenotypic and genotypic information available at each time is employed to re-estimate the marker effects, with various genomic evaluations being released along the year (Wiggans et al., 2011).

-*Selection:* Once the genomic predictions are found reliable enough, GS can be used to predict the genetic merit of animals, as soon as they are genotyped. Schaeffer (2006) highlighted the potential benefits of this strategy in terms of reducing generation intervals, increasing prediction accuracies and selection intensities and reducing breeding organization costs as well, while it has been argued that GS could be an effective way to enhance genetic evaluations of difficult-to-measure traits as well as of traits expressed late in life (Goddard et al., 2010; Dekkers, 2012).

Statistical methods to estimate markers effects

The estimation of marker effects can be treated as a multiple regression problem, in which phenotypes for a trait of economic relevance are the response variable, while the genotypes for SNP markers are the explanatory variables. This situation typically constitutes a “large p, small n problem” because the number of phenotypes is generally much lower than the number of markers whose effects need to be estimated (Solberg et al., 2009). Thus, there are not enough degrees of freedom to estimate all maker effects through ordinary least squares regression and

models may suffer of multicollinearity, especially because markers in close positions are expected to be highly correlated.

The basic statistical model employed for marker effect estimation in the influential study of Meuwissen et al. (2001) can be described as:

$$y = 1_n\mu + \sum_{j=1}^{p} X_j g_j + e, \quad (1)$$

where **y** is the vector of phenotypes (of length **n**), **μ** is an overall mean, 1_n is a vector of ones (of order **n)**, $\mathbf{g_j}$ and $\mathbf{X_j}$ represent the allele substitution effect of the j-th marker and an incidence matrix relating **y** to the effect of marker **j**, respectively, and **e** is a random residual, under the assumption that $\mathbf{e \sim N(0, \sigma^2_e)}$.

In practical terms, the main difference among the different methods proposed in Meuwissen et al. (2001) relies on the assumptions about the variances of marker effects. In one method, that can be considered a ridge regression procedure and that will be regarded to as RR-GBLUP hereafter, the best linear unbiased prediction of marker effects was obtained under the assumption that they follow the same normal distribution, with variance σ^2_m, so that mixed model equations are solved considering a variance ratio (λ) obtained as $\lambda = \sigma^2_e/(\sigma^2_m/k)$, $k=2\sum p_i - (1 - p_i)$ and p_i is the allelic frequency of the i^{th} marker, what reflects the fact that more polymorphic loci contribute more to the genetic variation.

In the Bayes A method (Meuwisen et al., 2001), a model similar to that in (1) was implemented under a Bayesian regression framework, so that marker-specific variances (σ^2_j) are fitted. The priors for such variances were defined as $\mathbf{\sigma^2_j \sim \chi^{-2}(v,S)}$, and the hyperparameters were chosen so that the prior distribution of marker effects is a scaled t distribution. The Bayes B method (Meuwisen et al., 2001) included an additional Metropolis-Hastings step, that induces a posterior distribution of marker effects having higher density at 0. In this method, marker effects are assumed to be either zero with a high probability (usually 0.95 or larger) or have their variances estimated assuming a prior distribution similar to that described for Bayes A.

A variant of RR-GBLUP consists in setting up a genomic relationship matrix based on genotypes from high-density panels and use it to replace the numerator relationship matrix in the regular BLUP equations employed for genetic evaluation. This procedure will be referred to as GBLUP hereafter. Goddard (2009)

demonstrated that both approaches (RR-GBLUP and GBLUP), under some assumptions, are equivalent.

Besides the methods proposed in Meuwissen et al. (2001), many other approaches have been applied for estimation of marker effects since then. Some of them involve dimensionality reduction techniques commonly used in other fields, among which partial least squares (Solberg et al., 2009), principal component analysis (Macciotta & Gaspa, 2009; Solberg et al., 2009) and the LASSO (Usai et al., 2009). Bayesian implementations of the LASSO (de los Campos et al., 2009; Legarra et al., 2011) have been applied in genomic prediction as well.

While some simulation studies had suggested larger predictive ability for Bayes B, when compared to other methods (Meuwissen et al., 2001; Clark et al., 2011), one drawback of such approach is that it can be too demanding computationally. Thus, to overcome such limitations, alternative Bayesian regression approaches of similar concept have been proposed: SSVS (Verbyla et al., 2009), Bayes Dpi and Bayes Cpi (Habier et al., 2011) and BayesR (Erbe et al., 2012).

A comprehensive review of genomic prediction methods can be found in de los Campos et al. (2013). These authors classified genomic prediction methods into three groups: penalized methods, Bayesian methods and non-parametric methods. In the case of penalized methods, marker effect estimation can be viewed as an optimization problem in which the objective function includes both goodness of fit to the training data (usually through the residual sum of squares) and a penalty function on model complexity.

Common penalty functions in penalized methods are based on either the sum of squares of regression coefficients (RR-GBLUP) or on the sum of absolute regression coefficients (LASSO), while a weighted average of these two penalty functions would originate the elastic net (EN) method, in an attempt to combine features of both methods (de los Campos, 2013).

Most Bayesian regression procedures used in genomic prediction share the same sampling model and basically differ regarding to the prior density of marker effects, what would define the type and extent of shrinkage induced on estimated marker effects as well as could induce variable selection (de los Campos et al., 2013). A Bayesian counterpart of RR-GBLUP is Bayesian Ridge Regression (BRR),

for which a normal prior is assumed for maker effects. According to de los Campos et al. (2013), this assumption could lead to sub-optimal solutions, for instance when some markers are in regions not linked to any QTL.

At higher marker density, a small proportion of the markers is expected to be statistically associated to QTL and thus have effect on the trait. For this reason, other prior densities that have higher mass at zero and thicker tails have been postulated for marker effects, aiming to induce stronger shrinkage of estimates of markers with small effects, while allowing less shrinkage of estimates of markers with larger effects. The most common prior densities with these features are the scaled t (Bayes A) and the double exponential (Bayesian LASSO).

Another group of Bayesian regression methods assume that marker effects follow a mixture of densities with large and small variances (e.g. SSVS, Bayes R), what could induce differential shrinkage of estimates of marker effects. In order to induce variable selection, other group of methods postulate a combination of point mass at zero and a density with large variance (slab) as a prior for marker effects (de los Campos et al., 2013), thus resulting in a proportion (π) of marker not having effect on the trait. Bayes B (for which the slab is a scaled t) and Bayes Cpi (for which the slab is a normal density) are examples of this last group.

As mentioned earlier, in many Bayesian regression methods, some hyperparameters are often fixed at arbitrary values (e.g. π in Bayes B), what has motivated some criticism. As suggested in Gianola (2013), a more elegant approach to this problem would be to assign priors for such hyperparameters and infer them from the resulting model. However, in a situation when $\mathbf{p} >> \mathbf{n}$ and with long-range LD, it may not be possible to estimate all these parameters jointly (de los Campos et al., 2013).

Moreover, Gianola (2013) alerted to the problem of making inferences based marker effects estimated obtained using Bayesian linear regression methods when $\mathbf{p} >> \mathbf{n}$. For such methods, unless $\mathbf{n} >> \mathbf{p}$, priors will always be influential, thus Bayesian learning would be imperfect. This author argue that such Bayesian regressions could be able to deliver reasonable predictions of complex traits but would be of limited value to infer their genetic architecture.

While in most of the methods discussed earlier, genomic values can be represented as parametric functions of marker genotypes, this does not hold in the case of other methods used to estimate marker effects, including non-parametric regression approaches (Gianola et al., 2006; Bennewitz et al., 2009) and machine learning methods (e.g. Long et al., 2007; González-Récio & Forni, 2011). An important motivation for using non-parametric methods in genomic prediction would be the possibility to enhance predictions for traits influenced by non-additive effects. It is worth to emphasize that in this case these methods predict genotypic values rather than purely additive breeding values.

Although simulation studies have suggested the superiority of methods based on some sort of variable selection over GBLUP and RR-GBLUP (Meuwissen et al., 2001; Habier et al., 2007; Solberg et al., 2008; Clark et al, 2011), this advantage rarely has been verified with real data. Previous studies comparing genomic predictions with different methods using real data indicated that RR-GBLUP performed comparably or better than variable selection methods (Hayes et al., 2009; Luan et al., 2009; Moser et al., 2009; Legarra et al., 2011), although there is evidence that substantially higher accuracy can be achieved using variable selection methods for some traits (e.g. traits affected by DGAT1) (VanRaden et al., 2009; Legarra et al., 2011).

It has been argued that shrinkage methods with assumptions close to the infinitesimal model (i.e. RR-GBLUP and its variants) are robust with respect to the underlying genetic architecture of the traits, while the predictive performance of variable selection methods is more variable in function of the genetic background of the traits (Daetwyler et al., 2010).

Herein, most of the discussion conducted on statistical methods was focused on genomic prediction applied in a multi-step framework. In many situations, either genotyped animals do not have their own phenotypic records and/or there is information available on relatives of the genotyped animals that could be used to enhance predictions. Under a multi-step approach, the first step would be preprocessing the phenotypic data in a way that a pseudo-phenotype would be generated for each genotyped animal (de los Campos et al., 2013), after which pseudo-phenotypes could be used as response variable to estimate marker effects

and compute genomic predictions. In addition, genomic predictions can be combined (blended) to traditional predictions of genetic merit (EBV), as is the case of some dairy cattle breeding programs (e.g. VanRaden et al., 2009).

One alternative to the multi-step approach is single-step genomic prediction (Legarra et al., 2009). In such procedure, all available information (including genotypic information, pedigree and phenotypes) can be employed in genetic evaluation, so that the pedigree relationship matrix is augmented with genomic information. This procedure allows to generate predictions for both genotyped and non-genotyped animals simultaneously, being that model assumptions are close to those of GBLUP.

Impact of genomic selection on inbreeding incidence and long-term genetic gain

When compared to traditional selection based on BLUP, in addition to enable increased rates of genetic progress, GS schemes could also reduce the levels inbreeding accumulated along the selection process (Daetwyler et al., 2007). This would be justified by the fact that genomic prediction would allow better estimation of the Mendelian sampling term (MS), i.e. of the component of the true breeding value (TBV) defined as the deviation of the parent average (PA). This feature would result in smaller co-selection of closely related animals when compared to BLUP.

On the other hand, large reductions in generation intervals under GS prediction could lead to an increase of annual inbreeding rates. However, the risks associated to inbreeding incidence are more relevant on a generation basis, since the processes that counterbalance this phenomenon also occur in time horizon equivalent to one generation (Daetwyler et al, 2007). According to these same authors, the possibility of more intense selection in GS schemes (e.g. screening of a larger number of selection candidates) could also contribute to increase the levels of inbreeding.

According to Woolliams et al. (1999), selection criteria with more emphasis on the MS term are crucial to guarantee long-term genetic gains as well as to reduce inbreeding rates per generation. The more accurate prediction of the genetic merit of females would allow more emphasis on this selection path, leading to more balanced

contributions between both sexes in the long-term and possibly allowing to reduce inbreeding rates when compared to BLUP selection, at a same level of genetic gain. Avendaño et al. (2004) also argued that the MS term is the main component that allow optimum contribution selection (OCS) strategies to be effective at controlling inbreeding rates and increase genetic progress.

As a general rule, many aspects previously discussed draw attention for the potential benefits of GS in situations in which genomic predictions allow more accurate estimation of the Mendelian sampling term. Anyway, further studies, possibly involving more detailed simulations, are needed to investigate the relationship between genomic selection and inbreeding incidence, as well as other long-term consequences of such a strategy.

The large amount of genomic information made available by high-density panels could allow the development of estimators of inbreeding that would be more informative than pedigree-based estimators (Villanueva et al., 2005; Daetwyler et al., 2007).

Outline of the thesis

This thesis comprises four different studies carried out to evaluate alternative strategies aimed to improve efficiency of genomic selection in animal breeding programs. The study presented in Chapter 2 was devoted to test the performance of different statistical methods used to predict the genetic merit of animals based on information of genome-wide dense molecular markers, after analyzing a real mice dataset from a public database. The focus was on the interplay between the genetic background of the traits and the relative performance of different methods as well as on the influence of the extent of relatedness between selection candidates and the reference population on such predictions. In this scope, better understanding of such inter-relationships as well as the identification of more suitable statistical methods to estimate marker effects could make an important contribution to improve the efficiency of GS.

The accuracy delivered by genomic predictions is a key factor that would determine the feasibility of GS schemes and it is directly associated to the precision with which the marker effects are estimated. In Chapter 3, the adequacy of using

different response variables (pseudo-phenotypes) to estimate marker effects was investigated. A large beef cattle population was simulated to evaluate the impact of using different pseudo-phenotypes in a multi-step genomic evaluation, aiming to identify more appropriate procedures for both model training and validation of genomic predictions in such situation.

The knowledge about the potential of different selection strategies to enable genetic gain and maintain genetic diversity in longer time horizons is essential for the sustainability of breeding schemes. A third study (Chapter 4) was designed to assess long-term consequences of application of GS in a simulated cattle population undergoing selection. Such study aimed to provide some insight on the potential consequences of GS under different scenarios of selection, while alternative strategies to monitor and control inbreeding incidence were investigated.

Using selective genotyping strategies has been suggested as an alternative to improve cost-effectiveness of GS, while some drawbacks have also been identified for some of such strategies. Chapter 5 addresses the question about how different selective genotyping approaches would perform when updating the reference population used in GS, considering a simulated beef cattle population undergoing selection, also evaluating the quality of genomic predictions obtained using different statistical methods, in both short and long-term.

Literature cited

AVENDAÑO, S.; WOOLLIAMS, J. A.; VILLANUEVA, B. Prediction of accuracy of estimated Mendelian sampling terms. **Journal of Animal Breeding and Genetics,** Berlin, v. 122, n. 5, p. 302–308, 2005.

BENNEWITZ, J.; SOLBERG, T.; MEUWISSEN, T. H. E. Genomic breeding value estimation using nonparametric additive regression models. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 41, n. 20, 2009. Available at: < http://dx.doi.org/10.1186/1297-9686-41-20>.

BOLIGON, A. A.; LONG, N.; ALBUQUERQUE, L. G.; WEIGEL, K. A.; GIANOLA, D.; ROSA, G.J.M. Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection. **Journal of Animal Science,** Champaign**,** v. 90, n. 13, p. 4716-4722, 2012.

BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM. The genome sequence of taurine cattle: a window to ruminant biology and evolution. **Science,** Washington DC**,** v. 324, n. 5926, p. 522-528, 2009.

CALUS, M. P. L. Genomic breeding value prediction: methods and procedures. **Animal**, Cambridge, v. 4, n. 2, p 157-164, 2010.

CLARK, S. A.; HICKEY, J. M.; VAN DER WERF, J. H. J**.** Different models of genetic variation and their effect on genomic evaluation. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 43, n. 18, 2011. Available at: <http://dx.doi.org/ 10.1186/1297-9686-43-18>.

DAETWYLER, H. D.; CALUS, M. P.; PONG-WONG, R.; DE LOS CAMPOS, G.; HICKEY, J. M. Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking. **Genetics,** Bethesda, v. 193, n. 2, p. 347-365, 2013.

DAETWYLER, H. D.; PONG-WONG, R.; VILLANUEVA, B.; WOOLLIAMS, J.A. The impact of genetic architecture on genome-wide evaluation methods. **Genetics,** Bethesda, v. 185, n. 3, p. 1021-1031, 2010.

DAETWYLER, H. D.;VILLANUEVA, B.; BIJMA, P.; WOOLIAMS, J. A. Inbreeding in genome-wide selection. **Journal of Animal Breeding and Genetics,** Berlin, v. 124, n. 6, p. 369–376, 2007.

DE LOS CAMPOS, G.; NAYA, H.;GIANOLA, D.; CROSSA, J.; MANFREDI, E.; WEIGEL, K. A.; COTES, J. M. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics,** Bethesda, v. 182, n. 1, p. 375–385, 2009.

DE LOS CAMPOS, G.; HICKEY, J.; PONG-WONG, R.; DAETWYLER, H. D.; CALUS, M. P. L. Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. **Genetics,** Bethesda, v. 193, n. 2, p. 327-345, 2013.

DE ROOS A. P. W.; HAYES, B. J.; SPELMAN, R.; GODDARD, M. E. Linkage disequilibrium and persistence of phase in Holstein Friesian, Jersey and Angus cattle. **Genetics,** Bethesda, v. 179, n. 3, p. 1503-1512, 2008.

DEKKERS, J. C. M. Application of Genomics Tools to Animal Breeding. **Current Genomics**, Sharjah, v. 13, n. 3, p. 207-2012, 2012.

ERBE, M.; HAYES, B. J.; MATUKAMULLI, L. K.; GOSWAMI, S.; BOWMAN, P. J.; REICH, C. M.; MASON, B. A.; GODDARD, M. E. Improving accuracy of genomic predictions within and between dairy cattle breeds with high density SNP panels. **Journal of Dairy Science,** Champaign, v. 95, n. 7, p. 4114-4129, 2012.

FAO. 2011. Mapping supply and demand for animal-source foods to 2030, by T.P. Robinson & F. Pozzi. Animal Production and Health Working Paper. No. 2. Rome.

GARRICK, D.J; TAYLOR, J.F.; FERNANDO, R .L. Deregressing estimated breeding values and weighting information for genomic regression analyses. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 41, n. 55, 2009. Available at: <http://dx.doi.org/ 10.1186/1297-9686-41-55>.

GIANOLA, D.; FERNANDO, R. L.; STELLA, A. Genomic assisted prediction of genetic value with semiparametric procedures. **Genetics,** Bethesda, v. 173, n. 3, p. 1761–1776, 2006.

GIANOLA, D. Priors in Whole-Genome Regression: The Bayesian Alphabet Returns. **Genetics,** Bethesda, v. 194, n. 3, p. 573-596, 2013.

GODDARD, M.E.; HAYES, B.J.; MEUWISSEN T.H. Genomic selection in livestock populations. **Genetics Research,** Cambridge, v. 92, n. 5-6, p. 413-421, 2010.

GODDARD, M. E**.** Genomic selection: Prediction of accuracy and maximisation of long term response. **Genetica**, Dordrecht, v. 136, n. 2, p. 245-257, 2009.

GONZÁLEZ-RECIO, O.; FORNI, S. Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 43, n. 7, 2011. Available at: < 10.1186/1297-9686-43-7>.

GUO, G.; LUND, M.S.; ZHANG, Y.; SU, G. Comparison between genomic predictions using daughter yield deviation and conventional estimated breeding value as response variables. **Journal of Animal Breeding and Genetics,** Berlin, v. 127, n. 6, p. 423-432, 2010.

HABIER, D.; FERNANDO, R. L.; KIZILKAYA, K.; GARRICK, D. J. Extension of the bayesian alphabet for genomic selection. **BMC Bioinformatics,** London, v. 12, n. 186, 2011. Available at: <http://dx.doi.org/10.1186/1471-2105-12-186>.

HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. The impact of genetic relationship information on genome-assisted breeding values. **Genetics,** Bethesda, v. 177, n.4, p. 2389–2397, 2007.

HARLIZIUS, B.; VAN WIJK, R.; MERKS, J. W. Genomics for food safety and sustainable animal production. **Journal of Biotechnology,** Amsterdam, v. 113, n. 1-3, p. 33-42, 2004.

HARRIS, B.L.; JOHNSON, D. L.; SPELMAN, R. J. Genomic selection in New Zealand and the implications for national genetic evaluation. In: International Committee for Animal Recording Biennial Session, 36., 2008, Niagara Falls, Canada. **Proceedings...**Niagara Falls: ICAR, 2008, p. 325-330.

HAYES, B.J.; BOWMAN, P.J.; CHAMBERLAIN, A.J.; GODDARD, M. E. Invited review: Genomic selection in dairy cattle: Progress and challenges. **Journal of Dairy Science,** Champaign, v. 92, n. 3, p. 433–443, 2009.

LEGARRA, A.; ROBERT-GRANIÉ, C.; CROISEAU, P.; GUILLAUME, F.; FRITZ, S. Improved Lasso for genomic selection. **Genetics Research,** Cambridge, v. 93, n. 1, p. 77-87, 2011.

LEGARRA, A.; ROBERT-GRANIÉ, C.; MANFREDI, E.; ELSEN, J-M. Performance of genomic selection in mice. **Genetics,** Bethesda, v. 180, n. 1, p. 611-618,2008.

LEGARRA, A.; AGUILAR, I.; MISZTAL, I. A relationship matrix including full pedigree and genomic information. **Journal of Dairy Science,** Champaign, v. 92, n. 9, p. 4656-4663, 2009.

LONG , N.; GIANOLA, D.; ROSA, G. J. M.; WEIGEL, K. A.; AVENDAÑO S. Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. **Journal of Animal Breeding and Genetics,** Berlin, v. 124, n. 6, p. 377-389, 2007.

LUAN, T.; WOOLLIAMS, J.A.; LIEN, S.; KENT, M.; SVENDSEN, M.; MEUWISSEN, T.H.E. The accuracy of genomic selection in Norwegian Red cattle assessed by cross-validation. **Genetics,** Bethesda, v. 183, n. 3, p.1119-1126, 2009.

MACCIOTTA, N. P. P.; GASPA, G. Use of principal component and factor analysis to reduce the number of independent variables in the prediction of genomic breeding values. **Italian Journal of Animal Science,** Pavia, v. 8 (suppl. 2), n. 1, p. 105-107, 2009.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics,** Bethesda, v. 157, n. 4, p. 1819–1829, 2001.

MOSER, G.; TIER, B.; CRUMP, R. E.; KHATKAR, M.S.; RAADSMA, H. W. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 41, n. 56, 2009. Available at: < http://dx.doi.org/10.1186/1297-9686-41-56>.

MUIR, W. M. Comparison of genomic and traditional BLUP estimated breeding value accuracy and selection response under alternative trait and genomic parameters. **Journal of Animal Breeding and Genetics,** Berlin, v. 124, n. 6, p. 342–355, 2007.

NEJATI-JAVAREMI, A.; SMITH, C.; GIBSON, J.P. Effect of total allelic relationship on accuracy of evaluation and response to selection. **Journal of Animal Science,** Champaign**,** v. 75, n. 7, p. 1738–1745, 1997.

OSTERSEN, T.; CHRISTENSEN, O. F.; HENRYON, M.; NIELSEN, B.; SU, G.; MADSEN, P. Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in purebred pigs. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 43, n. 38, 2011. Available at: <http://dx.doi.org/10.1186/1297-9686-43-38>.

SAATCHI, M., SCHNABEL, R. D.; ROLF, M. M.; TAYLOR, J. F.; GARRICK, D. J. Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 44, n. 38, 2012. Available at: < http://dx.doi.org/10.1186/1297-9686-44-38>.

SCHAEFFER, L.R. Strategy for applying genome-wide selection in dairy cattle. **Journal of Animal Breeding and Genetics,** Berlin, v. 123, n. 4, p. 218–223, 2006.

SOLBERG, T. R.; SONESSON, A. K.; WOOLLIAMS, J. A., MEUWISEEN, T. H. E. Reducing dimensionality for prediction of genome-wide breeding values. **Genetics Selection Evolution,** Jouy-en-Josas**,** v. 41, n. 29, 2009. Available at: < http://dx.doi.org/ 10.1186/1297-9686-41-29>.

USAI, M.G.; GODDARD, M.E.; HAYES, B.J. LASSO with cross-validation for genomic selection. **Genetics Research,** Cambridge, v. 91, n. 6, p. 427-436, 2009.

VANRADEN, P.M.; VAN TASSELL, C.P.; WIGGANS, G.R.; SONSTEGARD, T.S.; SCHNABEL, R.D.; TAYLOR, J.F.; SCHENKEL, F.S. Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science,** Champaign**,** v. 92, n. 1, p. 16–24, 2009.

VERBYLA, K.L.; HAYES, B.J.; BOWMAN, P.J.; GODDARD, M.E. Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle. **Genetics Research,** Cambridge, v. 91, n. 5, p. 307-311, 2009.

VILLANUEVA, B.; PONG-WONG, R.; FERNANDEZ, J.; TORO, M. A. Benefits from marker-assisted selection under an additive polygenic genetic model. **Journal of Animal Science,** Champaign**,** v. 83, n. 8, p. 1747–1752, 2005.

WOOLLIAMS, J.A.; BIJMA, P.; VILLANUEVA, B. Expected genetic contributions and their impact on gene flow and genetic gain. **Genetics,** Bethesda, v. 153, n. 2, p. 1009–1020, 1999.

CHAPTER 2 - Statistical methods for genomic selection in a mice population

Abstract

Background

The availability of high-density panels of SNP markers has opened new perspectives for marker-assisted selection strategies, such that genotypes for these markers are used to predict the genetic merit of selection candidates. Because the number of markers is often much larger than the number of phenotypes, marker effect estimation is not a trivial task. The objective of this research was to compare the predictive performance of ten different statistical methods employed in genomic selection, by analyzing data from a heterogeneous stock mice population.

Results

For the five traits analyzed (W6W: weight at six weeks, WGS: growth slope, BL: body length, %CD8+: percentage of CD8+ cells, CD4+/ CD8+: ratio between CD4+ and CD8+ cells), within-family predictions were more accurate than across-family predictions, although this superiority in accuracy varied markedly across traits. For within-family prediction, two kernel methods, Reproducing Kernel Hilbert Spaces Regression (RKHS) and Support Vector Regression (SVR), were the most accurate for W6W, while a polygenic model also had comparable performance. A form of ridge regression assuming that all markers contribute to the additive variance (RR-GBLUP) figured among the most accurate for WGS and BL, while two variable selection methods (LASSO and Random Forest, RF) had the greatest predictive abilities for %CD8+ and CD4+/ CD8+. RF, RKHS, SVR and RR-GBLUP outperformed the remainder methods in terms of bias and inflation of predictions.

Conclusions

Methods with large conceptual differences reached very similar predictive abilities and a clear re-ranking of methods was observed in function of the trait analyzed. Variable selection methods were more accurate than the remainder in the case of %CD8+ and CD4+/CD8+ and these traits are likely to be influenced by a smaller number of QTL than the remainder. Judged by their overall performance

across traits and computational requirements, RR-GBLUP, RKHS and SVR are particularly appealing for application in genomic selection.
Keywords: kernel regression, LASSO, Random Forest, ridge regression, SNP, subset selection

Background

The availability of high-density panels of single nucleotide polymorphisms (SNP) containing thousands of markers opened new perspectives for the study of complex diseases, while has enhanced marker-assisted selection strategies in animal and plant breeding.

The possibility to predict accurately the genetic merit of selection candidates based on their genotypes for SNP markers, a process known as genomic selection [1], is revolutionizing breeding schemes. The reasoning of this process is that whenever marker density is high enough, most QTL will be in high linkage disequilibrium (LD) with some markers and estimates of marker effects will lead to accurate predictions of genetic merit for a trait.

Despite this, the amount of information to be analyzed in this situation poses new challenges from statistical and computational viewpoints. As the number of predictor variables (markers) is generally much higher than the number of observations (phenotypes), there is lack of degrees of freedom to estimate all marker effects simultaneously, what is aggravated by the fact that models may suffer from multicollinearity, especially because markers in close positions are expected to be highly correlated.

According to review in [2], some of the alternatives that have been employed to overcome these issues are fitting markers as random effects (e.g. shrinkage estimation and Bayesian regression) or applying some dimensionality reduction technique or machine learning method, although there is no consensus on the most appropriate method for genomic predictions. It has been argued that shrinkage methods with assumptions close to the infinitesimal model (i.e. RR-GBLUP and its variants) are robust with respect to the underlying genetic architecture of the traits, while methods based on some sort of variable selection are more sensitive to the genetic background of traits ([3],[4]).

There are still few extensive studies aimed to compare predictive performance of such methods in plants or in animals [5]. In the present study, we analyze a publicly available dataset, including pedigree, genotypic and phenotypic information of a mice population. Although this same dataset had already been analyzed previously ([6], [7], [8]), we focus on a broader comparison of statistical methods employed for genomic prediction, by studying five traits that probably have considerable differences in terms of genetic architecture.

Thus, the objective of this research was to compare the predictive performance of ten different statistical methods employed in genomic selection by using data from a heterogeneous stock mice population, aiming to provide some insight in the scope of statistical methods useful for genomic selection and in the interplay between the genetic background of traits and the performance of these methods.

Methods

Data

The data came from a heterogeneous stock mice population kept by The Welcome Trust Centre for Human Genetics (WTCHG) (data are available at http://gscan.well.ox.ac.uk). Briefly, this population was generated from the crossing of eight inbred lines, followed by 50 generations of random mating. As a result, this population exhibits a high level of linkage disequilibrium, even for pairs of markers separated by until 2Mb [9]. When considering genotypic information obtained with a panel with 11.558 SNP markers and average inter-marker distance of 204 kb, the average r^2 between adjacent markers was about 0.62 [6]. This amount of LD enhanced QTL mapping for complex traits in mice [10] and would be equally helpful in the context of genomic selection, besides the fact that knowledge of the origin of this population could improve interpretability of the results.

Only animals with both genotypes and phenotypes were considered and details of sampling and genotyping are described in Valdar et al. [11]. The raw data included genotypes for 12,226 SNP markers located in autosomes of 1,940 animals. Data were edited such that only polymorphic markers with MAF ≥ 5% and with no

evidence of departure from Hardy-Weinberg equilibrium were considered in analyses.

Missing genotypes (0.1%) were imputed using probabilistic PCA (PPCA, [12]). Although the accuracy of this procedure is slightly lower than that of other methods, computing time is much lower. In addition, the proportion of missing genotypes is small enough to neglect the effects of imputation. After data editing, a dataset including information of 1,884 animals for 9,917 markers was considered in marker effect estimation, such that 168 full-sib families with average size of 11 were represented.

Five traits whose heritabilities are quite different were analyzed: percentage of CD8+ cells (%CD8+, h^2=0.89), ratio between CD4+ and CD8+ cells (CD4+/ CD8+, h^2=0.80), body weight at 6 weeks (W6W, h^2 = 0.74), growth slope (WGS, h^2=0.30), body length (BL, h^2=0.13) [11]. Aiming to reduce computing times, phenotypes for each trait were pre-corrected for the significant environmental effects reported by [11].

Regarding to the genetic architecture of the traits in this study, an analysis of the supplementary material in [10] revealed that 17, 11, 19, 10 and 6 QTL were found to be significant on %CD8+, CD4+/ CD8+, W6W, GS and BL, respectively. For the first three of these traits, the QTL mapped were responsible for more than 30% of the their variance (Table 1). The largest QTL with effects on %CD8+ and CD4+/ CD8+ explained about 8.0% and 12% of the variance of these traits, respectively, while the largest QTL on the other traits only accounted for about 3% or less of their variance.

Table 1 - Available information* on the genetic architecture of the traits in study

Trait	Nº QTL	variance explained (%)	largest QTL (%)	heritability
%CD8+	17	36.3	8.00	0.89
CD4+/CD8+	11	33.1	11.90	0.80
W6W	19	38.3	3.20	0.74
WGS	10	20.6	2.40	0.30
BL	6	16.7	3.10	0.13

Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Nº QTL = number of QTL mapped, proportion of the variance explained by them (variance explained, in %) and proportion of the variance explained by the QTL with largest effect(largest QTL, in %). *Information published in Valdar et al. [10] and Valdar et al. [11].

When analyzing this dataset, Legarra et al. [6] alerted for the non-random allocation of animals between cages, in a way that many full-sib groups were kept in the same cage and thus additive and environmental effects were confounded at this level. For this reason, phenotypes were also adjusted for this random effect.

For each trait, REML estimates of variance components were obtained using an animal model including all information available (pedigree and phenotypic records) and then phenotypes were adjusted for the environmental effects described previously. In this way, the adjusted phenotypes were obtained as a sum of polygenic and residual effects estimated from the animal model.

Study design

As our focus rely on the comparison of the performance of methods employed to estimate marker effects using real data, we employed a design similar to that employed by [6]. A cross-validation strategy was applied, such that data were split in two sets, reference (REF) and validation (VAL). For all methods, only the information on REF was employed to train the model, then solutions obtained in this step were used to predict the phenotypes of the animals in the VAL set.

The Pearson's correlation between phenotypes and their respective predictions ($r_{y,\hat{y}}$), hereinafter regarded as “predictive ability”, would allow comparison of predictive performance across methods. This approach has also a valuable interpretation in the context of animal breeding: the prediction of unobserved phenotypes mimics the prediction of the future performance of individuals in the population, as discussed in [6], in a way that the expected responses to selection using different methods could be compared.

Two-strategies for sampling animals were applied: (1) within-families, full-sib families were split such that about 55% (45%) of animals with phenotypes were included in the REF (VAL) set; (2) across-families, entire full-sib families were included in the REF set and used to predicted the observations of animals of other families (VAL set), such that REF set also comprised about 55% of the animals with phenotypes (Table 2). For each trait, ten replicates of each splitting strategy were done, such that empirical standard errors of parameters of interest were calculated based on nearly equal-sized partitions, ensuring that results were not due to random

splitting of data. To ensure a more precise comparison, the different methods were applied in exactly the same partitions of the data.

Table 2 - Summary statistics* pertaining to phenotypic data employed in cross validation**

Trait	Split	N	Training size			Phenotypes training		Phenotypes testing	
			Min	Ave	Max	Mean	SD	Mean	SD
W6W	within	1925	1059	1061.9	1066	-0.155	1.96	-0.188	1.95
WGS	within	1917	1056	1059.6	1068	0.001	0.04	0.002	0.04
BL	within	1840	1013	1017.9	1035	-0.002	0.40	-0.006	0.40
%CD8+	within	1407	774	778.6	785	0.010	4.34	0.122	4.36
CD4+/CD8+	within	1403	772	774.5	781	0.003	0.07	0.004	0.07
W6W	across	1925	1059	1067.6	1081	-0.162	1.95	-0.180	1.96
WGS	across	1917	1057	1063.3	1076	0.001	0.04	0.002	0.04
BL	across	1840	1014	1022.3	1030	-0.002	0.40	-0.005	0.40
%CD8	across	1407	775	780.9	791	-0.149	4.23	0.319	4.43
CD4+/CD8+	across	1403	774	782.2	799	0.003	0.07	0.003	0.07

*N = total number of phenotypic records. Minimum, average and maximum size of training set (Min, Ave and Max) and mean and standard deviation (SD) of the adjusted records considered in training and testing sets (averaged across replicates). **Traits considered: weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Split = splitting strategy in cross-validation (within or across-family).

It is important to emphasize that full-sib families considered in REF and VAL sets in this study are linked by distant relationships in the case of across-family splitting [6]. Thus, within-family predictions are expected to account for more recent relationships, while across-family predictions would mostly pick up LD persistent among families (i.e. older relationships).

Genomic predictions

The following generic model was fitted to estimate the effect of markers on the trait Y:

$$\mathbf{y = \mu + Xg + e}, \quad (1)$$

where y is the vector of adjusted phenotypes, μ is an overall mean, X is a matrix of genotypes for p SNP loci (whose elements are indicator variables denoting number of copies of allele 1), g is a vector of SNP marker effects and e is a vector of random residual terms. It is worth to emphasize that the adjusted

For all traits and sampling strategies, the following statistical procedures were employed to predict the phenotypes of the animals in VAL set:

- *RR-GBLUP* [1]: shrinkage method in which markers were treated as random effects, by solving mixed model equations defined in (1) considering the variance ratios calculated with REML estimates of residual variance (σ^2_e) and additive genetic variance (σ^2_u), obtained in a previous step.

Under these assumptions, the direct solution for equation (1) would be obtained as:

$$\hat{\mathbf{g}} = (\mathbf{X'X}+\lambda\mathbf{I})^{-1} (\mathbf{X'y}) \quad (2),$$

where $\lambda = \sigma^2_e / (\sigma^2_u / k)$, $k = 2 \sum p_i - (1-p_i)$ and p_i is the allelic frequency of the i^{th} marker, as in [13], what reflects the fact that more polymorphic loci contribute more to the genetic variation.

In the present study, we employed an alternative method to solve (1) based on the SVD decomposition of **X** (i.e. X = UDV' = RV'), as proposed in [14]. These authors showed that identical solutions to those in (2) can be obtained by:

$$\hat{\mathbf{g}} = \mathbf{V} (\mathbf{R'R}+\lambda\mathbf{I})^{-1}(\mathbf{R'y}) \quad (3),$$

what could be computationally advantageous when **p** >> **n**.

-*emBayesB* : this procedure consists in a BayesB-like method implemented using the Expectation-Maximization algorithm proposed by [15]. A mixture distribution is assumed for marker effects - a proportion γ of them have effects drawn from a double exponential distribution, while the remainder effects are drawn from a Dirac Delta (DD) function, which has all its probability mass at 0. In the present study, the parameter γ was also estimated from the data.

- *SS_BY*: this method implemented subset selection through a two-step procedure. First step was carried out to select markers with significant effects on **y** through single-marker regression. The correction proposed by Benjamini & Yekutieli [16] was used to adjust p-values for multiple comparison (markers were selected using α = 1%). This procedure is often employed to control the false-discovery rate under *dependence* assumptions. In the second step, simultaneous estimation of the **s** selected markers was done similarly as in (2), by fitting them as random effects.

- *SS_ABS*: marker effects estimated with RR-GBLUP method were screened and those loci with larger contribution to the genetic variance (mean ± 1.5 SD) were selected. The variance at each locus was calculated as $2p_i (1-p_i) \hat{g}_i^2$, where p_i is the

allelic frequency and ĝ the estimated effect for the ith locus. In the second step, simultaneous estimation of the selected markers was done similarly as in (2).

-*RKHS*: Reproducing Kernel Hilbert Spaces regression using a Gaussian kernel was carried out by fitting the following model:

$$\mathbf{y = \mu + K_h\alpha + e},$$

under the assumption of the following prior distributions $\boldsymbol{\alpha}$ ~ N(0, $\mathbf{K_h}\sigma_\alpha^2$) and **e** ~ N(0,**I** σ_e^2). The entries of the kernel matrix $\mathbf{K_h}$ were defined as:

$$K_h(x_i, x_j) = \exp(-h\, d_{ij}),$$

where the d_{ij} the squared Euclidean distance between individuals i and j calculated based on their genotypes for SNP markers and the smoothing parameter **h** was defined as $h = 2/d^*$ and d^* is the mean of d_{ij}. This method was implemented in a Bayesian framework by using a Gibbs sampler, similarly as described by [17].

-*SVR*: Support vector regression was implemented using a radial basis kernel. Briefly, this method employs linear models to map (implicitly) the data to a higher-dimensional space via a kernel function. As discussed in [18], one feature of this method is to minimize a cost function that simultaneously includes model complexity and error in the training data. The regularization parameter was set to 1 as well as the default values of the tuning parameters of the function svm (R package 'e1071') were adopted.

-*BayesCpi*: By following notation from the equation (1), this method postulates a mixture model for marker effects such that the elements of vector **Xg** were calculated for each animal as $\sum_{j=1}^{N}(x_j a_j I_j)$, where x_j is the genotype of the jth marker, coded as the number of copies of one allele, a_j is the effect of marker j and I_j is an indicator variable that assumes the value of 1 whether the jth marker has any effect on the trait or 0, otherwise.

It was assumed that a_j ~ N(0, σ^2_a) and e ~ N (0,Iσ^2_e). Inverted scaled chi-squared distributions were postulated for σ^2_a and σ^2_e as described in [19]. A binomial distribution with probability (1-π) was assumed for I_j and an uniform prior was assigned for π. This model was implemented using a Gibbs sampler, such that a single chain of 50,000 iterations was simulated, the first 5,000 being discarded as burn-in. Notice that, unlike in BayesB method [1], this mixture model assumes that

marker effects are sampled from the same (normal) distribution, instead of estimating marker-specific variances.

-*BayesC*: a similar model to that described for Bayes Cpi was fitted, differing of that by the fact that the parameter π was kept fixed at 0.90.

- *LASSO* [20]: this method can be understood as a shrunken version of least squares estimates, obtained after minimizing the residual sum of squares subject to the restriction that L1-norm of $\hat{\mathbf{g}}$ (i.e. sum of the absolute value of marker effects) must be ≤ **t**. The threshold **t** was defined by means of internal cross-validation (10-fold).

-*RF*: the Random Forest algorithm [21] was applied in a regression framework, by assuming the matrix **X** as predictor of the phenotypes in **y**. A random forest of 1000 trees was built and this model was used to predict observations of VAL set.

Implementation

All analyses were performed using the R software [22]. In order to avoid the direct inversion of large matrices, the GSRU algorithm [23] was employed to solve iteratively the linear systems in RR-GBLUP, SS_BY, SS_ABS and emBayesB. To speed up computations, the implementations for RR-GBLUP, SS_BY, SS_ABS, emBayesB, BayesCpi and BayesC were compiled in C++ language, by using Rcpp package. The method RKHS was implemented using the R code provided by [17]. The other methods were implemented using specific R packages: e1071 (SVR), glmnet (LASSO) and randomForest (RF). REML estimates of variance components were obtained using ASREML-R package [24]. All the analyses were performed on a workstation with a Intel i7-2600 3.40GHz processor and 8GB RAM.

Analyses of results

All methods were compared based on their predictive ability($r_{y,\hat{y}}$), calculated as the Pearson's correlation between the phenotypes of each animal in the VAL set and the respective predicted values ($\hat{\mathbf{y}}$). This statistic was also computed for a situation in which only information of pedigree and phenotypes was considered in a BLUP model (polygenic model, POL), such that gains in predictive ability due to the consideration of genotypic information could be evaluated. For POL, the predicted values of observations were EBVs of VAL animals, obtained when considering exclusively the phenotypic information on animals in the REF set.

Significant differences between methods in terms of predictive ability were assessed by means of paired t tests (α = 5%), adjusted by Bonferroni correction.

The bias of prediction of each method was measured by the average prediction error, while the trend of inflation was measured by the slope of the regression of the observed phenotypes (**y**) on their predicted values ($\hat{\mathbf{y}}$). Mean squared error (MSE) was employed as a measure of the overall fit achieved with each method. As a general rule, values for bias (inflation) close to zero (close to 1) indicate better performance. As the phenotypes for each trait are in different scales, MSE was normalized (NRMSE). NRMSE was computed as the root mean-squared error divided by the range of the observed values. Values close to zero for NRMSE are associated with better overall fit.

Averages and standard errors (SE) were computed for each statistic by considering the results of the ten replicates available in each situation. The computing times required for the implementation of each method were also monitored and compared. In the case of the methods which explicitly estimate marker effects (i.e. apart from RF, SVR and RKHS), the distributions of marker effects were also examined and compared.

The accuracy of RR-GBLUP was calculated as its predictive ability divided by the square-root of the heritability of each trait ([25]) and then compared with the expected value for this statistic($r_{g,\hat{g}}$), derived according to the formula in Daetwyler et al. (2010):

$$r(g,\widehat{g}) = \sqrt{\left(\frac{Nh^2}{Nh^2+Me}\right)},$$

where N is the (average) size of the reference set, h^2 is the (pseudo)heritability of the trait and Me is the number of independent chromosome segments, calculated as *Me = 2NeL/ln(4NeL)* or *Me = 2NeL* ([26]). *L* is the length of the genome in Morgans and Ne is the effective population size (calculated in present study based on the estimates of r^2 between SNP markers). The values of h^2 considered in the formula accounted for the fact that phenotypes were adjusted for the effect of cage.

Variation in accuracy across genetic groups

Heslot et al. [5] verified that large differences in accuracy between subpopulations could not be explained only by differences in phenotypic variance and sample size. Although the definition of subpopulations is not so obvious in the

present study, it would be reasonable to investigate differences in accuracy of prediction between the unrelated families comprising the mice dataset. Because family sizes are not large enough to enable calculation of predictive ability within each of such families, we investigated this question by clustering the individuals into groups according to the genetic distance between them.

For this, a hierarchical clustering algorithm (Ward's method) was applied to a matrix of genetic distances calculated based on the genomic relationship matrix between the animals, in order to identify non-trivial partitions of the data. The Calinski-Harabaz statistic was employed to find the optimal number of clusters and after this procedure, the solution obtained with Ward's method was refined using k-means algorithm.

For both within-family and across-family splitting, predictive abilities were calculated within each one of the genetic groups obtained through clustering, for each combination of method, trait and replicate. A Fligner-Killeen test was applied to assess homogeneity of phenotypic variances across groups, such that we could investigate whether eventual differences in predictive ability between groups could be related to differences in phenotypic variances.

In order to test for differences in within-group predictive ability, a Fisher's z transformation was applied over predictive abilities, since these are computed as Pearson's correlations and thus their sampling distributions are not normal. Then, for each replicate, equality of predictive ability across groups was assessed using a chi-square test, after which p-values were averaged across replicates.

Results

Variance components

REML estimates of variance components are presented for each trait in Table 3. Estimated heritabilities for W6W, WGS, BL, %CD8+ and CD4+/CD8+ matched well the previous estimates published by Valdar et al. ([11]) and presented in Table 1. The estimates for W6W, WGS and BL were in agreement to those obtained by [6], being that the largest difference was observed for body length, whose heritability was 7% lower in the present study.

Table 3 - REML estimates of variance components (and related parameters) for traits of a heterogeneous stock mice population

Trait	σ^2_u	SE	σ^2_c	SE	σ^2_e	SE	h^2	SE
W6W	3.915	29.836	1.719	13.100	3.E-05	9.E-04	0.695	0.030
WGS	8.E-04	2.E-04	1.E-03	1.E-04	9.E-04	1.E-04	0.295	0.069
BL	0.036	0.012	0.039	0.007	0.148	0.009	0.161	0.051
%CD8	19.370	2.851	1.990	0.471	0.357	1.505	0.892	0.101
CD4+/CD8+	5.E-03	7.E-04	6.E-04	1.E-04	4.E-04	4.E-04	0.825	0.081

Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+).
σ^2_u: additive genetic variance ; $\sigma^2_{c:}$ variance due the random environmental effect of cage; σ^2_e: residual variance; h^2: heritability (standard error, SE, in brackets).

Within-family predictions

In Figure 1, results of predictive ability under within-family splitting are presented for all methods, grouped by trait, as well as the results obtained when considering only pedigree and phenotypic information (i.e. using the polygenic model, POL). The polygenic model achieved predictive abilities about 0.56, 0.30, 0.15, 0.61 and 0.52 for W6W, WGS, BL, %CD8+ and CD4+/ CD8+, respectively.

For a same trait, some methods had comparable predictive abilities, although significant differences between methods could be found. For all traits, at least two of the methods (SVR and RKHS) reached greater predictive abilities than POL (Figure 1).

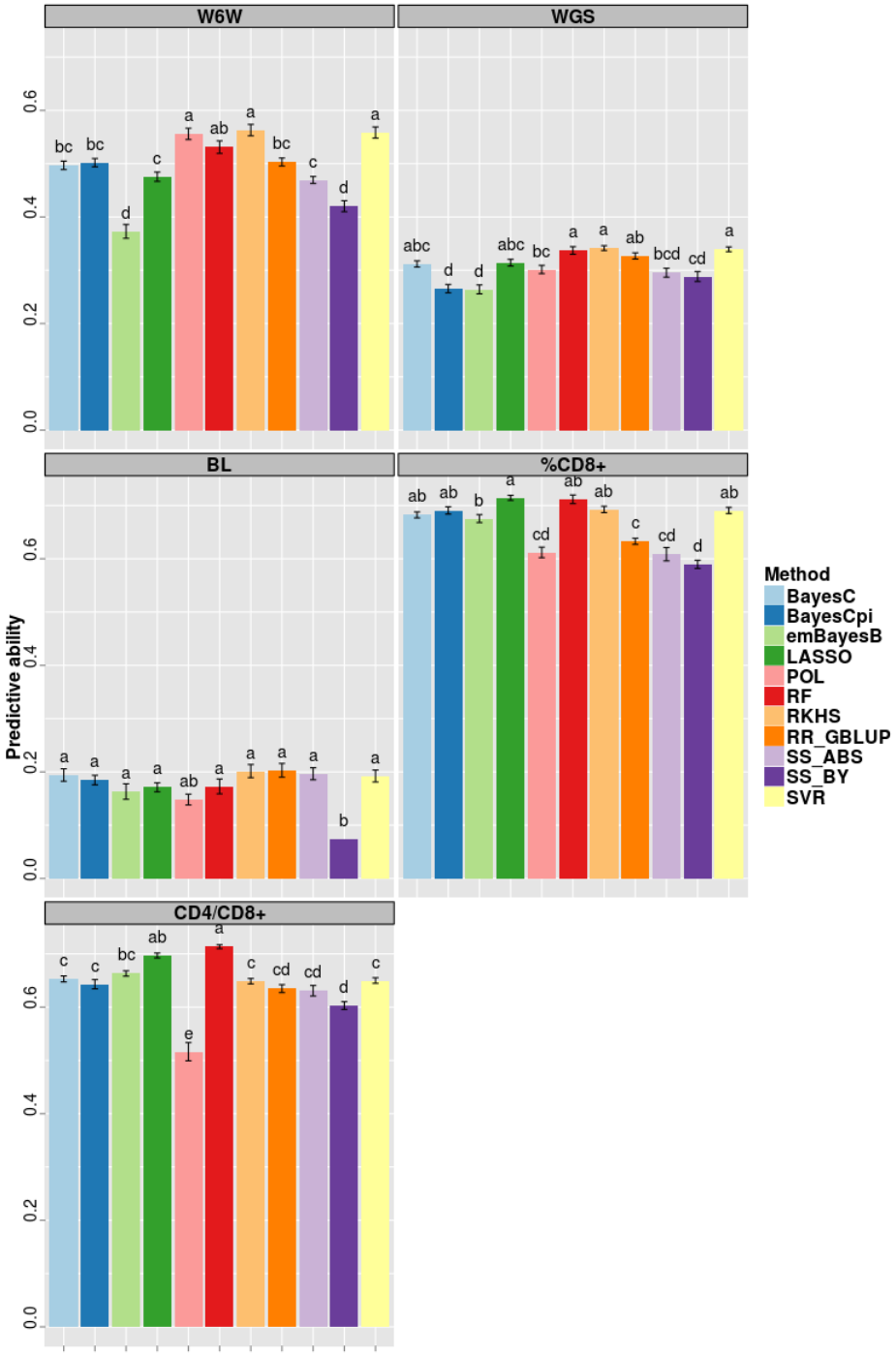


Figure 1 - Predictive ability* of the different methods employed in within-family predictions for five traits in a mice population.

*Average of ten replicates. Bars sharing the same letter are not different (P >0.05, pairwise t test adjusted with Bonferroni correction).

Traits: weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+).

The relative performance of the methods varied noticeably across traits. RKHS, POL and SVR (in this order) were the most accurate for W6W, while RKHS, SVR and RF outperformed the remainder methods with respect to the predictions for WGS. Predictions for BL did not differ greatly across methods, except by the worst performance of SS_BY. LASSO and RF were the two with greater predictive abilities for %CD8+ and CD4+/ CD8+. As a general rule, the methods based in some sort of variable selection (especially LASSO and emBayesB) had better performance in the case of %CD8+ and CD4+/ CD8+ compared to the other traits.

Overall, the subset selection methods (SS_ABS and SS_BY) did not rank among the best methods for none of the traits studied. It is important to mention that for BL, the significance threshold applied in SS_BY was possibly too stringent, since that in only one of the ten replicates significant markers were found, reason why error bars for predictive ability and fitting statistics are not presented in this situation.

Because methods with assumptions close to RR-GBLUP are among the most used in practical applications of genomic selection, it is meaningful to assess the additional gain in predictive ability that can be reached by methods with different assumptions. In present study, predictive ability of RR-GBLUP figured among the highest in the case of predictions for BL and WGS. For the remainder traits, the most accurate methods reached predictive abilities between 12% and 13% greater than RR-GBLUP.

An additional set of analyses was carried out by considering a smaller MAF threshold (1%) for genotypes, aiming to investigate whether lower frequency variants could be important for some of the traits under investigation. As a general rule, predictive abilities of the two sets of analyses did not differ by more than 0.5%, being that the largest increase (2.9%) was observed for WGS when using BayesCpi (data not shown).

Across-family predictions

It must be noted that across-family predictions using method POL are expected to have accuracy of zero, because the pedigree information do not include links between animals in REF and VAL sets, although the predictive ability cannot be explicitly computed in this case, since the SD of the predicted values for VAL set is zero.

For the remainder methods, predictive ability was consistently lower in across-family predictions (Figure 2) compared to within-family predictions (Figure 1). Across methods, the greatest decreases in predictive ability relative to within-family predictions were observed for W6W(66%), WGS (44%) and BL (41%), for which predictive abilities reached figures about 0.20 at most and no significant differences between methods were found.

For %CD8+ and CD4+/ CD8+, predictive ability averaged across methods was about 0.50 and thus about 22% lower than in within-family splitting. RF and LASSO (in this order) had the highest predictive abilities for both traits (Figure 2), although other methods reached comparable predictive ability for %CD8+. For these traits, the advantage over RR-GBLUP was more pronounced when compared to the results obtained for within-family predictions. For example, the most accurate method (RF) reached predictive abilities about 43% (%CD8+) and 58% (CD4+/ CD8+) greater than RR-GBLUP (Figure 2).

Bias, inflation and overall fit

Since the phenotypes for each trait are in different scales, the averages of bias are presented as proportions of the respective phenotypic SD in Table 4. Except for W6W and regardless of the splitting strategy, the largest amount of bias were observed for BayesC and BayesCpi, for which there was a trend of overestimation in the case WGS and BL, while the predictions for the other traits were underestimated. Predictions for LASSO and emBayesB were considerably underestimated in the case of W6W, %CD8+ and CD4+/ CD8+ and overestimated for WGS (Table 4). Overall, the less biased predictions were obtained with RF, RKHS, SVR and RR-GBLUP.

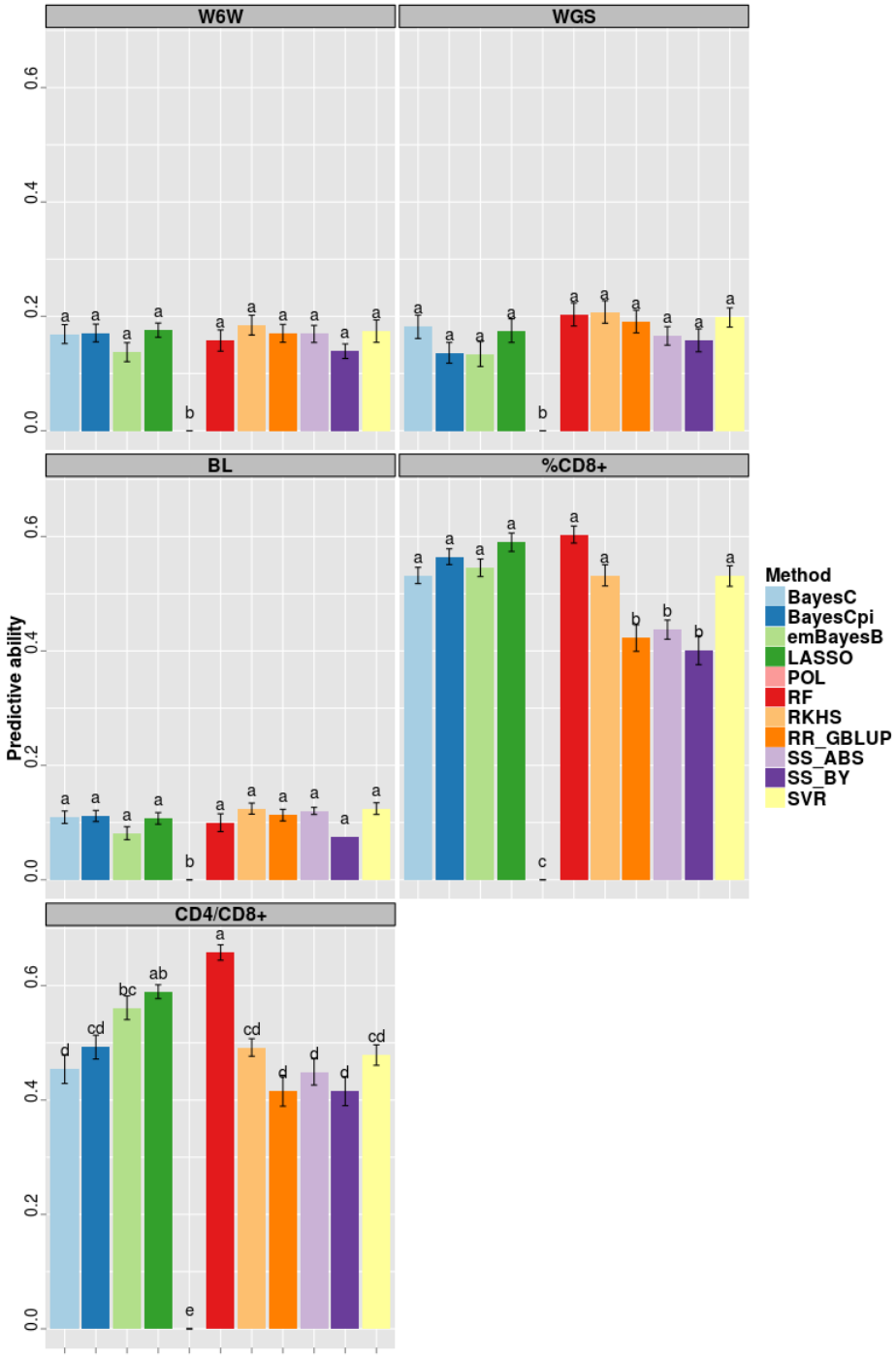


Figure 2 - Predictive ability* of the different methods employed in across-family predictions for five traits in a mice population.

*Average of ten replicates. Bars sharing the same letter are not different (P >0.05, pairwise t test adjusted with Bonferroni correction).

Traits: weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+).

Table 4- Bias* of genomic predictions from different methods, obtained for five traits of a mice population

Splitting	Method	Trait				
		W6W	WGS	BL	%CD8	CD4+/CD8+
Within	POL	-4%	1%	-1%	1%	1%
	emBayesB	-62%	24%	2%	-42%	-52%
	RR-GBLUP	-9%	8%	13%	2%	4%
	SS_BY	-18%	22%	-	5%	5%
	SS_ABS	-35%	4%	-30%	5%	17%
	RKHS	-1%	1%	-1%	1%	2%
	SVR	-4%	3%	-6%	5%	4%
	BayesCpi	-26%	66%	81%	-568%	-91%
	BayesC	-43%	263%	69%	-397%	-420%
	LASSO	-85%	53%	2%	-73%	-73%
	RF	-1%	1%	1%	0%	1%
Across	POL	-	-	-	-	-
	emBayesB	-51%	5%	3%	-58%	-87%
	RR-GBLUP	-31%	8%	8%	-2%	14%
	SS_BY	-44%	18%	-	-1%	17%
	SS_ABS	-54%	-7%	-38%	-16%	31%
	RKHS	-23%	4%	-4%	9%	10%
	SVR	-24%	6%	-8%	14%	10%
	BayesCpi	-15%	83%	46%	-719%	-196%
	BayesC	23%	280%	115%	-239%	-359%
	LASSO	-116%	40%	7%	-70%	-98%
	RF	-13%	2%	-1%	5%	5%

*Average of ten replicates. Bias was measured as the average difference between observed and predicted phenotypes of testing set and is the presented as a proportion of the standard deviation of each trait (in %). Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Splitting= splitting strategy in cross-validation (within or across-family).

The methods under investigation also differed greatly in terms of the inflation of genomic predictions, being that BayesCpi and BayesC were those producing the most inflated genomic predictions for all traits, followed by SS_BY and SS_ABS, while emBayesB was the only method which consistently resulted in deflation of genomic predictions, under within-family splitting (Table 5). Across traits, LASSO and RKHS had the coefficients of inflation closest to 1 for within-family predictions, followed by SVR, RF and RR-GBLUP. In most of the situations, the coefficients of inflation for across-family predictions showed greater deviation from 1.

Table 5- Inflation* of genomic predictions from different methods, obtained for five traits of a mice population

Splitting	Method	Trait				
		W6W	WGS	BL	%CD8	CD4+/CD8+
Within	POL	1.21	0.91	0.89	1.03	0.97
	emBayesB	1.43	1.57	3.70	1.65	1.55
	RR-GBLUP	0.95	0.70	0.88	0.66	0.75
	SS_BY	0.79	0.66	-	0.60	0.72
	SS_ABS	0.72	0.49	0.49	0.64	0.72
	RKHS	1.08	0.95	0.89	1.06	1.26
	SVR	1.24	0.91	0.68	1.16	1.18
	BayesCpi	0.55	0.29	0.40	0.17	0.29
	BayesC	0.18	0.12	0.21	0.19	0.20
	LASSO	0.93	1.01	1.08	1.00	1.07
	RF	1.48	1.10	0.78	1.12	1.11
Across	POL	-	-	-	-	-
	emBayesB	0.46	0.90	1.47	1.43	1.33
	RR-GBLUP	0.37	0.49	0.52	0.46	0.52
	SS_BY	0.30	0.46	-	0.42	0.52
	SS_ABS	0.29	0.30	0.32	0.50	0.52
	RKHS	0.55	0.95	0.75	1.21	1.36
	SVR	0.61	0.89	0.62	1.42	1.27
	BayesCpi	0.22	0.21	0.14	0.17	0.24
	BayesC	0.07	0.08	0.11	0.18	0.15
	LASSO	0.38	0.71	0.56	0.98	1.01
	RF	0.72	1.22	0.65	1.27	1.22

*Average of ten replicates. Inflation was measured as the slope of the regression of observed phenotypes on predicted phenotypes of testing set.
Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Splitting= splitting strategy in cross-validation (within or across-family).

In terms of overall fit, measured by the normalized root-mean squared error, BayesC and BayesCpi were again those with worst performance, what is expected judged by their performance in terms inflation and mainly bias, which are accounted for by MSE (Table 6). By averaging NRMSE across methods, it can be noticed that predictions for %CD8+ and CD4+/ CD8+ showed greater values for NRMSE than those of the other traits.

By considering the overall fit (Table 6), it can be noted a more consistent ranking of methods when compared to the results for predictive ability (Figures 1 and 2). In the case of within-family predictions, RKHS was the best method for W6W, WGS and BL, while RF was the best for the other two traits. Typically, RF, RKHS and SVR were the best three methods in terms of overall fit what was also observed in the case of across-family predictions.

Table 6- Normalized root-mean squared error(NRMSE)* of genomic predictions from different methods, obtained for five traits of a mice population

Splitting	Method	Trait				
		W6W	WGS	BL	%CD8	CD4+/CD8+
Within	POL	0.099	0.114	0.123	0.138	0.145
	emBayesB	0.140	0.122	0.124	0.165	0.165
	RR-GBLUP	0.103	0.114	0.123	0.147	0.136
	SS_BY	0.112	0.122	-	0.158	0.144
	SS_ABS	0.117	0.120	0.132	0.161	0.143
	RKHS	0.098	0.112	0.122	0.127	0.131
	SVR	0.099	0.112	0.123	0.128	0.130
	BayesCpi	0.166	0.214	0.268	1.252	0.502
	BayesC	0.534	0.877	0.251	1.035	1.080
	LASSO	0.172	0.137	0.127	0.204	0.196
	RF	0.102	0.112	0.123	0.124	0.119
Across	POL	-	-	-	-	-
	emBayesB	0.154	0.132	0.126	0.199	0.230
	RR-GBLUP	0.128	0.121	0.125	0.184	0.170
	SS_BY	0.141	0.126	-	0.193	0.175
	SS_ABS	0.145	0.130	0.140	0.188	0.180
	RKHS	0.122	0.119	0.124	0.152	0.152
	SVR	0.122	0.119	0.125	0.155	0.153
	BayesCpi	0.163	0.243	0.240	1.519	0.593
	BayesC	0.637	0.747	0.275	0.896	1.091
	LASSO	0.201	0.142	0.129	0.213	0.225
	RF	0.119	0.119	0.124	0.144	0.130

*Average of ten replicates. Lower values are associated with better overall fit .
Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Splitting= splitting strategy in cross-validation (within or across-family).

Distribution of marker effects

The variation in excess kurtosis for estimated marker effects distribution of different traits could indicate that a method is able to fit marker effect distribution to the QTL distributions of such traits. Results for this statistic are presented in Table 7. As a general rule, although the magnitude of excess kurtosis of effect distribution differed greatly among methods for a same trait, estimates of this statistic were reasonably consistent for a same method. The only exception was found in the case of BayesCpi, for which the marker effect distribution was close to a normal distribution for W6W, while more peaked distributions were found for the other traits. For the sake of brevity, results for across-family splitting are not presented, as the findings were very similar to those observed under within-family splitting.

A further inspection on the distribution of estimated marker effects, showed that, for the variable selection methods, a given proportion of the markers contributed

a smaller proportion of the total genetic variance accounted by the markers in W6W compared to the other traits (Table 7). The estimates of the proportion of markers with effect on each trait obtained using BayesCpi, by averaging the posterior means of (1-π) across replicates, were 59.0% (W6W), 0.1% (WGS), 11% (BL), 2.1% (%CD8+) and 0.4% (CD4 +/ CD8+) and also suggested a more polygenic control on W6W.

Table 7- Summary statistics* associated with distributions of estimated marker effects (within-family splitting)

		Method					
Trait	Stat*	BayesCpi	emBayesB	RR-GBLUP	LASSO	SS_ABS	SS_BY
BL	t1000	0.83	1.00	0.52	1.00	1.00	1.00
	t500	0.75	1.00	0.35	1.00	0.99	0.97
	t100	0.53	0.99	0.12	0.97	0.52	0.76
	t20	0.34	0.96	0.04	0.68	0.18	0.43
	|g| > 0	9457	4076	9820	199	623	541
	kurt	284.0	2650.1	0.6	496.3	37.7	503.7
CD4+/ CD8+	t1000	1.00	1.00	0.53	1.00	1.00	0.96
	t500	0.99	1.00	0.37	1.00	0.99	0.88
	t100	0.92	0.99	0.13	0.97	0.57	0.60
	t20	0.64	0.97	0.04	0.71	0.21	0.32
	|g| > 0	9559	3938	9820	228	577	1172
	kurt	407.2	2790.4	1.0	597.7	48.6	319.9
%CD8	t1000	0.91	1.00	0.53	1.00	1.00	0.98
	t500	0.83	1.00	0.36	1.00	0.99	0.93
	t100	0.50	0.99	0.13	0.97	0.54	0.69
	t20	0.21	0.96	0.04	0.71	0.19	0.39
	|g| > 0	9820	3792	9820	203	603	754
	kurt	36.1	3009.2	0.7	553.6	41.2	443.6
W6W	t1000	0.51	1.00	0.52	1.00	1.00	1.00
	t500	0.35	1.00	0.36	1.00	0.99	1.00
	t100	0.12	1.00	0.12	0.93	0.54	0.85
	t20	0.04	1.00	0.04	0.49	0.19	0.46
	|g| > 0	9820	4711	9820	325	612	456
	kurt	0.7	1094.8	0.6	218.0	39.9	155.5
WGS	t1000	1.00	1.00	0.52	1.00	1.00	1.00
	t500	1.00	1.00	0.35	1.00	0.99	0.98
	t100	1.00	1.00	0.12	0.95	0.54	0.77
	t20	0.93	1.00	0.04	0.58	0.19	0.40
	|g| > 0	1716	4623	9820	260	617	655
	kurt	877.0	1998.6	0.6	310.9	40.1	134.7

*Average of 10 replicates. t1000, t500, t100 and t20 = proportion of the variance accounted for the markers (varM) explained by those with the largest 1000, 500, 100 and 20 absolute effects, respectively. varM = $\Sigma_i [\hat{g}_i^2\ 2\ p_i\ (1\text{-}p_i)]$, in which $\hat{g}_i$ and p_i are the estimated effect and the allele frequency for the i^{th} marker, respectively. |g| > 0 = number of markers with non-null estimated effect. kurt = excess kurtosis of the distribution of estimated marker effects. Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+).

Computing time

The elapsed time to perform model training was of order of minutes for all methods investigated and the ranking of the methods for this criterion was consistent across traits. The more demanding methods were BayesCpi, BayesC and RF (in this order), whose computing times were between 5-fold and 131-fold larger than those required by the other methods (Figure 3). emBayesB and RR-GBLUP required the lowest computing time. Differences between methods in terms of computing time also depends on code optimization. Although all methods were implemented using compiled code, it is worth to mention that differences between them can be also due to implementation-specific aspects.

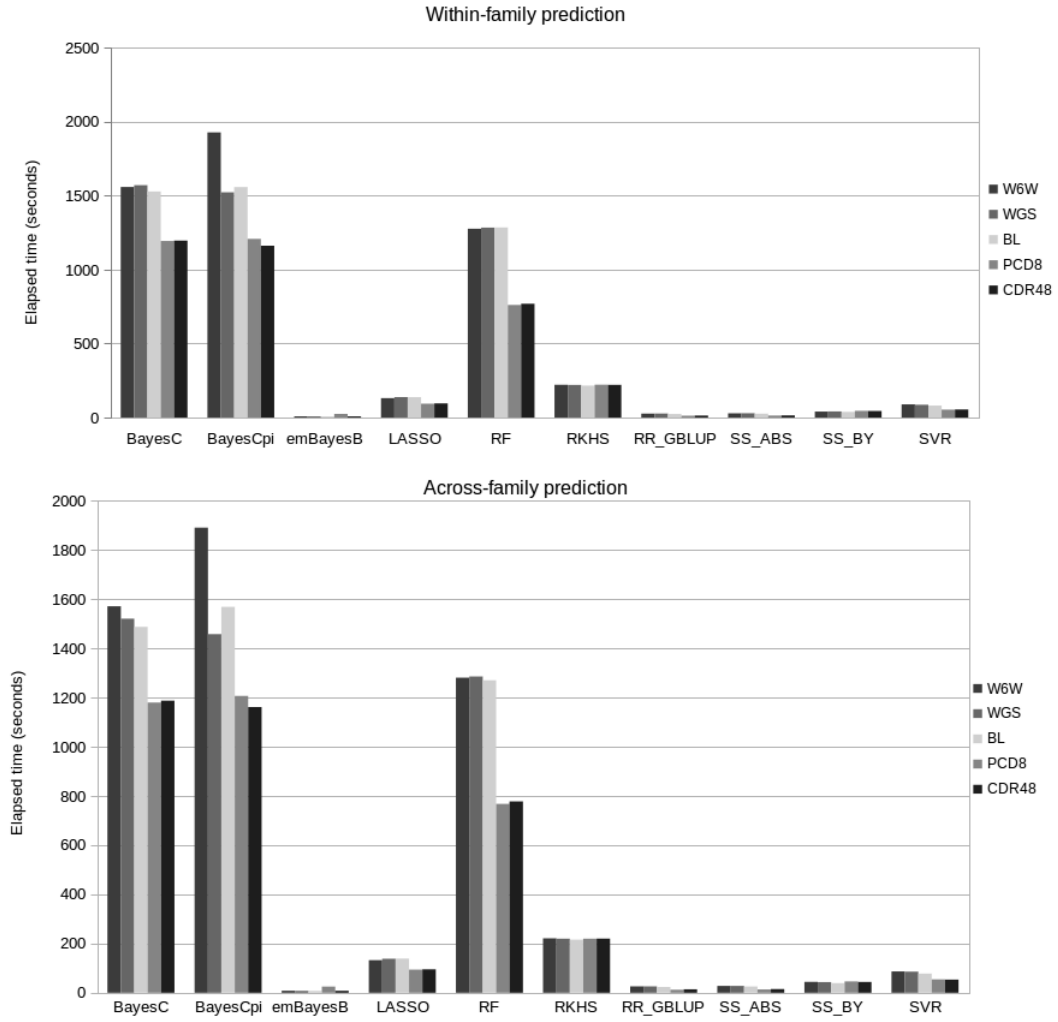


Figure 3 - Average computing time* required to perform model training using different statistical methods.

*Average of ten replicates. Elapsed times were measured during model training, carried out with the information available in the reference set, aiming to compute genomic predictions for five traits in a mice population.

Expected and realized accuracy of RR-GBLUP

In Figure 4, expected accuracies of RR-GBLUP are presented for the two approximations of the number of independent chromosome segments (Me1 and Me2), besides the realized values obtained in each situation. It can be seen that the way Me was approximated impacted heavily on expected accuracies. As a general rule, the realized accuracies of RR-GBLUP matched better the expected values in the case of within-family predictions and the expected values computed assuming *Me=2NeL* fitted best to the realized values.

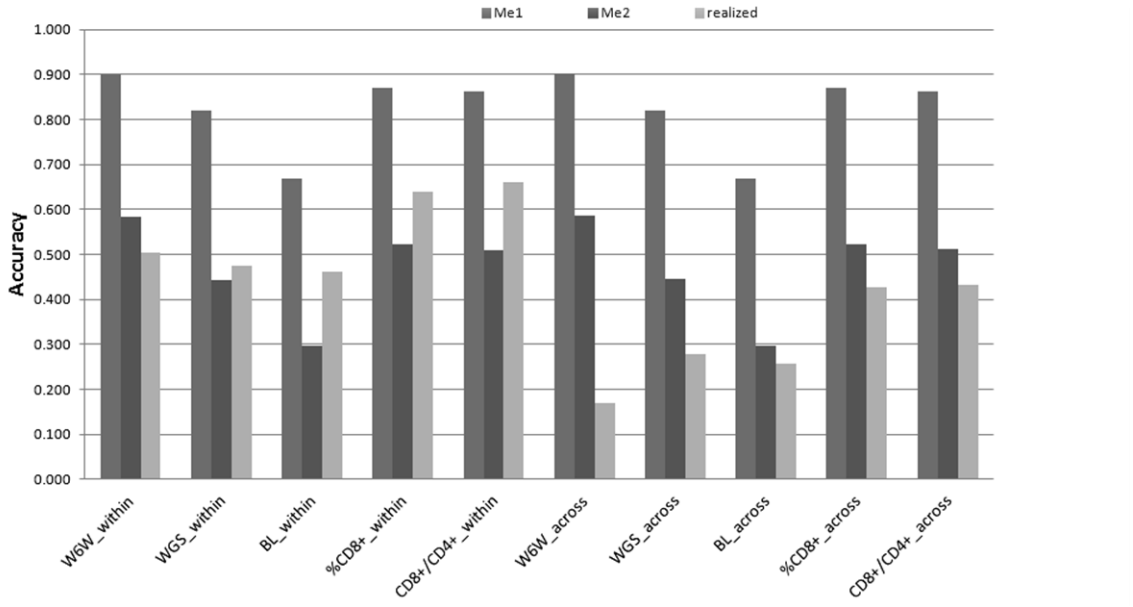


Figure 4 - Expected and realized accuracy of genomic predictions with RR-GBLUP.

Expected accuracies were calculated according to Daetwyler et al. (2010), by considering two approximations for the number of independent chromosome segments (Me): *Me1 = 2NeL/ln(4NeL)* or *Me2 = 2NeL*. realized = realized accuracy of GBLUP (average of ten replicates).

Five traits were considered: weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Two scenarios of prediction were considered within-family (_within) and across-family predictions (_across).

Variation in predictive ability across genetic groups

According to the Calinski-Harabaz statistic, the optimal number of groups was found to be four (data not shown). There was evidence of differences between these groups for phenotypic variance in the case of %CD8+ (P=0.00031), CD4+/CD8+

(P=0.00034) and WGS (P=0.04137), while significant differences were not found for the other traits (P>0.20).

Although the chi-square test for equality of within-group predictive abilities takes in account differences in sample size, there was remarkable differences in the p-values obtained across replicates for a same trait and method, such that, when differences were averaged, they did not provide strong evidence against the null hypothesis in most of the situations analyzed (Table 8).

Table 8. Summary of the results* of the test for equality of predictive ability across groups

Splitting	Within-family					Across-family				
Trait	W6W	WGS	BL	%CD8+	CD4+/ CD8+	W6W	WGS	BL	%CD8+	CD4+/ CD8+
emBayesB	0.42	0.30	0.26	0.28	0.05	0.22	0.47	0.33	0.06	0.03
RR-GBLUP	0.35	0.41	0.27	0.19	0.11	0.30	0.36	0.28	0.10	0.09
SS_BY	0.32	0.36	0.16	0.13	0.07	0.40	0.30	0.20	0.09	0.03
SS_ABS	0.39	0.34	0.28	0.04	0.17	0.20	0.28	0.34	0.06	0.06
RKHS	0.38	0.33	0.33	0.24	0.06	0.30	0.41	0.28	0.28	0.06
SVR	0.41	0.34	0.33	0.25	0.08	0.28	0.39	0.35	0.30	0.05
BayesCpi	0.37	0.28	0.27	0.28	0.07	0.31	0.38	0.25	0.12	0.07
BayesC	0.34	0.44	0.25	0.27	0.08	0.29	0.37	0.26	0.19	0.06
LASSO	0.45	0.29	0.24	0.24	0.10	0.30	0.36	0.38	0.10	0.03
RF	0.40	0.35	0.32	0.16	0.07	0.29	0.62	0.27	0.16	0.13

*P-values for the chi-square test for equality of predictive ability across groups (average of 10 replicates). In order to reduce the influence of discrepant replicates, the average of log10(p-value) was calculated for each trait and method, and then back-transformed to the original scale. Animals were clustered into 4 groups based on genetic distance between them.
Trait = weight at 6 weeks (W6W), weight growth slope (WGS), body length (BL), percentage of CD8+ cells (%CD8+), ratio between CD4+ and CD8+ cells (CD4+/CD8+). Splitting= splitting strategy in cross-validation (within or across-family).

Despite this, significant results (or at least suggestive) of differences in predictive abilities between groups were found in the case of CD4+/CD8+ and %CD8+, being that the magnitude of such differences also varied across methods. As a general rule, for these traits, stronger evidence for differences between groups in predictive ability was found under across-family splitting compared to within-family splitting.

Discussion

Previous studies with the mice dataset

As mentioned earlier, the mice dataset was also object of previous studies. [6] and [8] analyzed W6W, WGS and BL, using methods that are analogous to RR-GBLUP and LASSO, respectively, while [7] analyzed %CD8+ through a Bayesian variable selection method (RJMCMC) and thus somewhat analogous to the variable selection methods of this study (especially BayesCpi). Despite this, it is important to mention that results of these different studies are not fully comparable, due to the influence of data editing and random splitting procedures and even due to differences in method implementation.

For within-family predictions, as a general rule, predictive abilities were between 19% and 40% lower in the present study than in comparable situations reported by [6] and [8] what can be attributed mainly to differences in the way the cage effects were modeled. In present study, the phenotypes were adjusted for cage effects, while in that studies such effects were fitted simultaneously to marker effects and thus contributed to predictive ability. An additional argument in favor of this hypothesis is that similar differences were also observed when comparing model POL to an analogous model fitted by [6].

On the other hand, for %CD8+, emBayesB, BayesCpi and LASSO achieved predictive abilities 4%, 6% and 10% greater, respectively, than that obtained with reversible-jump Markov chain Monte Carlo (RJMCMC) under within-family prediction [7]. These authors also verified that, when dominance effects on this trait were fitted simultaneously to additive effects, predictive ability increased by about 6% and 10% for within-family and across-family prediction, respectively.

Comparison between statistical methods in genomic prediction.

While simulation studies suggest that variable selection methods (e.g. BayesB, LASSO) could outperform methods with assumptions close to those of GBLUP (e.g. [1], [4], [27]), the performance of RR-GBLUP has often been comparable to that of variable selection methods when real data were analyzed (e.g. [18], [28], [5]) . A possible explanation for the apparent divergence of results with simulated and real data could be that, for real data, the distribution of QTL effects for most traits is less extreme than has been simulated, as suggested by [3] and [29].

In the present study, methods with large conceptual differences reached very similar predictive abilities in some situations and a clear re-ranking of methods was observed in function of the trait analyzed. In other species, it has also been verified that methods with different assumptions had similar performance (e.g. [28], [30],[5]) , while method by trait interaction often takes place (e.g. [31,32]).

Overall, for at least one of the five traits analyzed in present study, most of the methods figured among the most accurate, being that the only exception was observed for SS_BY and SS_ABS, reason why these methods do not seem to be recommendable for application in genomic selection.

One important point that needs to be taken into account when comparing methods employed in genomic selection is that their performance is often influenced by key parameters required in their implementations, like the assignment of prior distributions in Bayesian regression methods and setting tuning parameters of machine learning methods. In this way, such parameters can be optimized for each situation analyzed in order to improve predictive performance.

Although we expected that the implementations of the LASSO (tuned through internal cross-validation) and emBayes (in which parameters related to the distribution of marker effects are estimated from the data) also could fit well to QTL distribution, results of excess kurtosis suggest that statistical learning was more effective in the case of BayesCpi, although some drawbacks were found for this method, especially in terms of bias.

Even though RKHS, SVR and RF figured among the best methods in some of the situations analyzed, some of their tuning parameters were not optimized for each trait, due to the additional computational effort that would be required by this task, in a way that their performance might still be improved.

According to [33], predictive ability (or accuracy) is currently the main statistic employed to compare genomic prediction methods. However, bias and inflation of genomic predictions should be matter of concern, especially if animals from different generations and with different amounts of information (e.g. progeny-tested and newborn animals) are among selection candidates. If predictions are biased upwards, genetic trend will be overestimated, benefiting newborn animals unduly, while inflation would exaggerate differences among predicted values compared to the

true differences, also having negative impact in selection schemes. Judged by their overall performance in terms of bias and inflation, RF, RKHS, SVR and RR-GBLUP outperformed the remainder methods.

It is also important to mention that in the present study the phenotypes used in model training were pre-corrected for environmental effects in order to reduce computing times, as well as the target phenotypes in the validation sets, what is not an optimal approach from a statistical perspective and could be an additional source of noise in the present results. For instance, in the case of overcorrection for some fixed effect in the validation set, it would be expected that methods that lead to more shrunk estimates of marker effects perform better in terms of bias and inflation in this situation.

Specifically with respect to SVR, this method had the lower overall predictive ability among the ten methods studied by [5] when considering eight different datasets of crop species, even after optimizing the model for each trait. In present study, SVR figured among the best methods for W6W and WGS, and had very similar predictive performance to RKHS, what can be explained by the similar kernel definition in both methods. One possible explanation for the worse performance of SVR in [5] could be the fact that SVR was fitted using a linear kernel while a radial basis kernel was employed in this study.

It has been suggested that the predictive ability of BayesCpi is not strongly affected by the starting value of π [19] and even by the lack of convergence in this parameter [34], although our results may suggest the need for further investigation of the impact of these factors with regard to bias and inflation of genomic predictions. An additional set of analyses was carried out to investigate the issue of convergence in π, by simulating two independent chains with different starting values for this parameter (0.90 and 0.10) and using the Gelman and Rubin's convergence test. The results varied across traits and also across replicates (data not shown), being that most of the replicates did not converge in the case of W6W and BL. For all traits, the averages of the posterior means of π across replicates were not strongly affected by the starting values, although the considerable variation in the estimates across replicates may suggest lack of information in the data to estimate π properly. The

poorer results with Bayes C could be justified by the misspecification of the value for π.

Heslot et al.[5] also pointed out that BayesCpi should not be recommended for application in genomic selection because it achieved very similar predictive ability to that of RR-GBLUP at a much greater computational cost. Conversely, BayesCpi presented some advantage over RR-GBLUP in the case of %CD8+ in the present study as well as in two of the 17 traits analyzed by [30]. Such disagreement demonstrates how difficult is to take a broad view on the relative performance of different methods and reinforces the hypothesis of interplay between relative performance of methods and genetic background.

The optimal method for genomic selection should be reliable across traits and computationally efficient, besides, obviously, being the highest accurate possible and less prone to overfitting [5]. Some authors have also alerted to the fact that methods relying more on LD between markers and QTL would be preferable to those whose accuracy result basically from the genetic relationships captured by the markers [27], because in this last case the accuracies are expected to decrease considerably in generations subsequent to estimation of marker effects. On the other hand, in situations of continuous updating of training populations and re-estimation of prediction equations, as typically occurs in dairy cattle (e.g. [35]), this not seems to be a major issue.

Given the imminent increase in dimensionality of genomic selection problems, due to both increase in the number of genotyped animals and especially in the density of marker panels [25], it is mandatory to take computing requirements in consideration when comparing statistical methods. Judged by their overall performance across traits and computational requirements, RR-GBLUP, RKHS and SVR seem to be particularly appealing. Although these methods have some conceptual differences, in both of them, problem dimensionality is reduced to the number of genotyped animals, what gives significant computational advantage over other methods when p >> n.

Genetic architecture and predictive performance.

It seems to be consensual that accuracy of genomic predictions is dependent of the genetic architecture of traits (number of underlying QTL, mode of inheritance),

as well as of the size of the training set, the number of independent chromosome segments (which is function of genome size and effective population size), the heritability of (pseudo)phenotypes used to train models and the marker density (e.g. [3],[25]).

Regarding to the relative performance of the prediction methods, Daetwyler et al. [3] suggested that the accuracy of RR-GBLUP is invariant to number of QTL affecting the trait (N_{QTL}), while the accuracy of methods based on variable selection is expected to be greater than that of RR-GBLUP when N_{QTL} is lower than the number of independent chromosome segments .

In the present study, the predictive abilities were considerably greater in the case of within-family predictions when compared to across-family predictions, what was also reported by [6] and [7]. Despite this, this superiority of predictive ability for within-family predictions varied markedly across traits and this pattern was consistent across methods, what could suggest that predictive abilities for some traits are more dependent on close relationships than the others.

Another hypothesis for the differences in the superiority of within-family predictions across traits, also mentioned by [7], is that resemblance between relatives, due to shared common environment (not properly accounted for in the model), could inflate predictive abilities and thus traits more influenced by such common environmental effects would exhibit larger superiority for within-family over across-family prediction.

Especially for W6W, the relatively good performance of POL and the different pattern for the distribution of estimated marker effects in variable selection methods could be indicative of a more polygenic background than the other traits.

In the other extreme, based in the previous results regarding QTL mapping of these traits [10] and given the superior predictive ability of variable selection methods for %CD8+ and CD4+/CD8+, one could expect that these traits are mostly influenced by a smaller number of QTL. In addition, there is some previous evidence that %CD8+ is influenced by dominance effects [7], although the two most accurate methods for this trait (LASSO and RF) do not explicitly take such effects into account.

It has been advocated RKHS, SVR and RF are able to capture complex interactions (e.g. dominance and epistatic effects), in a way that could be

hypothesized that part of their predictive ability could be due to such interactions, although it can be difficult to confirm this theory, given the difficult to model interactions explicitly. It is worth to emphasize that in this case these methods predict genotypic values rather than purely additive breeding values.

Within-family predictions, as defined in this study, are more similar to the situation expected in most of animal breeding applications [6], because selection candidates are expected to be related at some degree with animals of reference populations.

Another important question regarding to the predictive ability of genomic prediction methods regards to its variation across subpopulations or families. In the present study, there was some evidence that such differences in predictive ability are trait-specific, being that for the two traits for which suggestive differences between groups in predictive ability were found under some methods, phenotypic variances also differed between these groups. Thus, conversely to what was pointed out by [5], there was no evidence in the present study to reject the hypothesis that differences in the within-group predictive ability could be explained by differences in phenotypic variance.

According to [5], groups with larger phenotypic variance would have larger genetic variance and larger influence in model training, what would lead to higher predictive ability within these groups. A possible explanation for such differences could be sampling, since predictive ability is expected to be higher for groups with larger number of individuals in the training set, although the use of 10 replicates is expected to ensure that all groups were properly represented in training set. In this way, this question still deserve further investigation in order to confirm the existence of such differences and to investigate their origin.

Some authors argue that genotyping by whole-genome resequencing will become a regular practice in the near future ([25,36]). As the number of SNPs increases, the assumption that many of them do not have effect on a trait is more likely to be true ([37]), what would be a typical scenario in which variable selection methods would be preferable. Despite this, the first studies on analyses of full sequences did not signalize advantage of BayesB over RR-GBLUP ([25]), although more research is needed to answer this question properly. For these reasons,

comprehensive comparisons of statistical methods are still appealing in genomic selection, while further developments in terms of computational efficiency will be required to deal with problems of increasing dimensionality and complexity. More powerful tools will be probably needed to extend current models to account for pleiotropy (multi-trait predictions), non-additive genetic effects and genotype-by-environment interactions.

Conclusions

Methods with large conceptual differences reached very similar predictive abilities in some situations and a clear re-ranking of methods was observed in function of the trait analyzed. For all traits and situations analyzed, at least two of the genomic prediction methods (SVR and RKHS) led to more accurate predictions than the polygenic model.

Variable selection methods were more accurate than the remainder in the case of %CD8+ and CD4+/CD8+ and these traits are likely to be influenced by a smaller number of QTL than the remainder. Judged by their overall performance across traits and computational requirements, RR-GBLUP, RKHS and SVR are particularly appealing for application in genomic selection.

References

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

2. Jannink J-L, Lorenz AJ, Iwata H: **Genomic selection in plant breeding: from theory to practice.** *Brief Funct Genomics* 2010, **9**: 166-177.

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

4. Clark SA, Hickey JM, van der Werf JHJ**: Different models of genetic variation and their effect on genomic evaluation.** *Genet Sel Evol* 2011, **43**:18.

5. Heslot N, Yang H-P, Sorrells ME, Jannink J-L: **Genomic selection in plant breeding: A comparison of models.** *Crop Sci* 2012, **52**:146-160.

6. Legarra A, Robert-Granie C, Manfredi E, Elsen J-M: **Performance of genomic selection in mice.** *Genetics* 2008, **180**: 611–618.

7. Lee SH, Van Der Werf JHJ, Hayes BJ, Goddard ME, Visscher PM: Predicting unobserved phenotypes for complex traits from whole-genome SNP data. *PLoS Genet* 2008, 4:e100023.

8. Usai MG, Goddard ME, Hayes BJ: **LASSO with cross-validation for genomic selection.** *Genet Res* 2009, **91**: 427-436.

9. Laurie CC, Nickerson DA, Anderson AD, Weir BS, Livingston RJ, Dean MD, Smith KL, Schadt EE, Nachman MW: Linkage disequilibrium in wild mice. *PLoS Genet* 2007, 3:e144.

10. Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO, Taylor MS, Rawlins JN, Mott R, Flint J: **Genome-wide genetic association of complex traits in heterogeneous stock mice.** *Nat Genet* 2006, **38**: 879-887.

11. Valdar W, Solberg LC, Gauguier D, Cookson WO, Rawlins JN, Mott R, Flint J: **Genetic and environmental effects on complex traits in mice.** *Genetics* 2006, **174**: 959-984.

12. Moser G, Khatkar MS, Raadsma HW: **Imputation of missing genotypes in high density SNP data.** In *Conference Of The Association For The Advancement Of Animal Breeding And Genetics*: 2009, Barrosa Valley.

13. VanRaden P: **Efficient methods to compute genomic predictions.** *J Dairy Sci* 2008, **91**:4414 – 4423.

14. Hastie T, Tibshirani R: **Efficient quadratic regularization for expression arrays.** *Biostatistics* 2004, **5**: 329–340.

15. Shepherd RK, Meuwissen THE, Wooliams, JA: **Genomic selection and complex trait prediction using a fast EM algorithm applied to genome-wide markers.** *BMC Bioinformatics* 2010, **11**: 529.

16. Benjamini Y, Yekutieli D: **The control of the False Discovery Rate in multiple testing under dependency.** *Ann Statist* 2001, **29**:1165-1188.

17. Crossa J, de los Campos G, Pérez P, Gianola D, Burgueño J, Araus JL, Makumbi D, Dreisigacker S, Yan J, Arief V, Banziger M, Braun H-J: **Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers.** *Genetics* 2010, **186**:713–724.

18. Moser G, Tier B, Crump, RE, Khatkar MS, Raadsma HW: **A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers.** *Genet Sel Evol* 2009, **41**:56.

19. Habier D, Fernando RL, Kizilkaya K, Garrick DJ: **Extension of the bayesian alphabet for genomic selection.** *BMC Bioinformatics* 2011, **12**: 186.

20. Tibshirani R: **Regression shrinkage and selection via the lasso**. *J R Stat Soc B* 1996, **58**: 267-288.

21. Breiman L: **Random Forests**. *Machine Learning* 2001, **45**: 5-32.

22. R Development Core Team: **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria, 2012. [http://www.r-project.org].

23. Legarra A, Misztal I: **Technical note: Computing strategies in genome-wide selection.** *J Dairy Sci* 2008, **91**: 360-366.

24. Butler DG, Cullis BR, Gilmour AR, Gogel BJ: ASREML-R Reference Manual. Release 3.0. VSN International Ltd, Hemel Hempstead; 2009. OpenURL20.

25. Ober U, Ayroles JF, Stone EA, Richards S, Zhu D, Gibbs RA, Stricker C, Gianola D, Schlather M, Mackay TFC, Simianer H: **Using Whole-Genome Sequence Data to Predict Quantitative Trait Phenotypes in Drosophila melanogaster.** *PLOS Genet* 2012, **8**: e1002685.

26. Clark SA, Hickey JM, Daetwyler HD, van der Werf JHJ: **The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes.** *Gen Sel Evol* 2012, **44**:4.

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

28. Hayes BJ, Bowman PJ, Chamberlain AC, Klara Verbyla K, Goddard ME: **Accuracy of genomic breeding values in multi-breed dairy cattle populations.** *Gen Sel Evol* 2009, **41**:51.

29. Calus MPL: **Genomic breeding value prediction: methods and procedures.** *Animal* 2010, **4**: 157-164.

30. Resende Jr MF, Muñoz P, Resende MD, Garrick DJ, Fernando RL, Davis JM, Jokela EJ, Martin TA, Peter GF, Kirst M: **Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda L.).** *Genetics* 2012, **190:** 1503-1510**.**

31. Verbyla KL, Hayes BJ, Bowman PJ, Goddard ME: **Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle.** *Genet Res (Camb)* 2009, **91**: 307-311.

32. Hayes BJ: **Genomic Selection in the era of the $1000 genome sequence.** In *Symposium Statistical Genetics of Livestock for the Post-Genomic Era, 2009; Madison*. [http://dysci.wisc.edu/sglpge/pdf/Hayes.pdf]

33. Vitezica ZG, Aguilar I, Misztal I, Legarra A: **Bias in genomic predictions for populations under selection.** *Genet Res (Camb)* 2011, **93**:357-366.

34. Wolc A, Stricker C, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Habier D, Fernando R, Garrick DJ, Lamont SJ, Dekkers JCM: **Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model.** *Gen Sel Evol* 2011, **43**:5.

35. Wiggans GR, Sonstegard TS, Van Raden PM, Matukumalli LK, Schnabel RD, Taylor JF, Chesnais JP, Schenkel FS, Van Tassell CP: **Genomic Evaluations in the United States and Canada: A Collaboration.** In *International Committee on Animal Recording (ICAR),June 22-26, 2009; Munich*. ICAR Tech Ser. 13; 2009:347–353.

36. Meuwissen THE, Goddard ME: **Accurate prediction of genetic values for complex traits by whole-genome resequencing.** *Genetics* 2010, **185**: 623–631.

37. Goddard ME, Hayes BJ, Meuwissen THE: **Genomic selection in livestock populations.** *Genet Res (Camb)* 2010, **92**: 413-421.

CHAPTER 3 – Adequacy of using different pseudo-phenotypes for model training and validation of genomic predictions in a simulated beef cattle population

Abstract

Background

Genomic selection can enable accurate predictions of the genetic merit of selection candidates as soon as they are genotyped, provided that proper estimates of marker effects are available. Estimation of marker effects requires an adequate response variable (pseudo-phenotype) to summarize genetic information as well as efficient statistical methods to track the association between pseudo-phenotypes and genotypes of reference animals. This study was carried out to investigate the consequences of using three different pseudo-phenotypes for model training and validation of genomic predictions, obtained after applying different statistical methods for estimation of marker effects in a simulated beef cattle population.

Methods

A forward-in-time simulation framework was adopted to mimic a large beef cattle population. A forward prediction scheme was simulated in the recent generations, in which pseudo-phenotypes of proven bulls were used as response variables for model training, while information on young bulls did not contribute to estimate marker effects and was used for validation. Three different types of pseudo-phenotypes were considered: estimated breeding values (EBV), deregressed proofs (dEBVs) and progeny-yield deviations (PYD). Genomic predictions (GEBV) were obtained using four methods: two BLUP models based on a genomic relationship matrix, obtained after assigning weights of 20% (GBLUP20) or 0 (GBLUP0) for pedigree-based relationships, a variable selection method (LASSO) and a Bayesian regression model (BayesCpi). Further consequences of using different pseudo-phenotypes as prediction targets in validation were investigated, so that inferences from such procedure were contrasted with those obtained when using true breeding value as target.

Results

In the scenario with smaller amount of information to estimate marker effects, the relative performance of different methods was affected by the type of pseudo-phenotype employed as response variable. As a general result, using dEBV provided the most robust scenario for estimation of marker effects, both in terms of accuracy and scale of GEBV, while using such pseudo-phenotype as target predictand in validation also allowed more reasonable inferences than using EBV or PYD for the same purpose.

Conclusions

Deregressed proofs (dEBV) comprise a more suitable response variable to estimate marker effects than estimated breeding values (EBV) or progeny-yield deviations (PYD). When validating genomic predictions, proper inferences about the accuracy and scale of GEBV can be obtained using dEBV as target predictand. The accuracy of GEBV should be assessed by the Pearson's correlation between GEBV and dEBV divided by the average accuracy of dEBV in the validation set.

Background

Genomic selection (GS) has been considered a hot topic since the pioneer simulation study of Meuwissen et al. [1] highlighted that the genetic merit of selection candidates could be accurately predicted using only information from dense molecular marker panels. Nowadays, this information is routinely incorporated in some animal and plant commercial breeding programs.

While the cost-effectiveness of such technology was evident for dairy cattle breeding, its adoption in other fields is still limited. For instance, in beef cattle breeding programs, the different structure of the breeding schemes, the larger effective population sizes and the smaller generation intervals could result in more modest benefit through using genomic selection, compared to that obtained in dairy schemes, especially for the traits that are routinely recorded before selection decisions.

In this context, the accuracy delivered by genomic predictions plays an important role in the evaluation of the feasibility of GS schemes and it is directly associated to the precision with which the marker effects are estimated. Proper

estimation of marker effects would require an adequate response variable (hereafter, pseudo-phenotype) to summarize the genetic information on reference animals, as well as statistical methods that efficiently associate such response variable to marker information [2].

According to [3], the ideal response variable to estimate marker effects would be the true genetic merit (TBV) of the individuals observed on unrelated genotyped animals and in the absence of selection. Since TBV is unknown in real applications, different pseudo-phenotypes have been used to estimate marker effects, including estimated breeding values (EBV) (e.g. [4], [5], [6]), daughter-yield deviations (DYD) (or progeny-yield deviations, PYD) (e.g. [7], [8]) and deregressed proofs (dEBV) (e.g. [9], [10]).

The EBV, typically derived from BLUP, combines animal's own phenotypic records to those of its relatives, so that the smaller the amount of information, the more EBV is regressed towards the mean. The DYD (PYD) is a unregressed variable, obtained after adjustment of progeny performance for fixed effects, non-genetic random effects as well as for the genetic merit of mates [11]. The dEBV is a unregressed variable similar to DYD, but derived in a different manner, using only information from genetic evaluations (EBVs). While the dEBV can be obtained by the iterative procedure described by [12], a simpler procedure, proposed by Garrick et al. [3], can also be employed for this task.

Garrick et al. [3] highlighted important drawbacks of using EBV to estimate marker effects. Because EBV is a regressed variable, the contrast in EBV between genotypes at a given locus is shrunk relatively to the contrast verified if TBV or phenotypes were used as response variable. The extent of this shrinkage is proportional to the reliability (r^2) of each EBV and thus when r^2 varies among reference animals, using EBV as response would lead to heterogeneous residual variances as well. EBV includes ancestral information and its use to estimate marker effects can result in double-counting of information when relatives are included in the reference population. For genomic analyses, the authors of [3] suggested the use of an unregressed response variable (dEBV), obtained after parent average (PA) is removed and applying appropriate weighting.

There is no consensus on the most appropriate response variable for marker effects estimation. Garrick et al. [3], in a theoretical work, suggested that the relative benefit of using dEBVs would depend on the extent to which the amount of information used to predict EBV varies among reference animals. A simulation study [13] reported that using EBV as response resulted in slightly larger reliability of the GEBV, compared to using DYD. With real data, a study with pure-bred pigs found benefit for using dEBV as response variable instead of EBV [2], while other study with dairy cattle did not verify such advantage in GEBV reliability [14], even though pointed out that using EBV as response would result in problems in the scale of the genomic predictions.

One factor that hampers the comparison between studies about the use of different pseudo-phenotypes is that, while in simulations the TBV is the obvious predictand (the target to be predicted by GEBV), allowing a ideal scenario to evaluate the impact of response variables on accuracy of genomic predictions, different surrogate for TBV are employed to validate predictions with real data. As a general rule, the same type of variable used for model training is employed as predictand in the validation (e.g. [7], [8], [6]), but studies employing different pseudo-phenotypes for model training and model validation were also reported (e.g. [2], [14]).

In real applications, the correlation between GEBV and predictand in the validation set is considered as a proxy for GEBV accuracy. Because the reliability of predictands is often less than unity, it is a common practice to divide such correlation by an estimate of the accuracy of the predictand [15]. Besides the effect of different pseudo-phenotypes in marker effects estimation, it can be hypothesized that the type of variable used as predictand could influence the validation results. A simulation study could be helpful to address such question.

Gao et al. [14] compared different statistical methods for marker effect estimation, using either EBV or dEBV as response variable. Although the reliability achieved with different methods and response variables was relatively similar across the 17 traits analyzed, important changes in the scale of the GEBV obtained with different methods occurred depending on the response employed for model training. Such result could rise the hypothesis that the relative predictive performance of different methods could change depending on the pseudo-phenotype employed as

response variable. Because such conclusions were drawn from real data, a simulation study could be useful to investigate this question in more detail.

This study was carried out to compare the use of EBV, dEBV or PYD for model training and validation of genomic predictions obtained with different statistical methods for estimation of marker effects in a simulated beef cattle population.

Methods

Population structure. The QMSim software [16] was employed to generate a population for which the extent and pattern of linkage disequilibrium (LD) should be consistent with that verified in beef cattle populations. The simulations were carried out in a forward-in-time framework, so that the parameters related to the historical generations were defined similarly as in [17].

First, 1,000 historical generations were simulated, with effective population size (Ne) kept constant at the value of 1,000, followed by 1,020 historical generations in which Ne was gradually reduced from 1,000 to 200, such that mutation-drift equilibrium was established and initial LD was generated. In both steps, the number of individuals of each sex was the same and the mating system was based on random union of gametes.

After the simulation of these 2,020 historical generations, the population was expanded by randomly selecting 100 founder males and 100 founder females from the last generation of the historical population. Eight generations were simulated with five offspring per dam and an exponential growth of the number of dams, also under random union of gametes and without selection.

At the end of this process, 640 sires and 32,000 dams were randomly sampled from the last generation of the expanded population, from which 10 recent generations were simulated. This step reproduced a selected beef cattle population, with one offspring per dam and about 50% progeny of each sex. The animals were culled and selected based on their EBVs for a polygenic trait (with heritability of 0.10) and randomly mated after that. This simulated polygenic trait was uncorrelated with the traits for which genomic predictions were obtained (described later) and was intended to generate some linkage disequilibrium between markers due to selection in the recent generations. Replacement ratios were 60% and 20% for sires and

dams, respectively. In this way, a reduction in the Ne of recent generations should be expected as a consequence of selection, what has been verified with real data (e.g. [18], [19]). At the end of this process, the recent population comprised 352,640 animals, including sires, dams and their offspring, which was intended to mimic the structure and size of a large beef cattle population.

Simulated genome. The simulated genome consisted of 29 pairs of autosomes, with length varying from 40 cM to 146 cM, thus identical to the real bovine genome based on Btau_3.1 assembling [20] and totaling 2333 cM. The simulation of such a large genome, despite the higher computational requirements, intended to reproduce a more realistic scenario regarding the number of physically unlinked marker and QTL loci [17]. Values of 1% and 0.5% were assumed for the rates of missing marker genotypes and marker genotyping errors, respectively, and a recurrent mutation rate of 10^{-4} for both markers and QTLs was considered to approach mutation-drift equilibrium in historical generations. No mutation was generated in the recent generations [16].

Markers were evenly distributed along the genome and the initial number of markers was defined such that it allowed to obtain the desired number of segregating bi-allelic loci in the last historical generation. In this generation, 40,000 segregating loci, with minor allele frequency (MAF) greater than 0.05 were selected to mimic the genotypic information, so that the marker density approached that observed after the quality control of genotypes for 50K bovine SNP chips (e.g. [18], [21]).

In order to reduce computing requirements in the simulation process, for each of the 10 replicates simulated for each scenario (described later), 2,000 loci were selected as potential QTL loci and removed from the marker panel, so that the remainder 38,000 loci were assumed to be SNP markers available for use in genomic predictions.

Simulated traits. Two different traits were simulated, each one characterized by a level of heritability and sex in which phenotypes were measured (Table 1).

Both traits were influenced by 1000 QTL, so that the QTL positions were randomly assigned (among the potential QTL loci) and their effects were drawn from a gamma distribution with shape equal to 0.40, what is justified by the fact that such distribution was found to be a reasonable approximation for the QTL previously

mapped in cattle [22]. The scale of QTL effects was defined such that the sum of the variance explained by 5 loci randomly assigned (major QTL) was equal to 50% of the additive variance. Each locus had 50% of probability of having positive or negative effect.

Table 1 - Definition of the simulated traits

Trait	Heritability(h^2)	QTL effects*	MQTL	Vm (%)	Sex
A	0.30	gamma (shape = 0.4)	5	50	M/F
B	0.10	gamma (shape = 0.4)	5	50	F

*QTL effects were drawn from a gamma distribution with shape equal to 0.40. MQTL = Number of loci for which larger QTL effects were simulated (i.e. major QTL). Vm = proportion of the additive variance explained by the major QTL (in %). The simulated effects for major QTL were re-scaled so that the variance explained by each loci was the same for all of them. Sex = sex in which phenotypes were available: male (M), female (F).

The trait A mimicked moderately heritable selection criteria, expressed in both sexes, representing selection criteria similar to growth-related traits. Trait B mimicked a low-heritable selection criteria, only expressed in females, represent criteria similar to female fertility or reproductive longevity.

The simulation of such a extreme scenarios for the QTL distribution (5 QTL explaining 50% of the additive variance) could magnify differences between prediction methods in terms of accuracy, what was opportune to assess the power to detect such differences under different validation strategies (described later).

The true breeding values (TBV) of the 352,640 animals of the recent generations were simulated as:

$$TBV = \sum q_i \, v_i,$$

in which q_i and v_i are, respectively, the genotype (coded as the number of copies of the reference allele) and the substitution allelic effect for the i^{th} loci. Thus, only additive effects were simulated in all situations. For each trait, a continuous phenotype Y was simulated for each animal by adding an overall mean ($\mu = 10$) and a random residual term (ε) to its respective TBV. Residuals were drawn from a normal distribution, $\varepsilon \sim N(0, \sigma^2_\varepsilon)$, and σ^2_ε was defined so that the target heritability was met in the first generation simulated, by considering the additive variance in the

founder population which originated the recent generations in each scenario and replicate.

Pseudo-phenotypes and study design. The study design mimicked a situation in which only progeny-tested bulls were genotyped, in the same way as previously applied in studies in dairy and beef cattle populations (e.g. [7], [9], [10]). A forward prediction scheme was simulated. The older bulls (generations 0 to 7) comprised the training set used to estimate marker effects (reference population, N=3,328), while younger bulls (generations 8 and 9) were included in the set used for validation of the genomic predictions (validation population, N=768).

Three different types of pseudo-phenotypes were generated for each genotyped bull and employed as response variable to estimate marker effects: 1) estimated breeding value (EBV), 2) deregressed EBV(dEBV) and 3) twice the progeny-yield deviation (PYD).

A “traditional” genetic evaluation was simulated for each replicate and scenario, in which EBV and their respective reliabilities (based on the estimated prediction error variance) were obtained after fitting an animal model. The pedigree and the phenotypic information available in the recent generations were considered in such model, fitted using the BLUPF90 software [23].

The estimates of EBVs and associated reliabilities of genotyped bulls as well as those of their respective sires and dams were employed to compute the dEBV following [3]. The PYD was obtained as an average of adjusted progeny records, following [11]. Because phenotypes were expressed in animals of both sexes for one of the simulated traits (Table 1), the more generic term 'progeny-yield deviation' (PYD) was adopted to describe this pseudo-phenotype, instead of daughter-yield deviation (DYD).

For proper validation of genomic predictions, it is recommended to avoid data overlapping between the reference and validation sets ([24], [25]). For this reason, separate evaluations were carried out to obtain pseudo-phenotypes for model training and validation, in a way that pseudo-phenotypes of training bulls did not contain any phenotypic information on validation animals or on their descendants, while pseudo-phenotypes of validation animals included all available information until the 10^{th} recent generation.

Genomic predictions. It was assumed that only the sires from the recent generations were genotyped. Because missing markers and genotyping errors were simulated, the genotypes were submitted to a quality control procedure, such that only markers with MAF >2%, call rate > 98% and p-value for Hardy-Weinberg equilibrium test > 10^{-5} were considered, as well as the samples with call rate > 90%. After this process, genotypes for 4,096 bulls (3,328 of them included in the training set) were available for 37,686 bi-allelic markers. Then, for each marker, missing genotypes were replaced with its respective mean.

The following statistical methods were used to compute genomic predictions:

-GBLUP: The model for this method can be represented as:

$$\mathbf{y}=\mathbf{1}_n\boldsymbol{\mu} + \mathbf{Zg} + \mathbf{e}, \quad (1)$$

in which **y** is the vector of pseudo-phenotypes (EBV, dEBV or PYD) for each trait and scenario, **μ** is an overall mean, $\mathbf{1}_n$ is a vector of 1's, **Z** is an incidence matrix relating the vector of additive breeding values to **y**, **g** is a vector of additive breeding values and **e** is a vector of random residual terms. It was assumed that $\mathbf{g} \sim N(0, \mathbf{G}^*\sigma^2_g)$ and $\mathbf{e} \sim N(0, \mathbf{D}\sigma^2_e)$, where **G*** is a relationship matrix and **D** is a diagonal matrix, whose elements were defined in order to account for the differences in the reliabilities of the pseudo-phenotypes in **y,** as in [26]. In this way, the k^{th} diagonal entry of D (D_{kk}) was defined as: $D_{kk} = (1/R^2_k) - 1$, where R^2_k is the reliability of the pseudo-phenotype of the k^{th} animal. The **G*** matrix was obtained as $\mathbf{G}^*=(1-w)\mathbf{G}+w\mathbf{A}$, where **G** is the genomic relationship matrix and **A** is the numerator relationship matrix (based on pedigree information), both of order equal to the number of genotyped animals, and w is the proportion of the additive variance not captured by the markers. **G** was calculated as $\mathbf{G}=\mathbf{MM}'/\Sigma 2p_i(1-p_i)$, in which **M** is an incidence matrix for marker effects, whose elements were set to $0-2p_i$, $1-2p_i$ e $2-2p_i$ depending on if the genotype contained 0, 1 or 2 copies of the reference allele, whose frequency was p_i among the genotyped animals [26].

The genomic predictions were obtained using the software gebv version 1.0 [27]. In the computation of **G***, setting a value different of 0 for w is equivalent to fit residual polygenic effects that are not captured by the markers [28]. After testing different values for w (ranging from 0 to 0.40), Gao et al. [28] found that w=0.20 provided the best compromise in terms of reliability and scale of GEBV. In the

present study, GBLUP predictions were obtained either assuming w=0 (**GBLUP0**) or w=0.20 (**GBLUP20**).

-LASSO: This method can be understood as a penalized version of the least squares solutions, which minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant **t** [29]. In the present study, the tuning parameter **t** was defined by an internal validation procedure based on the Cp statistic [30]. Similarly as in GBLUP, differences in reliabilities of the pseudo-phenotypes used for model training were taken into account via weighted analyses. This method was implemented using the estimates obtained with the package 'glmnet' of the R software, version 2.15.0 [31].

-Bayes Cpi (Cpi) : This Bayesian regression model has some similarity with Bayes B method [1], in the sense that both methods fit only a portion of the markers in the model and thus can be considered variable selection methods. By following the same notation from the equation (1), Bayes Cpi postulates a mixture model for marker effects such that the elements of vector **Zg** (i.e. the vector of breeding values) were calculated for each animal as:

$$\sum_{j=1}^{N} x_j a_j l_j,$$

where x_j is the genotype of the j^{th} marker, coded as the number of copies of one allele, a_j is the effect of marker j and l_j is an indicator variable that assumes the value of 1 whether the j^{th} marker has any effect on the trait or 0, otherwise. It was assumed that $a_j \sim N(0, \sigma^2_a)$ and $\boldsymbol{e} \sim N(0, \mathbf{D}\sigma^2_e)$, where **D** is a diagonal matrix, whose elements were defined similarly as described previously for GBLUP**.** Inverted scaled chi-squared prior distributions were assigned for σ^2_a and σ^2_e, as described in [32]. A binomial distribution with probability (1-π) was assumed for l_j and an uniform prior was assigned for π.

Note that, unlike in BayesB method [1], this mixture model assumes that the marker effects are sampled from the same (normal) distribution, instead of estimating marker-specific variances. Another important difference is that the proportion π is estimated from the data, while this parameter is often treated as known in Bayes B. Thus, the choice for testing Bayes Cpi in the present study was justified by the fact that this tries to address some of the drawbacks often associated to BayesB [33].

Bayes Cpi was applied using the software GS3 [34] to obtain samples of the marginal posterior distribution of the parameters through an MCMC algorithm, carried out for 100,000 iterations (considering a burn-in period of 20,000 iterations and a thin interval of 100 iterations).

Analyses of results. Different model training strategies were defined by each combination of pseudo-phenotype and statistical method employed to estimate marker effects. For each strategy, genomic predictions (GEBV) were obtained for all genotyped animals as the sum of the estimated marker effects, or equivalently, by solving the mixed model equations in the case of GBLUP.

The results were analyzed assuming two different validation scenarios: 1) the TBV of validation animals was the predictand (i.e. the target to be predicted by GEBV); 2) the predictand was of the same type of pseudo-phenotype employed in model training (EBV, dEBV or PYD).

Under validation scenario 1, the consequences of using different statistical methods and/or pseudo-phenotypes for marker effects estimation were assessed according to different statistics (described below). Because TBV was used as predictand, this would be the ideal validation scenario, that would allow proper inferences about the relative benefit of different model training strategies under investigation.

Using pseudo-phenotypes as predictands (validation scenario 2) would mimic criteria that are available to compare model training strategies in a real application. It could be argued that some results of validation scenario 2 could be anticipated from conclusions of previous studies, e.g. using EBV for model training should result in scaled down GEBV [13]. On the other hand, validation results with real data can deviate markedly from such expectations (e.g. [14]). Thus, the comparison between validation scenarios 1 and 2 would underline situations where the use of a given pseudo-phenotype as predictand could lead to improper inferences about model training strategies, possibly helping to identify more robust predictands for validation.

The following statistics were employed to evaluate genomic predictions: i) the Pearson's correlation between predictand and GEBV in the validation set was employed to evaluate the accuracy of the genomic predictions (acc). ii) the slope of the regression of predictand on GEBV was employed to evaluate the scale

(inflation/deflation) of the genomic predictions (b1). The smaller the deviation of the estimated b1 from 1, the smaller the scale differences between predictand and GEBV. iii) the normalized root-mean squared error (NRMSE) was used to evaluate the overall fit. The mean squared error (MSE) quantifies both squared bias and variance of the estimator. Because pseudo-phenotypes used as predictands in validation scenario 2 have different reliability (thus being expected to have different scales) the MSE was normalized to provide a more comparable measure of overall fit among scenarios. Smaller values of NRMSE are associated to better fit.

For validation scenario 2, alternative estimators of acc and b1 were obtained using procedures intended to adjust for differences in signal-to-noise ratios of the predictands [15]. Two alternative estimators of accuracy were obtained: acc_adj was obtained after dividing acc by the square-root of the average reliability of predictand in the validation set, as in [28]; acc_w was obtained as a weighted Pearson product moment correlation, as in [35], so that weights (w_p) for each data point were obtained as a function of predictand reliability, similarly as described previously. The alternative estimate of b1 (b1_w) was obtained after fitting a weighted linear regression of predictand on GEBV, using w_p as weights, similarly as in [28].

In addition, for GBLUP predictions, the estimated prediction error variance (based on the inverse of the coefficient matrix) was employed to compute individual accuracies of GEBV (rPEV) for each validation animal, similarly as in [36]. The average of rPEV in the validation set was considered as a proxy for genomic prediction accuracy and thus compared to the other estimators of GEBV accuracy described earlier.

The results were compared based on the averages (and standard errors, SE) of the statistics related to accuracy, inflation and overall fit, obtained with the 10 replicates in each scenario.

Results

Linkage disequilibrium and summary statistics. The average (SD) linkage disequilibrium (LD) between adjacent markers over the 29 simulated chromosomes, measured by the r^2 statistic, was 0.203 (0.229), while the median of r^2 was 0.113. The pattern of LD decay as a function of the distance was also consistent with that observed in beef cattle breeds (Supplementary Figure 1). It was verified a reduction

of the estimates of Ne calculated based on the estimates of r^2 along the recent generations, decaying from 170 (generation 1) to about 105 in the 10^{th} generation, with an average of 145 over the ten recent generations.

The true accuracies of the pseudo-phenotypes considered in the model training and validation are presented in Figure 1, for each trait simulated. Notice that, for each trait, method and replicate, the same set of pseudo-phenotypes was employed, namely the 3,328 sires from generations 0 to 7 as training set and the 768 sires from generations 8 and 9 as validation set.

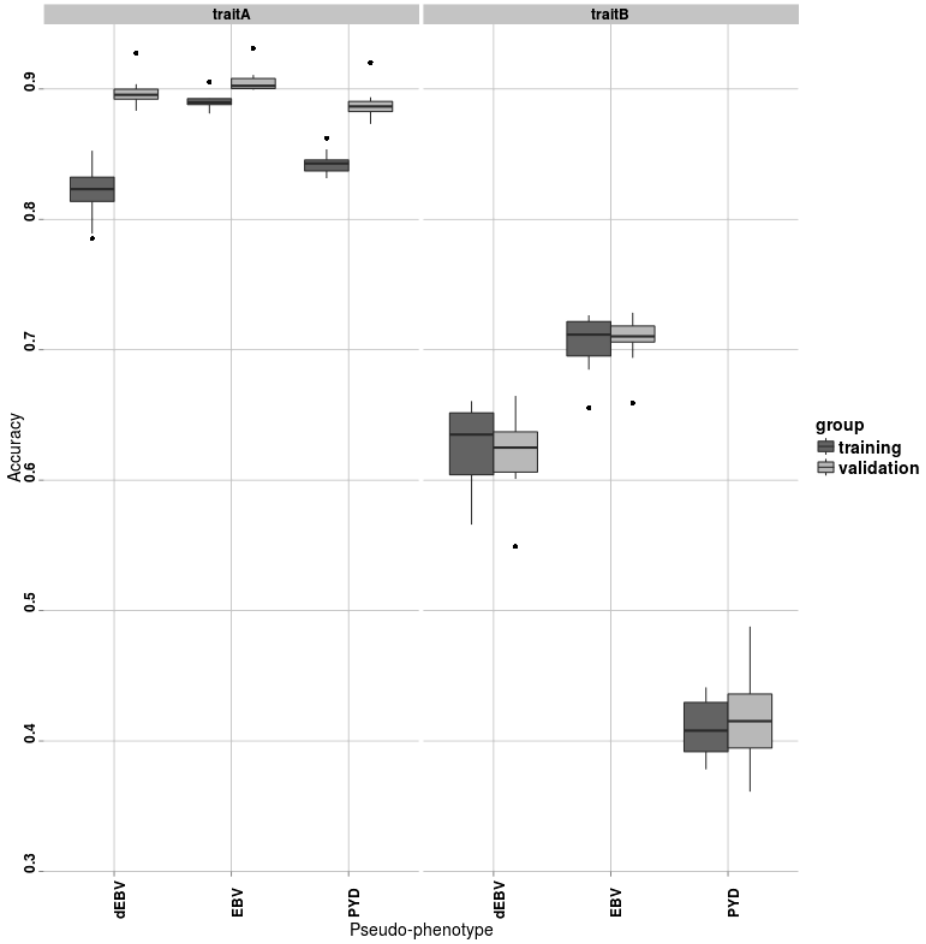


Figure 1 - Accuracies* of the pseudo-phenotypes employed for model training and model validation**

* Boxplots of the Pearson's correlation between the true breeding value and each pseudo-phenotype, considering either training or validation animals of each scenario (10 replicates). Trait A: heritability = 0.30; 5 QTL explained half of the additive variance/ expressed in both sexes; Trait B: heritability = 0.10 / 5 QTL explained half of the additive variance / expressed only in females. **dEBV = deregressed proof (Garrick et al. [3]); EBV = estimated breeding value; PYD = progeny-yield deviation (VanRaden & Wiggans [11]).

For the trait of moderate heritability (trait A), the average accuracy of EBV was about 0.89 for training animals, while similar figures were smaller for dEBV and PYD (average about 0.82 and 0.84, respectively). For the validation animals, the accuracies of all pseudo-phenotypes were more similar (average about 0.90). The slightly larger accuracies of pseudo-phenotypes of the validation animals are related to the fact that, due to the forward prediction scheme adopted, additional information (phenotypes of animals from generations 8 to 10) was considered in such predictions, but not in the genetic evaluation in which pseudo-phenotypes of training animals were generated. For trait B (low heritability), the average accuracy of a given pseudo-phenotype was quite similar between training and validation sets. The average accuracy of EBV was about 0.70, while correspondent figures for dEBV and PYD were 0.62 and 0.41, respectively.

Accuracy of genomic predictions. For trait A, when the true breeding value of the validation animals was employed as predictand to assess the accuracy of genomic predictions (validation scenario 1), the accuracy achieved with a given statistical method was quite similar regardless of the type of pseudo-phenotype employed to estimate marker effects (Table 2). The accuracies achieved with variable selection methods (Bayes Cpi and LASSO) were between 10% and 14% greater than those verified with GBLUP and there was no difference between the accuracy achieved with the two alternative implementations of GBLUP. Given the standard errors of such estimates of accuracy, there was no important difference between Bayes Cpi and LASSO.

For trait A, the validation strategy in which the same type of pseudo-phenotype employed to estimate marker effects was used as predictand (validation scenario 2) underestimated GEBV accuracy in all scenarios (estimates between 2% and 8% smaller than true accuracies) (Table 2). The smallest deviance from the true accuracy was achieved using EBV as predictand. In validation scenario 2, if the accuracies achieved with a same method and different pseudo-phenotypes were contrasted, the validation results could lead to the (wrong) inference of a small advantage for using EBV to estimate marker effects. For trait A, the inferences about the relative performance of the different statistical methods would be quite similar in both validation scenarios, with Bayes Cpi and LASSO outperforming GBLUP.

For trait B, greater differences of the true GEBV accuracy (validation scenario 1) were observed as a function of the pseudo-phenotype employed to estimate marker effects (Table 2). For both Bayes Cpi and GBLUP, using dEBV or EBV resulted in similar GEBV accuracy, while a slight advantage for using dEBV was observed for LASSO (accuracy 4% greater than that obtained using EBV). On the other hand, using Bayes Cpi and PYD to estimate marker effects lead to a 6% increase in GEBV accuracy compared to using EBV and this same method, while, for the other methods, using PYD resulted in accuracy about 20% smaller than using EBV.

Table 2 - Accuracy[1] of genomic predictions as a function of statistical method and pseudo-phenotype[2] employed in model training, for different validation scenarios[3]

		Pseudo-phenotype used to estimate marker effects								
		EBV			dEBV			PYD		
Trait	Method	Validation scenario			Validation scenario			Validation scenario		
		1	2	diff(%)	1**	2**	diff(%)	1**	2**	diff(%)
A	Bayes Cpi	0.80*	0.76*	-4.9	0.80*	0.74*	-7.3	0.80*	0.74*	-8.0
	GBLUP0	0.70*	0.68*	-3.4	0.70*	0.65*	-7.3	0.69*	0.64*	-8.0
	GBLUP20	0.70*	0.68*	-2.3	0.70*	0.65*	-7.2	0.69*	0.64*	-8.0
	LASSO	0.78*	0.75*	-4.6	0.78*	0.72*	-7.4	0.78*	0.71*	-8.0
B	Bayes Cpi	0.62*	0.54*	-12.7	0.63*	0.40*	-36.3	0.66*	0.29*	-56.3
	GBLUP0	0.56*	0.52*	-6.4	0.55*	0.35*	-35.8	0.43*	0.18*	-58.6
	GBLUP20	0.56*	0.54*	-2.7	0.55*	0.35*	-36.0	0.44*	0.18*	-60.2
	LASSO	0.64*	0.55*	-15.1	0.67*	0.42**	-36.6	0.52*	0.22*	-58.1

[1] Average of 10 replicates for the Pearson's correlation between the breeding value estimated using genomic information (GEBV) and the target predictand in each validation scenario. *Standard error $\leq$ 0.01. **Standard error $\leq$ 0.02.

[2] EBV = estimated breeding value; dEBV = deregressed proof [3]; PYD = progeny-yield deviation [11].

[3] Validation scenario 1: GEBV was correlated to the true breeding value. Validation scenario 2: GEBV was correlated to the same type of pseudo-phenotype employed in model training. For both scenarios, the correlations were calculated considering information from animals of the validation set. diff(%) : difference between the estimates of accuracy in validation scenarios 2 and 1, expressed as proportion of the true values (scenario 1), in %.

For trait B, under validation scenario 1, no important differences were verified between the GEBV accuracies achieved with the two versions of GBLUP and the variable selection methods outperformed GBLUP (Table 2). As a general rule, such advantage was larger when unregressed variables were used, so that the largest advantage for a variable selection method was observed when PYD was employed for model training, when Bayes Cpi achieved accuracy about 50% greater than

GBLUP, thus suggesting an interaction between prediction method and pseudo-phenotype.

In validation scenario 2, the extent of underestimation of the true GEBV accuracy was more pronounced in the case of trait B. Using EBV as predictand resulted in underestimation of GEBV accuracy about 15% at most, while the underestimation of GEBV accuracy was more severe when using dEBV and PYD (estimated accuracies about 35% and 56% smaller than the true values, respectively). For this reason, such validation scenario would lead to erroneous inferences about the relative benefit of using different pseudo-phenotypes in marker effects estimation, suggesting an undue advantage in using EBV over dEBV.

When EBV was used in validation scenario 2, the estimated GEBV accuracy of variable selection methods was only 4% greater than that of GBLUP, while the true difference (estimated in scenario 1) varied between 11% and 15% (Table 2). When either dEBV or PYD were used for validation, the differences in accuracy between methods were more consistent with the true differences (at least proportionally).

As mentioned previously, under validation scenario 2, the estimates of GEBV accuracy based on the raw correlation between predictand and GEBV (acc) lead to underestimation of the true accuracy, especially when using dEBV or PYD (Table 2). This underestimation was more severe in the scenario where the predictands used for validation had lower reliability (trait B). For trait A, where the accuracy of all pseudo-phenotypes is larger, the different estimates of accuracy did not differ from the true accuracy by more than 8%, regardless of the type of pseudo-phenotype considered.

Among the different alternatives investigated to estimate GEBV accuracy, the best approximation to the true accuracy was verified when using dEBV and the estimate obtained after dividing acc by the average accuracy of predictand in the validation set (acc_adj), for which the largest absolute difference between true accuracy and acc_adj was about 3% (Table 3) (compared to 36% without adjustment, acc). For trait B, the calculation of acc_adj using EBV overestimated GEBV accuracy, while resulted in underestimation of accuracy when using PYD. The weighted Pearson correlation (acc_w) resulted in negligible benefit, when compared to the raw estimates of correlation (acc) (Table 3).

Table 3 - Relative difference between alternative estimates of accuracy of genomic prediction[3] and correspondent true values[2], using different pseudo-phenotypes[1] and methods to estimate marker effects

EBV						
trait	method	true	acc	acc_adj	acc_w	rPEV
traitA	Bayes Cpi	0.80	-5	3	-4	
	GBLUP0	0.70	-3	5	-3	9
	GBLUP20	0.70	-2	6	-2	-3
	LASSO	0.78	-5	3	-4	
traitB	Bayes Cpi	0.62	-13	20	-11	
	GBLUP0	0.56	-6	29	-5	17
	GBLUP20	0.56	-3	34	-2	4
	LASSO	0.64	-15	17	-14	
dEBV						
trait	method	true	acc	acc_adj	acc_w	rPEV
traitA	Bayes Cpi	0.80	-7	1	-7	
	GBLUP0	0.70	-7	1	-6	1
	GBLUP20	0.70	-7	1	-6	-5
	LASSO	0.78	-7	1	-7	
traitB	Bayes Cpi	0.63	-36	-3	-32	
	GBLUP0	0.55	-36	-2	-32	-7
	GBLUP20	0.55	-36	-2	-32	-18
	LASSO	0.67	-37	-3	-33	
PYD						
trait	method	true	acc	acc_adj	acc_w	rPEV
traitA	Bayes Cpi	0.80	-8	1	-7	
	GBLUP0	0.69	-8	2	-7	6
	GBLUP20	0.69	-8	1	-7	-6
	LASSO	0.78	-8	1	-7	
traitB	Bayes Cpi	0.66	-56	-43	-55	
	GBLUP0	0.43	-59	-46	-58	53
	GBLUP20	0.44	-60	-48	-60	32
	LASSO	0.52	-58	-45	-56	

[1]In each scenario and replicate, the following pseudo-phenotypes were employed to estimate marker effects: EBV = estimated breeding value; dEBV = deregressed proof [3]; PYD = progeny-yield deviation [11]. Validation was carried out assuming that the same type of pseudo-phenotype employed in model training was available on validation animals.

[2]True accuracy (true) was computed as the Pearson's correlation between breeding values estimated using genomic information (GEBV) and the true breeding value, considering the animals in the validation set.

[3]GEBV accuracy was estimated based on the following criteria: i) Pearson's correlation between GEBV and pseudo-phenotype of validation animals (acc); ii) acc divided by the average accuracy of pseudo-phenotypes in the validation set (acc_adj), as in [28]; iii) acc weighted by a function of the reliability of pseudo-phenotypes in the validation set (acc_w), as in [35]; iv) accuracy based on the estimated PEV, as in [36]. Values are expressed as the relative difference between estimated GEBV accuracy and the true value, as a proportion of the true value, in %. (average of 10 replicates).

When GBLUP was used for model training, another alternative to estimate GEBV accuracy was based on the estimated prediction error variance (rPEV). A comparison between the two alternative implementations of this method evidenced that GBLUP20 consistently resulted in smaller estimates of GEBV accuracy when compared to GBLUP0 (Table 3). As a general rule, using either GBLUP0 and dEBV or GBLUP20 and EBV resulted in averages of rPEV that provided reasonable approximations for the true accuracy.

Scale of genomic predictions. In most of the simulated situations, for validation scenario 1, the estimates for the slope of regression of TBV on GEBV were greater than 1, regardless of prediction method, suggesting that genomic predictions were deflated (i.e. too regressed) when compared to the scale of TBV (Table 4). As a general rule, greater departure from 1 was verified in the case of trait B.

Table 4 - Scale[*] of genomic predictions as a function of statistical method and pseudo-phenotype[1] employed in model training, for different validation scenarios[2]

Trait	Method	Pseudo-phenotype used to estimate marker effects								
		EBV			dEBV			pYD		
		Validation scenario			Validation scenario			Validation scenario		
		1	2	diff(%)	1	2	diff(%)	1	2	diff(%)
A	Bayes Cpi	1.12 (0.01)	0.96 (0.01)	-13.9	1.12 (0.02)	1.11 (0.02)	-0.4	1.13 (0.01)	1.12 (0.02)	-0.9
	GBLUP0	1.15 (0.02)	1.01 (0.02)	-12.6	1.07 (0.02)	1.07 (0.02)	-0.4	0.99 (0.01)	0.98 (0.02)	-0.9
	GBLUP20	1.27 (0.02)	1.13 (0.02)	-11.6	1.05 (0.02)	1.05 (0.02)	-0.3	1.10 (0.02)	1.09 (0.02)	-0.8
	LASSO	1.28 (0.01)	1.10 (0.02)	-13.7	1.15 (0.01)	1.14 (0.01)	-0.5	1.18 (0.02)	1.17 (0.02)	-0.8
B	Bayes Cpi	1.28 (0.02)	0.81 (0.01)	-36.8	1.29 (0.01)	1.27 (0.02)	-2.2	1.46 (0.01)	0.71 (0.02)	-51.3
	GBLUP0	1.50 (0.02)	1.01 (0.02)	-32.3	1.24 (0.02)	1.23 (0.03)	-1.1	0.95 (0.02)	0.44 (0.03)	-53.7
	GBLUP20	1.62 (0.03)	1.14 (0.03)	-29.6	1.38 (0.02)	1.36 (0.04)	-1.4	1.09 (0.02)	0.48 (0.04)	-55.5
	LASSO	1.72 (0.03)	1.06 (0.02)	-38.6	1.05 (0.02)	1.02 (0.02)	-2.8	1.23 (0.04)	0.57 (0.02)	-53.5

[*] Average of 10 replicates for the slope (b1) of the regression of predictand on the breeding value estimated using genomic information (GEBV). Standard errors are presented in brackets. [1]EBV = estimated breeding value; dEBV = deregressed proof [3]; PYD = progeny-yield deviation [11]. [2]Validation scenario 1: the predictand was the true breeding value. Validation scenario 2: the predictand was the same type of pseudo-phenotype employed in model training. For both scenarios, the regression was performed with data of the validation set. diff(%) : difference between the estimates of b1 in validation scenarios 2 and 1, expressed as a proportion of the true values (scenario 1), in %.

For both traits, using EBV to estimate marker effects induced a slope with the greatest departure from 1, except for Bayes Cpi. In this situation, using GBLUP20 and LASSO resulted in predictions considerably deflated (Table 4). For trait A, using GBLUP0 and PYD to estimate marker effects provided the best result in terms of scale of GEBV (0.99). For trait B, the combinations GBLUP0-PYD and LASSO-dEBV provided the best results for scale of GEBV (0.95 and 1.05, respectively).

Another important result regards to the correspondence between estimates of b1 between both validation scenarios. As a general rule, using dEBV as predictand (validation scenario 2) induced a slope quite similar to that estimated using TBV (validation scenario 1). Especially for trait B, using EBV or PYD as predictand could lead to erroneous inferences about the scale of genomic predictions. For instance, while using EBV and GBLUP0 resulted in a estimate of b1 equal to 1.5 in validation scenario 1, the use of EBV as predictand in validation scenario 2 induced an estimated slope of 1.01, suggesting (erroneously) that this scenario would provide the best result in terms of scale of GEBV.

Under validation scenario 2, it was not verified benefit for using the alternative estimate of b1 (b1_w), obtained after fitting a weighted linear regression. Such procedure was intended to account for differences in the reliabilities of the pseudo-phenotypes employed as predictands to estimate b1, but resulted in slope estimates very close to those obtained in the non-weighted regressions (data not shown).

Overall fit. Due to the differences in the scale of the different pseudo-phenotypes used as predictands in the validation scenario 2, the results for MSE would not be comparable across all pseudo-phenotypes. For this reason, only the results of overall fit for validation scenario 1 are presented (Table 5).

For both traits, the smallest estimated MSE was obtained when training the model using LASSO and dEBV (Table 5). In the case of trait B, a comparison between pseudo-phenotypes within prediction method evidenced significantly worse results for using PYD, except in the case of Bayes Cpi.

Table 5 - Overall fit* of genomic predictions as a function of statistical method and pseudo-phenotype[1] employed in model training

	Trait A			Trait B		
Method	EBV	dEBV	PYD	EBV	dEBV	PYD
Bayes Cpi	2.96 (0.75)	2.94 (0.71)	2.81 (0.65)	2.08 (0.17)	2.03 (0.16)	2.12 (0.21)
GBLUP0	1.46 (0.07)	1.47 (0.07)	1.47 (0.08)	2.13 (0.19)	2.11 (0.19)	4.12 (0.24)
GBLUP20	1.52 (0.07)	1.45 (0.07)	1.49 (0.07)	2.17 (0.19)	2.13 (0.19)	4.10 (0.24)
LASSO	1.17 (0.04)	1.14 (0.04)	1.17 (0.04)	1.95 (0.18)	1.66 (0.16)	3.85 (0.24)

* Average of 10 replicates for mean squared error (MSE). Standard errors are presented in brackets. The target predictand used to compute MSE was the true breeding value, considering the information in the validation set.

[1] EBV = estimated breeding value; dEBV = deregressed proof [3]; PYD = progeny-yield deviation [11].

Discussion

Study design. The pattern of linkage disequilibrium (r^2) between SNP markers obtained in the present study was consistent with the figures previously reported for beef cattle breeds (e.g. [37], [38], [39]) and the estimates of effective population size in the recent generations were close to those reported for cattle populations undergoing selection (e.g. [40], [41]). Thus, the current simulation design was suitable to generate a population resembling real beef cattle populations, both in terms of LD and Ne.

In simulation studies on genomic selection in cattle, a smaller number of individuals is often simulated, and individual's own phenotypes and genotypes are used to train and validate genomic predictions, what is not a drawback, provided that effective population size and the pattern of linkage disequilibrium mimic the target population properly. Despite the higher computational cost, we opted to simulate a population of larger size, in order to mimic a situation where information on progeny-tested bulls is employed for genomic predictions, similarly as in many studies of genomic selection in real dairy (e.g. [7], [21]) and beef cattle populations (e.g. [9], [10]).

It would be worth to emphasize that the present results should be analyzed after taking into consideration two important assumptions done here. First, it was assumed that markers and QTL had similar properties (i.e. mode of inheritance,

mutation rate and allelic frequencies), similarly as in other simulation studies (e.g. [42], [43]). Visscher et al. [44] suggested that markers could not capture all QTL variance if the allelic frequencies of markers and QTL were much different, what would imply in low LD between markers and QTL and lower accuracy of genomic predictions, even when using high-density marker panels.

The second assumption relates to the fact that only five loci (major QTL) explained one half of the additive variance. Given the previous results regarding QTL mapping and genomic selection in cattle ([22], [21]), one can expect that this situation will not be often observed with real data. Despite this, the extreme scenarios simulated for the QTL distribution magnified differences between prediction methods in terms of accuracy, what was opportune to assess the power to detect such differences under different validation strategies, as mentioned earlier.

Statistical methods. It is known from previous simulation studies that variable selection methods tend to outperform GBLUP when the number of QTL is smaller than the effective number of chromosome segments and/or when traits are influenced by QTL of large effect ([43], [45], [46]). The superiority of variable selection methods over GBLUP in this study confirmed such expectation.

For the trait of low heritability and expressed only in females (trait B), the relative performance of different methods was affected by the type of pseudo-phenotype employed as response variable. While a slight advantage was verified for LASSO over Bayes Cpi in terms of accuracy when either EBV or dEBV information were considered in model training, Bayes Cpi outperformed LASSO when using PYD. An explanation for such result could be related to the fact that Bayes Cpi was the only method in which the residual variance was estimated, so that a proper variance ratio was considered when training the model with this method.

According to [26], a weighted analysis using weights defined as a function of the reliability of pseudo-phenotypes would account for heritability and differences in progeny size of each bull, resulting in an adequate variance ratio for model training. In the present study, because no fixed effects were simulated, the reliability of PYD was estimated using selection index theory, similarly as in [47], while reliabilities of EBV were based on the estimated PEV and reliabilities of dEBV were derived following [3]. For trait B, the estimated reliabilities of PYD deviated more markedly

from the true reliabilities and this could have influenced model training. In this situation, while GBLUP and LASSO predictions were affected by such misspecification, the more adaptive features of Bayes Cpi possibly contributed to obtain more accurate predictions and better overall fit in this scenario.

Regarding to the comparison between GBLUP0 and GBLUP20, no important differences between them were verified in terms of accuracy for the two traits simulated. GBLUP0 resulted in predictions with better scale for most of the situations investigated, while a slight advantage was verified for GBLUP20 when using high reliable dEBV (trait A). Liu et al. [48] and Gao et al. [28] found benefit for fitting GBLUP models with polygenic effects (similarly as GBLUP20), especially in terms of GEBV scale, although the optimal weighting factor differed across traits. Such discrepancies would suggest that trait-specific weighting factors should be used in genomic evaluations [28].

Moreover, the weight given to polygenic effects in GBLUP also influences the magnitude of individual reliabilities of GEBV (rPEV), with decreasing estimates of rPEV when increasing the weight attributed to pedigree-based relationships. When EBV was employed for model training, GBLUP20 resulted in rPEV more compatible with the true accuracy, while the opposite was verified when using dEBV. From the results of the present study, it seems that the optimal weighting would depend on the pseudo-phenotype employed for model training, although a further study, using a wider range of weights, is needed to enhance conclusions about this topic.

The implementation of the LASSO using the Cp criterion generated predictions as accurate as those from Bayes CPi in most scenarios and was effective at reducing computing time (average computing times for LASSO and Bayes Cpi were about 3 minutes and 8 hours, respectively). Despite this, the worse performance of LASSO for trait B when using PYD could cast doubt on the robustness of this approach, although this occurrence was not reported by [30]. Future studies could compare the Cp-based version of the LASSO to alternative implementations of this method, e.g. LASSO tuned through cross-validation and the Bayesian LASSO, in order to investigate the best compromise between computing time and predictive performance.

Pseudo-phenotypes. The present results suggest that the type of pseudo-phenotype used to estimate marker effects would be more influential when the amount of information available for model training is smaller, especially because the reliabilities of different pseudo-phenotypes will differ more markedly. As a general rule, using different pseudo-phenotypes was found to be more influential on the scale than on the accuracy of genomic predictions.

Although the accuracy is the main statistic used to assess the quality of genomic prediction methods, [49] alerted to the fact that inflation and bias should be matter of concern, especially if animals with different amounts of information are among the selection candidates and biased/inflated predictions are obtained, situation which is expected to benefit newborn animals unduly, while in case of deflation of genomic predictions, younger animals are expected to be under-evaluated.

For trait B, using PYD resulted in considerably smaller accuracy, except in the case of predictions obtained with Bayes Cpi and such result could be related to problems in the estimated reliability of PYD, as discussed previously. As a general rule, there was a slight advantage in accuracy for unregressed pseudo-phenotypes over EBV when using variable selection methods.

In most of the situations, using EBV as response variable resulted in deflated predictions and this problem was more severe when EBV reliability was lower (trait B). Guo et al. [13] also reported the same problems and suggested that GEBV should be scaled back, being divided by the average reliability of EBV in the reference set. However, real data often present more complex data structures that could hamper generalization of such recommendation. For instance, inflated estimates of GEBV were obtained when training the model with different types of pseudo-phenotypes (e.g. [50], [14]) and such occurrences could be related to non-random definition of reference and validation groups [51] and even to model deficiencies [15].

As a general result, using dEBV provided the most robust scenario for estimation of marker effects, both in terms of accuracy and scale of GEBV. Using DYD could also result in better scale than using EBV, but the fact that DYD is less

reliable (has larger residual variation) and typically not available from routine genetic evaluations could discourage its adoption.

Using pseudo-phenotypes as targets in validation. Real applications of genomic selection would require choosing an appropriate pseudo-phenotype as target for validation. In our perspective, an appropriate predictand for validation should allow inferences as close as possible as those obtained if the TBV was available, what could only be assessed through simulation. Some of the results presented in this study can be helpful to prevent problems associated to using improper targets in validation.

Because the reliability of the pseudo-phenotypes employed as predictands is less than 1, a simple Pearson correlation between GEBV and predictand would often underestimate the true GEBV accuracy, similarly as verified in this simulation study. Thus, it has been common to estimate the accuracy of GEBV by diving the correlation between GEBV and predictand in the validation set by the average accuracy of the predictand (e.g. [47], [2], [14]).

From the present results, one can expect that if the EBV is employed as predictand and such correction is applied (e.g. [21], [52]), the GEBV accuracy would be overestimated, especially when EBV reliability is low. Since the extent of overestimation seems to be larger for GBLUP, such validation strategy could contribute to diminish the relative advantage of variable selection methods over GBLUP.

According to [14], because both EBV and GEBV are derived using information from relatives, an autocorrelation between both variables could arise and using EBV for validation would be problematic, especially when a large portion of EBV reliability is due to parent average, what would inflate the estimated GEBV accuracy. Possibly due to these reasons, dEBV has been employed as predictand in some studies, even after training the model using EBV (e.g. [2], [14]). On the other hand, high reliable EBV has been advocated as an appropriate surrogate for TBV in validation (e.g. [53], [54]).

Another drawback of using EBV as target predictand in validation is that such procedure could induce erroneous inferences about the scale of genomic predictions, what was already suggested by [14], based on results with real data. In the present

study, when using GBLUP0, the estimated slope of regression of GEBV on EBV was very close to 1, regardless of EBV reliability, while, compared to TBV, GEBV was in fact scaled down. Such validation procedure would lead to wrong inferences about the scale of GEBV obtained with different methods, benefiting GBLUP0 unduly.

From the present results, the correlation between GEBV and dEBV adjusted by the average accuracy of the dEBV in the validation set was found to be quite robust to estimate the true GEBV accuracy, while the slope of the regression of dEBV on GEBV was also appropriate to assess adequacy of GEBV scale, reason why dEBV could be recommended as target predictand in validation.

This study focused on a comparison of different pseudo-phenotypes for both model training and validation, trough a multi-step procedure, as is the case of many real applications of genomic selection (e.g. [7], [8], [9]). An alternative to this type of procedure is single-step genomic prediction ([55],[56]), which would allow to combine marker-based and pedigree-based relationships to derive predictions for both genotyped and non-genotyped animals simultaneously, based on all phenotypic information available, thus without the need to compute pseudo-phenotypes for genotyped animals.

Some studies have compared both approaches (e.g. [49], [28]) and suggested a slight advantage for the single step-approach, especially in terms of the scale of predictions. However, large computational requirements can be a drawback of single-step if many animals are genotyped [57] and biases can occur if the model includes unknown-parent groups [58]. It is worth to mention that unknown-parent groups are expected to be common in many beef cattle evaluations, especially due to the use of multi-sire natural mating. Anyway, while the concern about pseudo-phenotype definition would matter whenever multi-step approaches are used, validation procedures would require careful definition of the target predictands, regardless of the prediction method employed.

Conclusions

Deregressed proofs (dEBV) comprise a more suitable response variable to estimate marker effects than estimated breeding values (EBV) or progeny-yield deviations (PYD). When validating genomic predictions, proper inferences about the accuracy and scale of GEBV can be obtained using dEBV as target predictand. The

accuracy of GEBV should be assessed by the Pearson's correlation between GEBV and dEBV divided by the average accuracy of dEBV in the validation set.

Literature Cited

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

2. Ostersen T, Christensen OF; Henryon M, Nielsen B, Su G, Madsen P: **Deregressed EBV as response variable yield reliable genomic predictions for pigs**. *Genet Sel Evol* 2011, **43**:38.

3. Garrick DJ, Taylor JF, Fernando RL: **Deregressing estimated breeding values and weighting information for genomic regression analyses.** *Genet Sel Evol* 2009, **41**:55.

4. Moser G, Tier B, Crump RE, Khatkar MS, Raadsma HW: **A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers.** *Genet Sel Evol* 2009*,* **41**: 56.

5. Verbyla KL, Hayes BJ, Bowman PJ, ME Goddard: **Short Note: Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle.** *Genet Res* 2009, **91**: 307-311.

6. 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**:1175-1183.

7. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS: **Invited review: Reliability of genomic predictions for North American Holstein bulls.** *J Dairy Sci* 2009 **92**:16–24**.**

8. Luan T, Woolliams JA, Lien S, Kent M, Svendsen M, Meuwissen THE: **The accuracy of genomic selection in Norwegian Red cattle assessed by cross-validation.** *Genetics* 2009, **183:**1119-1126.

9 . Saatchi M, Schnabel RD, Rolf MM, Taylor JF, Garrick DJ: **Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle.** *Genet Sel Evol* 2012, **44**:38.

10. Weber KL, Drake DJ, Taylor JF, Garrick DJ, Kuehn LA, Thallman RM, Schnabel RD, Snelling WM, Pollak EJ, Van Eenennaam AL: **The accuracies of DNA-based estimates of genetic merit derived from Angus or multibreed beef cattle training populations.** *J Anim Sci* 2012, **90**:4191-4202.

11. VanRaden PM, Wiggans GR: **Derivation, calculation, and use of national animal model information.** *J Dairy Sci* 1991, **74**:2737-2746.

12. Jairath L, Dekkers JC, Schaeffer LR, Liu Z, Burnside EB, Kolstad B: **Genetic evaluation for herd life in Canada.** *J Dairy Sci* 1998, **81**:550-562.

13. Guo G, Lund MS, Zhang Y, Su G: **Comparison between genomic predictions using daughter yield deviation and conventional estimated breeding value as response variables.** *J Anim Breed Genet* 2010*,* **127**:423-432.

14. Gao H, Lund M, Zhang Y, Su G: **Accuracy of genomic prediction using different models and response variables in the Nordic Red cattle population.** *J Anim Breed Genet* 2013*,* **130**: 333-340.

15. Daetwyler HD, Calus MP, Pong-Wong R, de Los Campos G, Hickey JM: **Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking**. *Genetics* 2013, **193**: 347-365.

16. Sargolzaei M, Schenkel FS: **QMSim: a large-scale genome simulator for livestock.** *Bioinformatics* 2009, **25**:680-681.

17. Brito FV, Braccini Neto J, Sargolzaei M, Cobuci JA, Schenkel FS: **Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle.** *BMC Genet* 2011, **12**:80.

18. Sargolzaei M, Schenkel FS, Jansen GB, Schaeffer LR: **Extent of linkage disequilibrium in Holstein Cattle in North America.** *J Dairy Sci* 2008, **91**:2106-2117.

19. Flury C, Tapio M, Sonstegard TS, Drögemüller C, Leeb T, Simianer H, Hanotte O, Rieder S: **Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium**. *J Anim Breed Genet* 2010, **127**: 339–347.

20. Snelling WM, Chiu R, Schein JE, Hobbs M, Abbey CA, Adelson DL, Aerts J, Bennett GL, Bosdet IE, Boussaha M, Brauning R, Caetano AR, Costa MM, Crawford AM, Dalrymple BP, Eggen A, Van der Wind AE, Floriot S, Gautier M, Gill CA, Green

RD, Holt R, Jann O, Jones SJM, Kappes SM, Keele JW, Jong PJ, Larkin M, Lewin HA, McEwan JC, McKay S, Marra MA, Mathewson CA, Matukumalli LK, Moore SS, Murdoch B, Nicholas FW, Osoegawa K, Roy A, Salih H, Schibler L, Schnabel RD, Silveri L, Skow LC, Smith TPL, Sonstegard TS, Taylor JF, Tellam R, Van Tassell CP, Williams JL, Womack JE, Wye NH, Yang G, Zhao S. for the International Bovine BAC Mapping Consortium: **A physical map of the bovine genome.** *Genome Biol* 2007, **8**:R165.

21. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME: **Invited review: Genomic selection in dairy cattle: Progress and challenges.** *J Dairy Sci* 2009, **92**:433-443.

22. Hayes BJ, Goddard ME: ***The distribution of the effects of genes affecting quantitative traits in livestock.*** *Genet Sel Evol* 2001, ***33****: 209–229.*

23. Misztal I: **BLUPF90 - a flexible mixed model program in Fortran 90.** Athens: Animal and Dairy Science -University of Georgia; 2008.

24. Amer PR, Banos G**: Implications of avoiding overlap between training and testing data sets when evaluating genomic predictions of genetic merit.** *J Dairy Sci* 2010, **93**:3320–3330.

25. Olson KM, VanRaden PM, Tooker ME, Cooper TA: **Differences among methods to validate genomic evaluations for dairy cattle.** *J Dairy Sci* 2011, **94**:2613-2620.

26. VanRaden PM: **Efficient methods to compute genomic predictions.** *J Dairy Sci* 2008, **91**: 4414–4423.

27. Sargolzaei M, Schenkel FS, VanRaden PM: ***gebv: Genomic breeding value estimator for livestock***. *In Technical Report to the Dairy Cattle Breeding and Genetics Committee***.** Guelph: University of Guelph; 2009.

28. Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS, Su G: **Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population.** *Genet Sel Evol* 2012, **44**: 8.

29. Tibshirani R: **Regression shrinkage and selection via the lasso.** *J R Stat Soc B* 1996, **58**: 267-288.

30. Usai MG, Carta A,Casu S: **Alternative strategies for selecting subsets of predicting SNPs by LASSO-LARS procedure .** *BMC Proc* 2012, **6**(Suppl 2):S9.

31. R Development Core Team: *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna; 2012. ISBN 3-900051-07-0, [http://www.R-project.org/]

32. Habier D, Fernando RL, Kizilkaya K, Garrick, DJ: **Extension of the bayesian alphabet for genomic selection.** *BMC Bioinformatics* 2011, **12**:186.

33. Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando RL: **Additive genetic variability and the Bayesian alphabet.** *Genetics* 2009, **183**: 347–363.

34. Legarra A, Ricard A, Filangi O: ***GS3 software: Genomic Selection, Gibbs Sampling and Gauss-Seidel*. 2013.**
[http://snp.toulouse.inra.fr/~alegarra/manualgs3_last.pdf].

35. Legarra A, Robert-Granié C, Croiseau P, Guillaume F, Fritz S: **Improved lasso for genomic selection.** *Genet Res* 2011, **93**:77-87.

36. Clark, S.A.; Hickey, J.M.; Daetwyler, H.D.; Van Der Werf, J.H.J: **The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes.** *Genet Sel Evol* 2012, **44**:4.

37. Lu D, Sargolzaei M, Kelly M, Li C, Van der Voort G, Wang Z, Plastow G, Moore S, Miller SP**: Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle.** *Front Genet* 2012, **3**:152.

38. O'Brien AMP, Garcia JF, Carvalheiro R, Neves HHR, Van Tassell CP, Sonstegard TS, Utsunomiya YT, Mcewan JC, Solkner J: **Comparing linkage disequilibrium between taurine and indicine cattle with a high density SNP chip.** In *Book of Abstracts of 63rd Annual Meeting of the European Association for Animal Production*, 27-31 August 2012; Bratislava. Wageningen Academic Publishers; 2012: 18.

39. Espigolan R, Baldi F, Boligon AA, Souza FR, Gordo DG, Tonussi RL, Cardoso DF, Oliveira HN, Tonhati H, Sargolzaei M, Schenkel FS, Carvalheiro R, Ferro JA, Albuquerque LG**: Study of whole genome linkage disequilibrium in Nellore cattle.** *BMC Genomics* 2013**, 14:**305**.**

40. Brito FV, Sargolzaei M, Braccini Neto J, Cobuci JA, Schenkel FS: **Pedigree analysis in a large Brazilian Nellore herd.** In *Proceedings of the 9th World*

Congress on Genetics Applied to Livestock Production: 1-6 August 2010; Leipzig. German Society for Animal Science; 2010: PP1-57.

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

42. Kizilkaya K, Fernando RL, Garrick DJ: **Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes**. *J Anim Sci* 2010, **88**:544-551.

43. Coster A, Bastiaansen JW, Calus MP, van Arendonk JA, Bovenhuis H: **Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance.** *Genet Sel Evol* 2010, **42**:9.

44. Visscher PM, Yang K, Goddard ME: **A commentary on 'Common SNPs explain a large proportion of the heritability for human height' by Yang et al.** *Twin Res Hum Genet* 2010, **13**: 517–524.

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

46. Clark SA, Hickey JM, Van Der Werf JHJ: **Different models of genetic variation and their effect on genomic evaluation.** *Genet Sel Evol* 2011, **43**:18.

47. Habier D, Tetens J, Seefried FR, Lichtner P, Thaller G: **The impact of genetic relationship information on genomic breeding values in German Holstein cattle.** *Genet Sel Evol* 2010, **42**:5.

48. Liu Z, Seefried FR, Reinhardt F, Rensing S, Thaller G, Reents R: **Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic prediction.** *Genet Sel Evol* 2011, **43**:19.

49. Vitezica ZG, Aguilar I, Misztal I, Legarra A: **Bias in genomic predictions for populations under selection.** *Genet Res* 2011, **93**:357-366.

50. Croiseau P, Legarra A, Guillaume F, Fritz S, Baur A, Colombani C, Robert-Granie C, Boichard D, Ducroq V: **Fine tuning genomic evaluations in dairy cattle through SNP pre-selection with the Elastic-Net algorithm.** *Genet Res* 2011, **93**: 409-417.

51. Patry C, Ducrocq V: **Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle.** *J Dairy Sci* 2011, **94**:1011–1020.

52. Daetwyler HD, Hickey JM, Henshall JM, Dominik S, Gredler B, Van der Wer JHJ, Hayes BJ: **Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population**. *Anim Prod Sci* 2010, **50**:1004–1010.

53. Cleveland MA, Hickey JM, Forni S: **A common dataset for genomic analysis of livestock populations**. *G3* 2012, **2**:429-435.

54. Pérez-Cabal MA, Vazquez AI, Gianola D, Rosa GJ, Weigel KA: **Accuracy of Genome-Enabled Prediction in a Dairy Cattle Population using Different Cross-Validation Layouts.** *Front Genet* 2012, **3**:27.

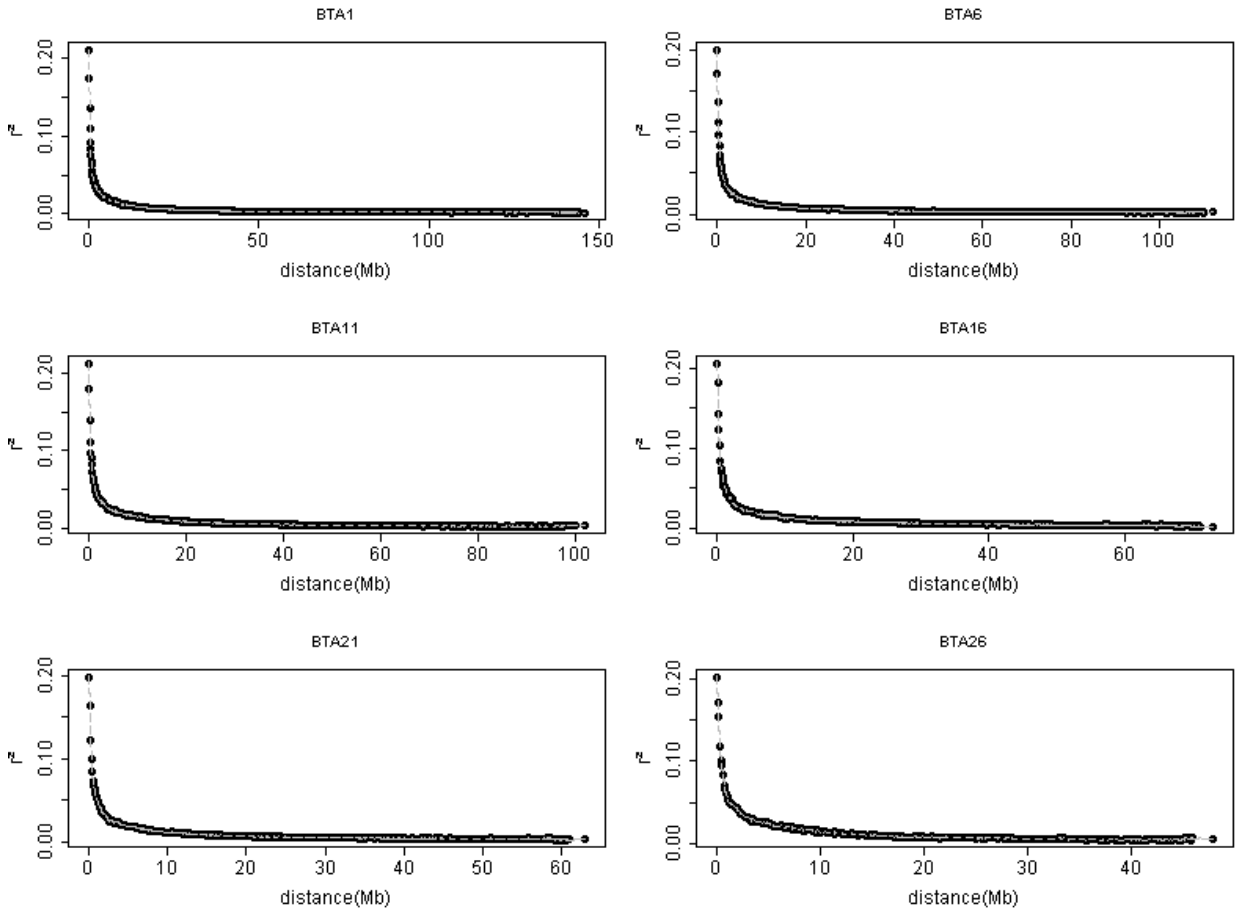
55. A Legarra, I Aguilar, I Misztal: **A relationship matrix including full pedigree and genomic information.** *J Dairy Sci* 2009, **92**: 4656-4666.

56. I Aguilar, I Misztal, DL Johnson, A Legarra, S Tsuruta, TJ Lawlor: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**: 743-752.

57. Legarra A, Ducrocq V: **Computational strategies for national integration of phenotypic, genomic, and pedigree data in a single-step best linear unbiased prediction.** *J Dairy Sci* 2012, **95**: 4629-4645.

58. Misztal I, Vitezica ZG, Legarra A, Aguilar I, Swan AA: **Unknown-parent groups in single-step genomic evaluation.** *J Anim Breed Genet* 2013, **130:** 252-258.

Supplementary Figure 1



Linkage disequilibrium* between simulated syntenic single nucleotide polymorphism (SNP) pairs as a function of physical distance.

*In order to exemplify the extent of linkage disequilibrium generated in this simulation study, the mean r^2 over successive intervals of 0.1 Mb is plotted. For the sake of brevity, instead of all simulated BTA, six chromosomes of varying length were arbitrarily chosen for this purpose.

CHAPTER 4 - Trait-specific long-term consequences of genomic selection in beef cattle

Abstract

Background

Simulation studies can be useful to address long-term consequences of selection schemes as applied to real breeding programs, helping to identify effective strategies to enable genetic gain and maintain genetic diversity in longer time horizons . The aim of this study was to evaluate the impact of application of genomic selection (GS) in beef cattle, by simulating scenarios mimicking situations of particular importance to these populations.

Methods

Forward-in-time simulation was employed to mimic a population with pattern of linkage disequilibrium (LD) close to that verified in real beef cattle populations. Different scenarios of genomic selection and traditional BLUP selection were simulated for 15 generations, mimicking selection of female reproduction, meat quality and growth traits. For GS scenarios, an alternative selection criterion was simulated (wGBLUP), intended to enhance long-term gain by attributing more weight to low-frequency favorable alleles. Marker-based estimators of inbreeding were contrasted to conventional pedigree-based estimators.

Results

GS allowed genetic progress up to 40% greater than BLUP under selection for female reproduction and meat quality traits, while no benefit was verified under selection for a beef cattle growth trait. The alternative criterion wGBLUP did not increased long-term response, while it was beneficial to reduce inbreeding rates and loss of favorable alleles. Estimation of inbreeding using marker information based on the concept of runs of homozygosity outperformed the remainder strategies to estimate homozygosis due to identity-by-descent.

Conclusions

Large benefits were envisaged for GS over traditional selection for scenarios mimicking selection for meat quality and female reproduction. There was evidence

that larger advantage can be expected for GS compared to BLUP when the selected trait is under less polygenic background and that attributing more weight to favorable alleles of low-frequency can contribute to reduce inbreeding rates and loss of favorable alleles in GS.

Background

Since pioneer studies highlighted the potential benefits of the application of genomic selection (GS) in breeding programs (e.g. [1], [2]), many researchers compared this strategy to traditional selection schemes, evaluating the potential gain in accuracy that could be achieved with GS (e.g. [3], [4], [5], [6], [7]). Such focus on accuracy indicate that GS schemes has been evaluated with respect to their potential to enhance short-term response [8].

Besides the need for knowledge about the potential of GS to promote genetic progress in the short-term, it is also important to investigate the long-term consequences of its application. The knowledge about the potential to enable genetic gain and maintain genetic diversity in longer time horizons is essential for the definition of short and long-term goals of breeding schemes and ensure their sustainability.

When compared to traditional selection based on BLUP, in addition to enable increased rates of genetic progress, GS schemes could also reduce the levels of inbreeding accumulated along the selection process [9]. This is justified by the fact that genomic prediction would allow better estimation of the Mendelian sampling term, resulting in smaller co-selection of closely related animals when compared to BLUP.

Bijma [8] suggested that, in absence of genomic information, the only way to increase long-term genetic gains would be invariably related to the maintenance of effective population size (Ne) large enough to cope with this aim, what could imply in some sacrifice of the short-term response. Acceptable levels of Ne can be handled by the adoption of strategies including restriction on inbreeding and optimum contribution selection (OCS). The latter would allow to limit the increase of inbreeding/co-ancestry at each generation while maximizing the genetic gain in a given time horizon (e.g. [10], [11],[12]). Nevertheless, the efficacy of these strategies relies on accurate estimates of relatedness between individuals as well as of the

genetic merit of selection candidates.In this way, the use of genomic information could open new opportunities, allowing more precise estimation of relatedness and inbreeding coefficients ([13], [9], [8], [14]).

Some studies had proposed the use of alternative selection criteria aiming to increase long-term response through proper weighting of the contribution of each locus to the genomic predictions ([15],[16]). Jannink [16] carried out a simulation study based on marker genotypes of 192 inbred lines of a real population of barley. This study highlighted the importance of giving larger weight to low-frequency favorable alleles in the beginning of the selection process, since this strategy would allow to increase long-term genetic gain without compromising the short-term response. Despite the importance of this result, large differences in terms of selection schemes and marker density make difficult to extrapolate such conclusions to the application of GS in cattle.

Sonesson et al. [12] compared the application of truncation selection and OCS, in scenarios using different methods to predict the genetic merit of individuals and using pedigree or genomic information to constraint inbreeding. Such study provide some evidence of differences in local patterns of inbreeding incidence as a function of the statistical method used to derive genomic predictions and that genomic relationships can be valuable to impose restriction on inbreeding.

In real applications of GS in cattle breeding programs, genomic predictions are obtained under routinely update of the reference population used to estimate marker effects (e.g. [17], [18]), thus it would be desirable to address long-term consequences of GS under such circumstances.

Several studies have suggested that larger benefit can be achieved under the application of GS for traits which are difficult/costly to measure as well as for traits measured late in life (e.g. [19], [20], [21]). In such situations, the use of genomic information could allow more accurate predictions of genetic merit, even before individuals have own performance recorded.

According to Carvalheiro el al. [22], some potential for application of GS in Brazilian beef cattle is envisaged, especially due to the large number of animals routinely recorded by breeding organizations and due to the fact that bulls are progeny-tested relatively late in life, what could justify GS, especially for traits in

which the progress achieved by conventional selection is currently limited. Thus, it would be important to evaluate the impact of such application of GS, by simulating scenarios mimicking situations of particular importance to these populations.

Thus, the present study was designed to achieve the following aims:

i) to investigate the long-term impact of genomic selection, when compared to traditional selection, in scenarios mimicking selection for female reproduction traits (E1), meat quality traits (E2) and growth traits (E3);

ii) to evaluate the impact of different genomic prediction methods, under different scenarios of genetic architecture, in terms of genetic progress, inbreeding incidence and maintenance of genetic diversity.

iii) to evaluate the long-term impact of giving larger weight to low-frequency favorable alleles in genomic predictions.

iv) to compare different estimators of inbreeding and genetic diversity based on the genomic information.

Methods

Population structure

The QMSim software [23] was employed to generate 10 populations (replicates) with pattern of linkage disequilibrium (LD) close to that verified in real beef cattle populations. The forward-in-time simulation was carried out with parameters defined similarly as in [21].

First, 1,000 historical generations (H1) were simulated in which the effective population size (Ne) was kept constant at the value of 1,000. After this step, starting from the last generation of H1, 1,020 historical generations (H2) were simulated, so that a bottleneck was simulated by gradual reduction of Ne from 1,000 to 200. Such a strategy allowed the initial establishment of mutation-drift equilibrium and the creation of linkage disequilibrium between the simulated loci. For both H1 and H2, the number of animals of each sex was the same and the matings were simulated assuming random union of gametes.

After the simulation of the 2,020 historical generations, 100 males and 100 females were randomly selected from the last generation of H2 and the population was expanded. In this step, 3 generations were simulated under random mating and

no selection, so that each female produced five offspring per generation and the number of females in reproduction increased exponentially at each generation.

From the last generation simulated in the expansion step, 40 males and 1000 females were randomly selected to compose the base population (G0) from which all recent generations (described later) were simulated. The same simulation framework was employed to produce ten different replicates (base populations), so that the same base populations were used to simulate the starting point of the different scenarios (described later), similarly as applied by [24].

For each scenario and base population (replicate), 15 recent generations were simulated, so that different selection strategies were applied to each scenario, which were succeeded by random mating of the selected animals at each generation to produce the subsequent generation. At each generation, it was assumed that each female and male in reproduction had 1 and 25 offspring, respectively, totaling 1,000 individuals.

Simulated genome

The simulated genome was composed by 29 pairs of autosomes (BTA), whose individual length varied from 40 cM to 146 cM, totaling 2333 cM [25]. Regardless of the larger computational demand, the simulation of such a large genome is convenient to reproduce a more realistic scenario regarding the number of markers and QTL physically unlinked [21]. In the historical generations, a mutation rate (u) of 10^{-4} was simulated for both markers and QTL, allowing to meet the target number of segregating loci in the last generation. For the sake of simplicity, missing genotypes and genotyping errors were not simulated.

In the present study, mutation events were only simulated in the historical generations, since the recent population involved a small number of generations and thus the effect of mutation could be neglected [23]. For both markers and QTL, a recurrent mutation model was adopted, in which one allelic state was changed to other pre-existing allelic state of the same loci. Markers were simulated as bi-allelic, aiming to resemble SNP markers, while the number of QTL alleles at each loci varied from 2 to 4. The markers were distributed evenly spaced along the genome, what resulted in a number of markers per chromosome proportional to length of each BTA.

While the number of QTL was also proportional to the length of each BTA, their position within each chromosome were set randomly.

When a large number of historical generations is simulated, aiming to achieve mutation-drift equilibrium, a large proportion of non-segregating loci is obtained in the last generation. Thus, in the last historical generation, 40,000 segregating markers were randomly selected among those with MAF > 0.01, in order to simulate the information available in marker panels in the recent generations. In addition, 1,000 QTL loci with MAF > 0.01 were randomly selected to simulate phenotypes in the recent generations. Such procedures allowed to simulate marker density similar to that obtained when using commercial SNP panels designed for cattle, after the quality control of genotypes is carried out (e.g. [26], [3]).

For each replicate, 3,064 marker loci (IBD loci) were removed from the simulated panel considered in genomic predictions and were used to monitor identity-by-descent, through the attribution of specific tags of each founder allele available in G0 at each IBD loci, similarly as in Sonesson et al. [12]. Markers with MAF < 5% were also removed from the panel available for genomic prediction and were used to monitor the evolution of loci potentially associated to deleterious mutations (homozygous mutation loci, HML), similarly as in [27].

Simulated scenarios

Three selection schemes were simulated (E1 to E3), aiming to replicate selection for female reproductive traits, meat quality and growth in beef cattle. For a given scheme, each scenario was defined by the application of a particular selection criterion, as a way to compare genomic prediction under different methods to traditional selection based on BLUP.

Since the version of QMSim software employed in this study did not allow to simulate repeated cycles of selection and mating using different genomic prediction methods, custom routines were coded in R language [28] to allow simulation of such scenarios. In order to simulate gametogenesis, the recombination was modeled by assuming that number of crossovers follow a Poisson distribution with average equal to the chromosome length (in Morgans), so that crossover positions were randomly assigned along each chromosome. For each scenario and replicate, the genotypes of

animals of a given generation were simulated assuming random union of gametes of the selected parents, 1000 dams and 40 sires.

The true breeding value (TBV) of each animal was simulated as: $TBV = Zv$, where Z and v are, respectively, the incidence matrix relating individuals to QTL alleles and the matrix of additive effects of QTL alleles. In all scenarios, a continuous phenotype Y was simulated by adding an overall mean (μ = 10) and a random residual (ε) to the TBV of each animal. The residuals were simulated so that ε ~ N (0, σ^2_ε), in which σ^2_ε was defined in each combination of scenario and replicate to met the pre-specified heritability.

The following selection schemes were simulated:

-Female reproduction (E1, Figure 1): A continuous trait of heritability equal to 0.15, influenced by 1000 QTL and only expressed in females was simulated. QTL effects were drawn from a gamma distribution (shape parameter = 0.40). The first generation was obtained after the random mating of the 1040 animals (1000 dams and 40 sires) of the base population (G0), producing 1,000 individuals (G1), about half of each sex.

After G1 was generated, the following selection criteria were employed to select the parents of the next generation (G2), as well as those of the following generations (up to G15):

i) BLUP: estimated breeding value (EBV) obtained using BLUP. This criterion is intended to allow evaluation of the consequences of traditional selection using BLUP.

ii) GBLUP: genomic estimated breeding value (GEBV), where $GEBV = \sum x_j \widehat{\beta_j}$, x_j is the genotype for the j-th marker and $\widehat{\beta_j}$ is the respective estimated allele substitution effect. This criterion is intended to allow evaluation of the consequences genomic selection schemes.

iii) wGBLUP: weighted GEBV (wGEBV), where $wGEBV = x_j \, \widehat{\beta_j} \, p_j^{-0.5}$, in which x_j and $\widehat{\beta_j}$ are the genotype and estimated allele substitution effect, respectively, and p_j is the respective frequency of the favorable allele for the j-th marker. This criterion is identical to that applied by Jannink (2010) and is intended to allow reducing the loss of favorable alleles of lower frequency, by attributing more weight to them when computing genomic predictions. By contrasting this scenario to that in which GBLUP was used, one could evaluate the potential benefit of such a weighting strategy.

iv) TBV: true breeding value, as described previously. The simulation of this criterion was intended to allow investigation of more extreme consequences of the selection process on genetic progress and genetic diversity.

For all scenarios, it was assumed that the phenotype of each female would be available only after the selection decisions. Thus, when selecting the parents of generation **i+1**, genetic predictions for females from generation **i** were based only on the phenotypic information accumulated until generation **(i-1)**. After this step, among the females of generation **i**, only the phenotypes of those selected for reproduction were made available and considered in the subsequent genetic evaluation. For the sake of simplicity, it was assumed that each female expressed the phenotype once.

At each generation, the BLUPF90 software [29] was employed to predict EBVs of all animals. All pedigree and phenotypic information accumulated and available until each generation was considered in BLUP equations, as well as the variance ratios calculated with the simulated values for the variance components.

In the scenarios under genomic selection, the 1000 females from G0 were assumed to be genotyped with a 50k panel, composing the initial reference population employed to estimate marker effects and needed to simulate the scenarios in which GBLUP or wGBLUP were applied. Marker effects were estimated using ridge regression, assuming that the effects of all loci follow the same normal distribution [1], using a framework similar to that described in [30].

Before each new cycle of selection and reproduction, the prediction equations and genomic predictions were updated with phenotypes and genotypes of the females selected in the previous generation. In order to select the parents of generation **i+1**, it was assumed that animals from generation **i** and their parents were available for reproduction (selection candidates). In addition to the females with phenotypes, all selection candidates were assumed to be genotyped. Genotyping of all selection candidates from a lower density and imputing genotypes to higher density has been suggested as a cost-effective strategy to implement GS [31], what could justify the assumption of genotyping of all selection candidates in the present study. For the sake of simplicity, all genotyped animals were assumed to have genotypes available for all simulated markers (mimicking a 50k panel).

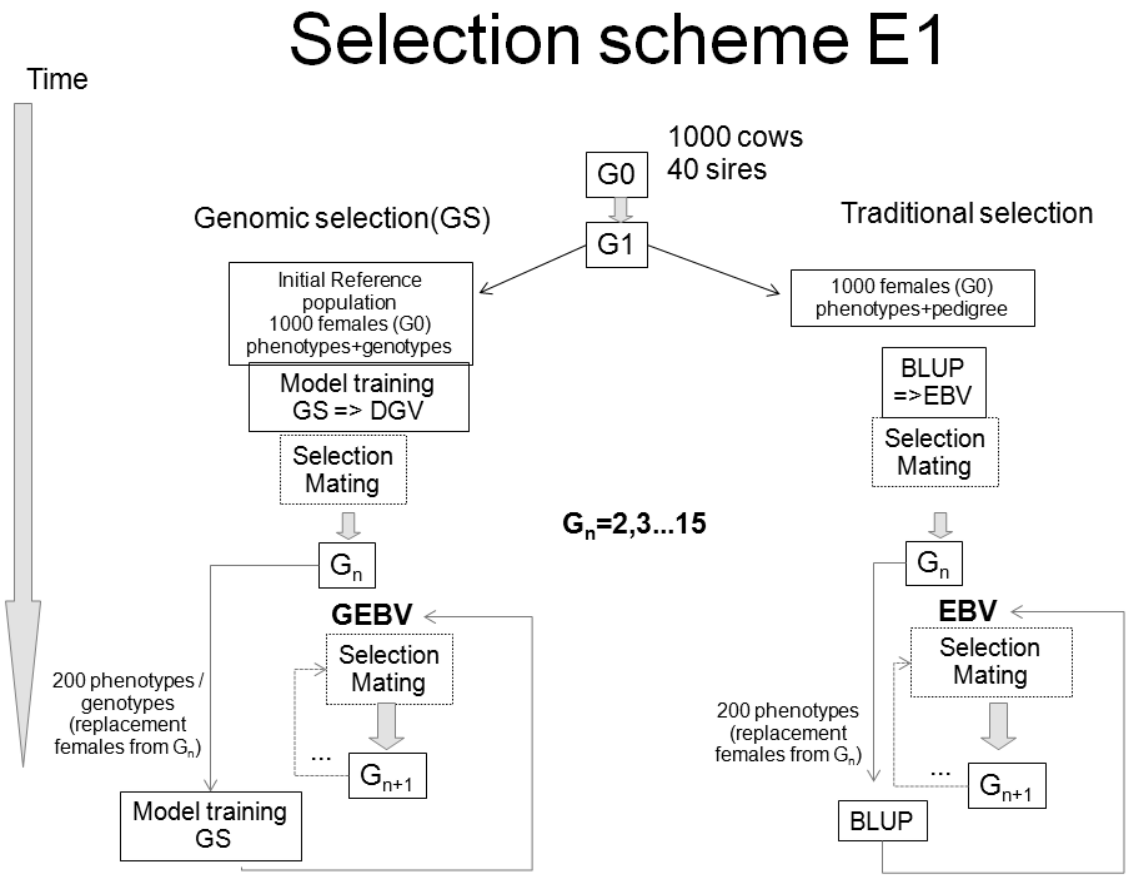


Figure 1. Fluxogram of the simulation of selection for female reproduction (E1)
Each generation (n=1,2,...15) consisted of 1000 animals (500 males and 500 females). After selection and mating, it was assumed that the phenotypes of the females selected for replacement (N=200) were recorded and added to the dataset used to compute GEBV (GS) or EBV (BLUP). In the case of GS, it was assumed that these replacement females were genotyped with a 50k panel.
After generation **n+1** was produced, these animals (500 females and 500 males) were included among the selection candidates, so that their genetic merit was predicted based on all information available on females from generation **n** or older.
For both GS and traditional selection, two scenarios of replacement strategies were simulated : v1) fixed replacement rate of 20% (for both males and females); v2) selection of the top 1000 females and 40 males among the selection candidates ("variable replacement rate"). In the case of GS, two alternative scenarios (GBLUP and wGBLUP) were simulated for the computation of genomic predictions.

For each selection criterion, two replacement strategies were simulated: 1) culling of the worse 20% animals (8 males and 200 females) among the parents of generation **i** (v1, "fixed replacement rate") and selection of the top 8 males and top 200 females from generation **i** to replace them; 2) selection of the top 1,000 females and top 40 males for that criterion, considering all animals available for reproduction, regardless of their generation coefficient (v2, "variable replacement rate").

Thus, for E1, eight different scenarios were simulated, as the combination between selection criterion (BLUP, GBLUP, wGBLUP or TBV) and replacement strategy (v1 or v2), for ten replicates.

-Meat quality traits (E2, Figure 2): Three different traits were simulated, all of them with the same heritability (h^2=0.35) and expressed by animals of both sexes:

-Trait A: influenced by 1000 QTL, whose effects were drawn from a gamma distribution (shape = 0.40).

-Trait B: influenced by 100 QTL, whose effects were drawn from a gamma distribution (shape = 0.40).

-Trait C: influenced by 1000 QTL, five of which had larger effect ("major QTL"), each one explaining 7% of the additive variance and with position randomly sampled among the QTL loci. The remainder QTL (995) had effects drawn from a normal distribution, which were re-scaled such that together they explained the remainder 65% of the additive variance.

Thus, the traits simulated were intended to mimic different scenarios of genetic architecture: more polygenic control (A), trait influenced by smaller number of loci (B) and one trait influenced by large number of loci, with a few of them having very larger effect (C).

Similarly as in scheme E1, the first generation was obtained after the mating of the 1,040 animals from the base population (G0), producing 1,000 offspring, about a half of each sex. After G1 was obtained the following selection criteria were employed to select the parents of the next generation (G2), as well as those of the following generations (up to G15):

-TBV (previously described)

-BLUP (previously described)

-GBLUP (previously described)

-wGBLUP (previously described)

-LASSO: GEBV calculated with estimated marker effects obtained through the use of the variable selection method elastic net (EN) [32]. Estimation of marker effects was carried out using glmnet R package [28]. This method can be understood as an extension of the LASSO method [33], in which solutions are obtained under constraints imposed by a mixed penalty parameter, resultant from the combination between the LASSO penalty and the ridge regression penalty, and it is expected to outperform LASSO when predictors are highly correlated, as is the case of SNP markers.

Because the present implementation considered a larger weight (95%) for the LASSO penalty, this method is regarded as LASSO hereinafter. As suggested in previous studies (e.g. [34], [35]), it is expected that variable selection methods outperform GBLUP in terms of predictive ability when the trait is influenced by a few QTL and/or by large QTL.

-wLASSO: genomic breeding value estimated similarly as in the case of wGBLUP (attributing larger weight to rare alleles), except by the fact that marker effect estimates were obtained via EN.

It was assumed that, starting from G1, phenotypes would be measured on 250 animals randomly chosen at each generation (Figure 2), being that phenotyping would require the slaughtering of these animals, as is the case of most carcass and meat quality traits.

In order to select the parents of generation **i+1**, the animals from generation **i** that were not slaughtered were available for reproduction (i.e. selection candidates) as well as the parents of the animals from generation **i**. For selection candidates from generation **i** or older, each selection criterion was computed considering all information accumulated until that generation (including phenotypes on half-sibs from generation **i**).

For scenarios under genomic selection, the 250 phenotypes from G1 animals comprised the initial reference population, employed to estimate marker effects under the different methods previously described. At each generation, marker effect estimates were updated, after including the new phenotypes and genotypes in the

reference set. In addition to the animals slaughtered, it was assumed that all animals available for reproduction (selection candidates) were genotyped, similarly as in E1.

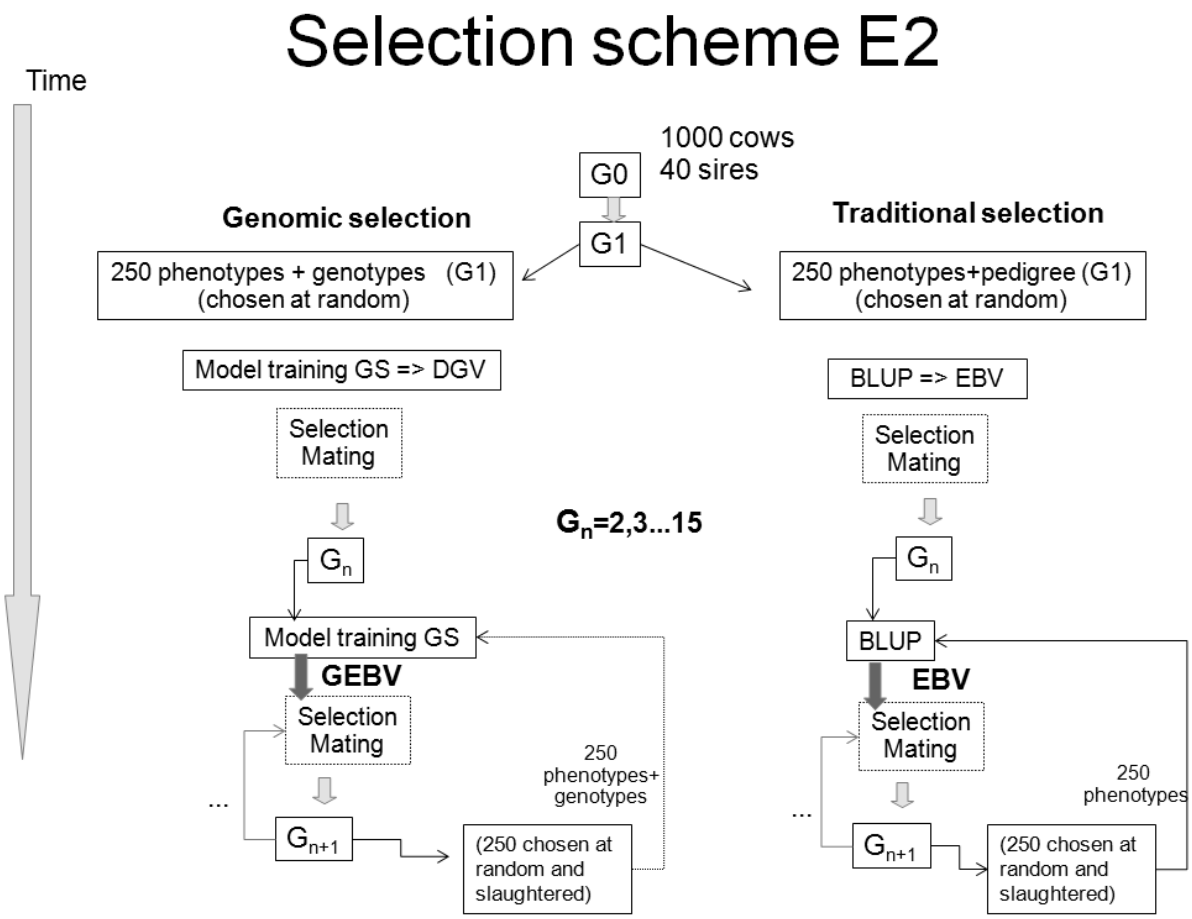


Figure 2. Fluxogram of the simulation of selection for meat quality (E2)
Each generation (n=1,2,...15) consisted of 1000 animals (500 males and 500 females). At generation n, 250 young animals were chosen at random to be slaughtered. The remainder animals (~375 males and ~375 females) were included in the set of selection candidates (jointly with the parents of generation n), in order to simulate the subsequent cycle of selection and mating.
Following this same simulation framework three different traits (A, B and C), mimicking different scenarios of genetic architecture, were simulated (see methods section for more detailed description).
In the case of trait A, for both GS and traditional selection, two scenarios of replacement strategies were simulated : v1) fixed replacement rate of 20% (for both males and females); v2) selection of the top 1000 females and 40 males among the selection candidates ("variable replacement rate"). For traits B and C, the replacement strategy v1 was adopted.
In the case of GS, four alternative scenarios (GBLUP,wGBLUP, LASSO and wLASSO) were simulated for the computation of genomic predictions.

Aiming to investigate the effect of replacement strategies in this scheme, the same strategies described as fixed (v1) and variable replacement rate (v2) were applied in the scenarios with selection for trait A, while only the strategy of fixed

replacement rate was applied in the case of traits B and C. Thus, for each selection criterion and replicate, four scenarios were simulated (A_v1, A_v2, B_v1 and C_v1), totaling 24 scenarios

- Growth trait (E3, Figure 3): A continuous trait, with heritability equal to 0.35, expressed in both sexes and influenced by 1,000 QTL was simulated. QTL effects were drawn from a gamma distribution (shape=0.40). The first generation was obtained after the mating of the 1,040 animals from the base population (G0), producing 1,000 offspring, about a half of each sex. After G1 was obtained, the same selection criteria simulated in E1 (BLUP, GBLUP, wGBLUP e TBV) were employed to select the parents of the next generations (G2 to G15). In order to select the parents of generation **i+1**, it was assumed that the parents of generation **i** were available for reproduction, as well as the animals from generation **i.**

For all scenarios, it was assumed that own performance records would be available before the animals were available for reproduction (and thus considered as selection candidates), similarly as occurs with weight traits and growth rate in cattle. The fixed replacement rate strategy (v1), previously described, was adopted for all scenarios.

At each generation, under BLUP scenario, all new information available in terms of phenotypes (n=1000) and pedigree was considered in the predictions before selection of the parents of the next generation.

For the scenarios under genomic selection, 500 animals from G0 and 500 animals from G1 were randomly chosen to compose the initial reference population considered for the estimation of marker effects, being that predictions under GBLUP and wGBLUP were obtained as described previously for E1. At each generation, 200 animals were selected to be added to the reference population, to contribute to update marker effect estimates. Two different strategies were adopted to genotype animals to update the reference population: genotyping the 20% with the largest phenotypic values (g1) genotyping the upper 10% and lower 10% extremes for phenotypic value (g2) . Thus, beside the scenarios under selection by BLUP or TBV, four genomic selection scenarios were simulated (GBLUP_g1, GBLUP_g2, wGBLUP_g1 e wGBLUP_g2). It was assumed that all selection candidates were genotyped, similarly as described in the case of E1 and E2.

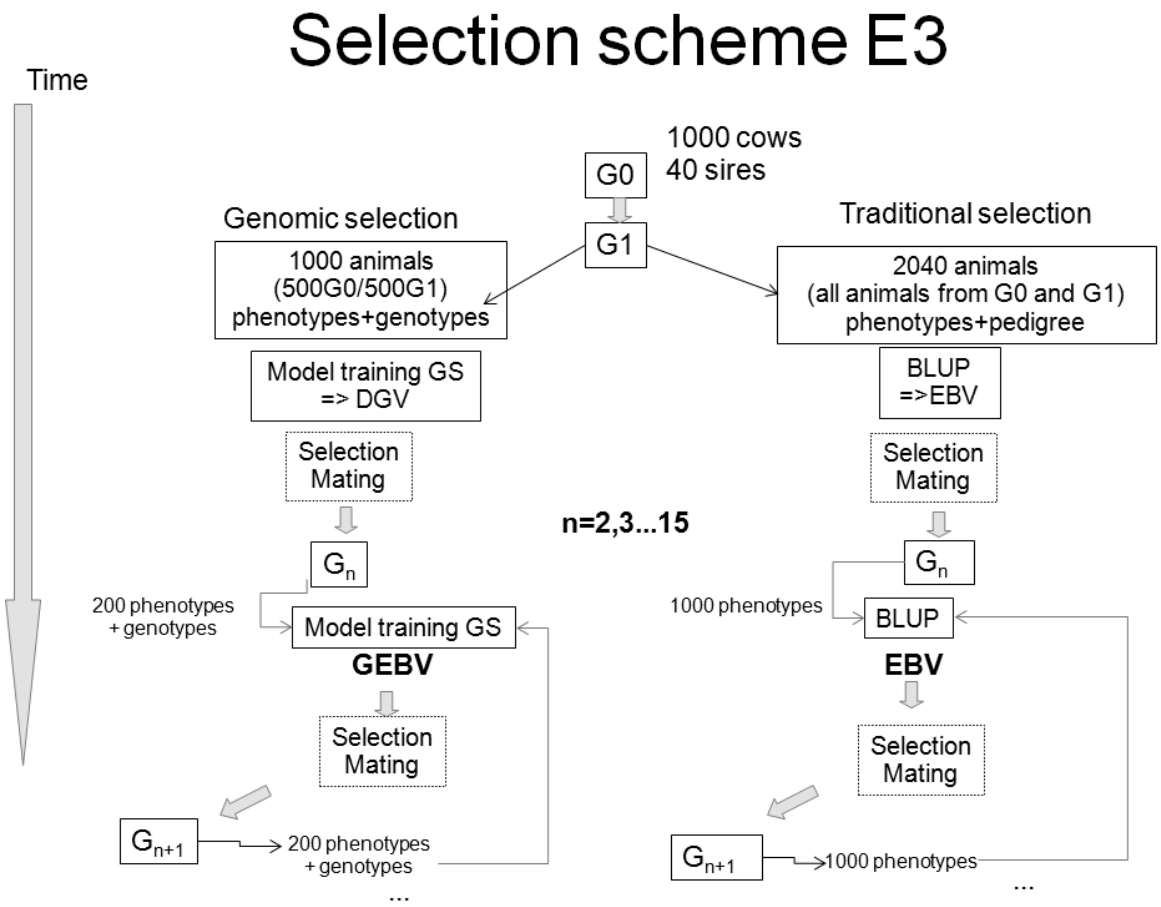


Figure 3. Fluxogram of the simulation of selection for growth (E3)
Each generation (n=1,2,...15) consisted of 1000 animals (500 males and 500 females). It was assumed that own performance information was available before selection decisions.
In the case of GS, two alternative scenarios (GBLUP and wGBLUP) were simulated for the computation of genomic predictions. In combination with them, two alternative scenarios were simulated for genotyping strategies: the reference population was updated with information of extreme 10% animals (g1) or top 20% (g2) for phenotypic value.
In all scenarios, a fixed replacement rate of 20% was simulated (for both males and females).

Criteria for analysis of results

For each scenario and replicate, the following statistics were computed at each generation:

-deltaG: accumulated genetic gain, computed as the difference between the average TBV at generation **i** and the average TBV in the respective base population (G0), standardized by the SD of TBVs in this base population.

-SD.TBV: standard deviation of the true breeding values (TBV) at generation **i**.

aveL: average generation coefficient of the parents of the animals from generation **i**, computed as a proxy for generation interval in each scenario.

-IBD: Percentage of genome that is homozygosis due to identity-by-descent, calculated as in Sonesson et al. [12]. For each locus simulated to monitor IBD, homozygosis was computed as fj = $\sum f_{kj}^2$, where f_{kj} is the frequency of the founder allele **k** at locus **j** in the generation **i**. The IBD coefficient was calculated as the average of fj over all monitored loci.

-Fped: average of pedigree-based inbreeding coefficient, over all animals from generation **i**.

-HML: Percentage of homozygosis for alleles potentially associated to deleterious mutations. Calculated as the average of the homozygosis for the lower frequency variants, i.e. for all "homozygous mutation loci" [27].

-Number of favorable alleles lost (Nlost): the accumulated number of favorable alleles lost along the selection process.

At each generation, all statistics were computed separately for each replicate (population). Aiming to enhance more objective comparisons among scenarios, the results of each statistic were analyzed by fitting the following linear model, considering the data regarding to selection for a same trait:

$$\mathbf{s_{ij}= c_i + p_j + e_{ij},}$$

where $\mathbf{s_{ij}}$ is the j-th observation for statistic **s** in the i-th scenario, $\mathbf{c_i}$ is the effect of the i-th scenario on s, $\mathbf{p_j}$ is the effect of the j-th base population and $\mathbf{e_{ij}} \sim N(0,\sigma^2 e)$. The assumptions of normally distributed and homoscedastic residuals were checked using Shapiro-Wilk and Breusch-Pagan tests, respectively.

For each statistic, generation and trait, least squares means of each scenario were contrasted using t tests, adjusted for false discovery rate (α=5%). In the case of IBD, Fped and HML, results of each statistic were log transformed in order to meet the assumptions of the linear model.

Association between different estimators of inbreeding

Aiming to evaluate the feasibility of using SNP marker information to monitor inbreeding incidence, the panels simulated in this study were employed to compute the following estimators of inbreeding at the level of individual:

-Fhet: statistic based on excess SNP homozygosity, so that: Fhet = $O(H_j)$-E(H)/[m-E(H)], where $O(H_j)$ is the observed homozygosis over all loci for individual **j**, E(H) is the expected homozygosis considering all sampled individuals and **m** is the number of markers considered [36].

-F_{ROH}: statistic based on the estimated proportion of the genome located in runs of homozygosity (ROH). ROH are defined as segments comprised by continuously homozygous SNPs, spanning at least a certain length (L), which are present in an animal due to parents transmitting identical haplotypes to their offspring [36]. For each animal, F_{ROH} was calculated as the ratio between the proportion of the genome located in ROH and the total genome length. The length of a ROH is associated to the number of generations **g** since the last common ancestor, so that ROH length follow an exponential distribution with mean equal to 1/2**g** Morgans [27]. For example, the expected length of ROH segments due to common ancestors at 2, 5 and 10 generations in the past are 0,25M, 0,1M and 0,05M, respectively (or approximately 25Mb, 10Mb and 5Mb).

The detection of ROH was carried out through an algorithm based on sliding windows of 1000kb along the genome, such that to be in a ROH a region must had at least 15 consecutive markers in homozygosis as well as a minimum density of 1 SNP at each 500 kb. To compute F_{ROH}, three different minimum length thresholds were applied to define ROHs (L = 2Mb, 4Mb or 8Mb), so that three inbreeding estimators were obtained (F_{ROH2}, F_{ROH4} e F_{ROH8}), what theoretically would allow to detect autozygous regions due to common ancestors at 25, 12 and 6 generations in the past, respectively. Both Fhet and F_{ROH} were obtained using PLINK software [37].

At each generation, the Pearson's correlation between the different estimators of inbreeding were computed, including Fhet, F_{ROH}, as well as individual estimates based on pedigree information (Fped), homozygosis due to identity-by-descent (using IBD loci) and homozygosis at HML loci. Such strategy would allow to evaluate the correspondence between different inbreeding estimators (based either on marker information or pedigree) and the estimated proportions of identity-by-descent and homozygosis for alleles potentially associated to deleterious mutations. Due to the extra computing time required in determination of ROH, these different inbreeding estimators were compared only considering data of scheme E3, for the scenarios

under selection by TBV, BLUP, as well as for the scenarios under GBLUP and wGBLUP in which extreme animals were selected to update the reference population.

Results

Scheme E1 (Selection for female reproduction)

Trends verified for the different statistics along the generations

-Genetic progress:

For a same selection criterion, greater genetic progress was verified under the scenarios with variable replacement rate, being that the advantage over scenarios with fixed replacement rate reached 20% in the last generation, for GBLUP and wGBLUP (Figure 4).

This greater genetic progress achieved with variable replacement rate is consistent with the smaller generation interval (aveL) verified in the scenarios with this replacement strategy (Figure 4). Conversely to what was observed for selection based on BLUP (for which aveL remained nearly constant since generation 7), under genomic selection there was a slight trend of reduction in aveL until the last generations, what could be possibly related to a increase in prediction accuracy, as a consequence of the increase in the size of the reference population. Since generation 7, selection based on BLUP resulted in larger generation intervals (Figure 4), what is also associated to the smaller genetic progress in this scenario compared to GBLUP and wGBLUP.

-Genetic diversity and inbreeding:

When compared to the remainder scenarios, the selection based on TBV and without constraint on the replacement rate resulted in much more pronounced reduction in the variability of breeding values (SD.TBV) (Figure 4). Starting from generation 7, the increase in the average inbreeding coefficients based on IBD loci was larger in the scenarios under BLUP selection, especially when this method was applied without constraint on the replacement rate (v2), reaching values close to 15% in generation 15. The smallest inbreeding levels (according to IBD) were achieved under selection based on TBV (Figure 4).

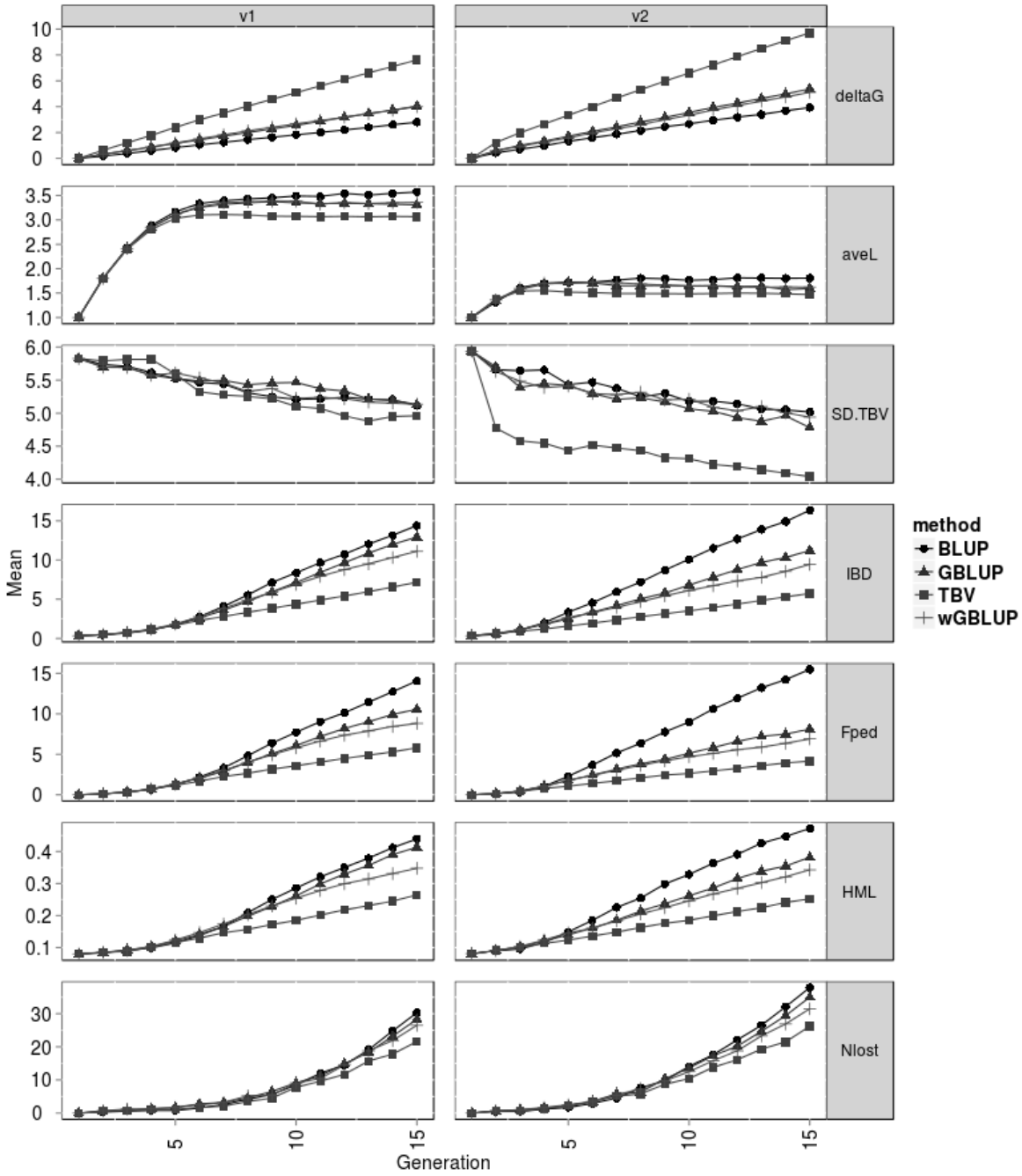


Figure 4. Trends for genetic progress and genetic diversity* under different scenarios of selection for female reproduction**

*deltaG = accumulated genetic gain (in units of additive SD of the base population); aveL = average generation interval; SD.TBV = 10*SD of true breeding values (TBV); IBD = proportion of genome that is homozygous due to identity-by-descent, in %; Fped = average of individual pedigree-based inbreeding coefficients, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Nlost= accumulated number of favorable alleles lost. Averages of 10 replicates are plotted for each statistic and generation. **Each scenario is defined by combination of: a) replacement strategy (replacement rate fixed at 20% or variable rate, v1 and v2, respectively) and b) selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP).

Trends for homozygosity at HML loci and average of pedigree-based inbreeding coefficients were very similar to those verified for IBD, being that the largest values were equal to 0.45% and 15%, respectively, for BLUP at generation 15 (under variable replacement rate). The number of favorable alleles lost exhibited a quadratic trend, with more pronounced losses after generation 10 (Figure 4).

Contrasts between averages of scenarios, at generation 15

For the sake of brevity, and since the focus of this study is on long-term consequences of the different strategies under investigation, detailed results for contrasts between scenarios at generation 15 are presented in this section.

-Replacement strategy:

For a same selection criterion, the allowance of variable replacement rate (v2) implied in genetic gain significantly larger than that verified for scenarios with replacement rate fixed at 20% (v1) (Table1). The genetic gain accumulated until the generation 15 was about 40% greater for v2 under selection based on BLUP, while such advantage was slightly inferior when genomic selection was applied (genetic gain about 30% greater for scenarios without constraint on the replacement rate).

Table 1. Genetic progress and genetic variability[1] under different scenarios[2] of selection for female reproduction at generation 15

		deltaG			SD.TBV			aveL		
Repl	Criterion	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
v1	BLUP	2.79	0.15	e	0.51	0.02	a	3.57	0.03	a
v1	GBLUP	4.04	0.12	d	0.51	0.01	a	3.31	0.02	b
v1	TBV	7.61	0.08	b	0.50	0.01	ab	3.07	0.00	c
v1	wGBLUP	4.00	0.17	d	0.51	0.01	a	3.36	0.03	b
v2	BLUP	3.93	0.15	d	0.50	0.01	a	1.81	0.03	d
v2	GBLUP	5.37	0.10	c	0.48	0.01	b	1.60	0.02	e
v2	TBV	9.70	0.12	a	0.40	0.01	c	1.47	0.01	f
v2	wGBLUP	5.11	0.08	c	0.49	0.01	ab	1.62	0.01	e

[1] deltaG = accumulated genetic gain (in units of additive SD of the base population); SD.TBV = standard deviation of true breeding values; aveL = average generation interval; ȳ = average of 10 replicates, SE= standard error. *For a same statistic and generation, averages with the same letter do not differ (t-test, adjusted p-value > 0.05).

[2] Each scenario is defined by combination of: a) replacement strategy: fixed rate (20%, v1) or variable rate (v2) and b) selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP).

Compared to the use of fixed replacement rate, a variable replacement rate resulted in smaller genetic variability for the scenarios under selection based on TBV

and GBLUP. At generation 15, for TBV and GBLUP, the SD_TBV was 19% and 7% smaller for v2 compared to v1, respectively.

Except for selection based on TBV, the variable replacement rate resulted in significantly larger inbreeding incidence, when Fped and IBD were considered in the first generations of selection (data not shown). Although the averages of Fped and IBD were still larger at generation 15, under BLUP selection, the difference between v1 and v2 was not significant ($p > 0.05$) (Table 2).

Under genomic selection, the results for inbreeding coefficients tended to be opposite to those verified for BLUP, being that smaller inbreeding levels were obtained in the scenarios with allowance of variable replacement rate, although some of these differences were not significant (Table 2). At generation 15, the variable replacement rate resulted in averages of Fped significantly smaller after selection based on GBLUP and wGBLUP, when compared to scenarios under fixed replacement, as well smaller values of Fped and IBD after selection based on TBV.

-BLUP vs genomic selection (GS):

Genomic selection (GBLUP and wGBLUP) resulted in genetic progress significantly greater than selection based on BLUP (Table 1), so that, under the fixed replacement rate, the genetic progress was about 40% greater for genomic selection, while it was between 25% and 35% greater in the scenarios without constraint on the replacement rate.

As a general rule, GS methods resulted in pedigree-based inbreeding coefficients significantly smaller than BLUP (between 25% and 54% smaller), regardless of the replacement strategy. However, at generation 15, under fixed replacement rate, only wGBLUP resulted in estimates significantly lower than BLUP for estimates of IBD, HML and number of alleles lost (Nlost) (Table 2).

-Weighting on low frequency favorable alleles (wGBLUP vs GBLUP):

No significant difference between GBLUP and wGBLUP was verified in terms of genetic progress (Table 1). Regarding to inbreeding incidence, wGBLUP resulted in smaller averages of Fped and HML, when compared to GBLUP (-15% e -16%, respectively), under fixed replacement rate. Under variable replacement rate, wGBLUP resulted in average IBD coefficients 15.5% smaller than GBLUP (Table 2),

although differences for the remainder statistics related to inbreeding were not significant.

Table 2. Inbreeding incidence[1] under different scenarios[2] of selection for female reproduction at generation 15

		IBD(%)			HML(%)			Fped(%)			Nlost		
Repl	Scenario	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
v1	BLUP	14.10	0.07	ab	0.43	0.06	ab	13.77	0.07	a	30.30	1.69	cd
v1	GBLUP	12.46	0.09	bc	0.41	0.05	ab	10.22	0.08	b	28.30	1.83	cde
v1	TBV	7.11	0.04	e	0.26	0.03	d	5.77	0.03	e	21.60	0.95	f
v1	wGBLUP	10.90	0.06	cd	0.34	0.06	c	8.67	0.06	c	26.50	1.40	de
v2	BLUP	15.95	0.08	a	0.46	0.07	a	15.02	0.09	a	37.90	1.32	a
v2	GBLUP	11.10	0.04	c	0.38	0.03	bc	8.03	0.04	cd	35.00	2.02	ab
v2	TBV	5.71	0.03	f	0.25	0.04	d	4.15	0.03	f	26.20	2.08	e
v2	wGBLUP	9.38	0.04	d	0.34	0.03	c	6.85	0.04	d	31.50	1.34	bc

[1] IBD = proportion of genome that is homozygous due to identity-by-descent, in %; Fped = average of individual pedigree-based inbreeding coefficients, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Nlost= accumulated number of favorable alleles lost. ȳ = average of 10 replicates, SE= standard error. *For a same statistic and generation, averages with the same letter do not differ (adjusted p-value > 0.05).

[2] Each scenario is defined by combination of: a) Repl: fixed replacement rate (20%, v1) or variable rate (v2) and b) selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP).

Scheme E2 (Selection for meat quality traits)

Trends verified for the different statistics along the generations

-Genetic progress:

Similarly as in the case of selection for female reproduction (scheme E1), greater rates of genetic progress were achieved in the scenarios under variable replacement rate (v2) compared to the scenarios under fixed replacement rate (v1) Figure 5). The trends for generation interval (aveL) were also similar to those verified for scheme E1, so that aveL was kept nearly constant starting from generation 7, under selection by BLUP, while under genomic selection it still decreased during this period (Figure 5).

The selection based on TBV, under variable replacement rate (v2), resulted in more accentuated reduction in the variability of true breeding values (SD_TBV), when compared to the remainder scenarios (Figure 5). Because the trends of inbreeding incidence were also similar to those obtained for scheme E1, they will not be detailed here.

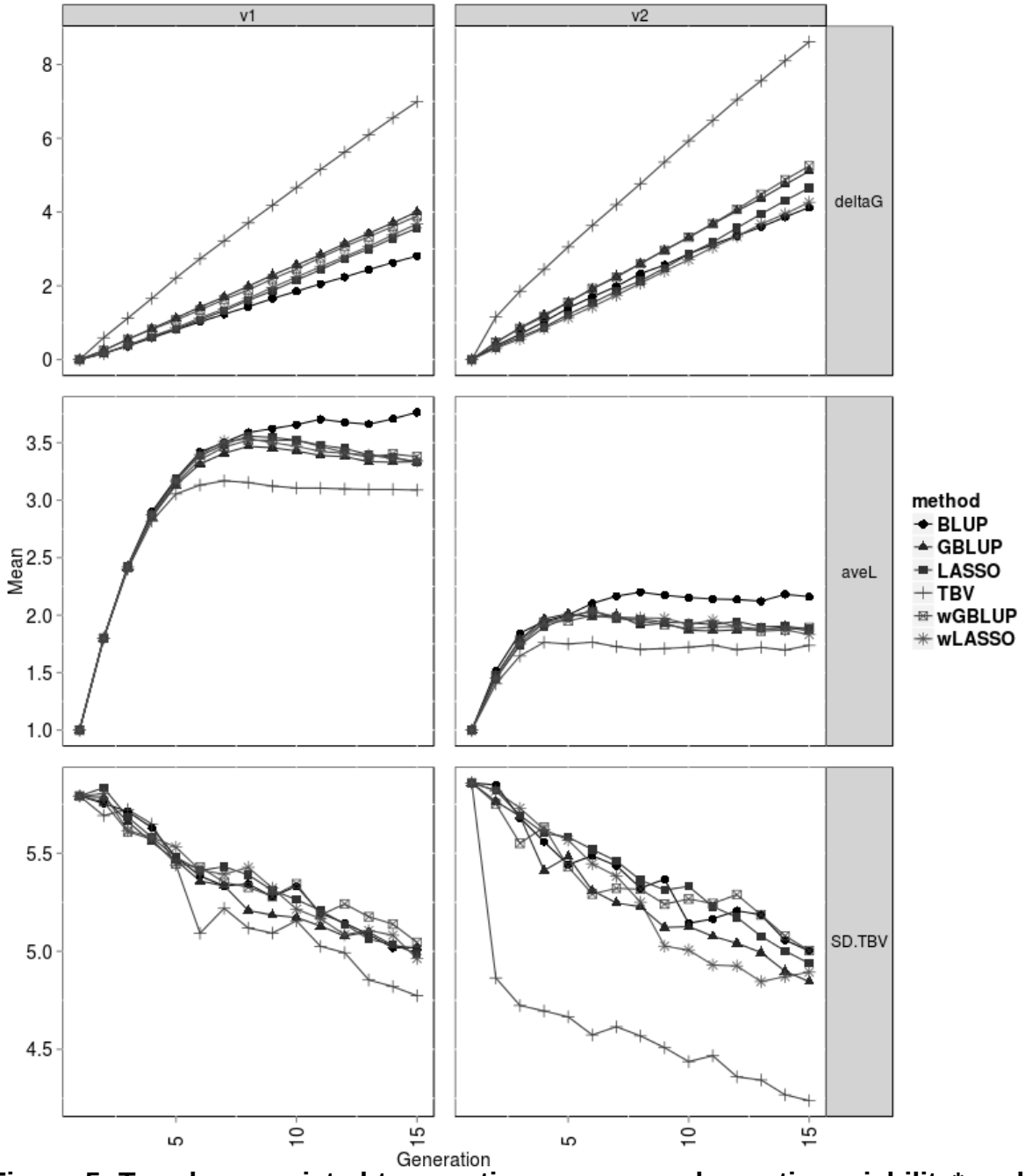


Figure 5. Trends associated to genetic progress and genetic variability* under different scenarios of selection for meat quality (trait A)**

*deltaG = accumulated genetic gain (in units of additive SD in the base population); aveL = average generation interval; SD.TBV = 10*SD of true breeding values (TBV); Averages of 10 replicates are plotted for each statistic and generation.

**Each scenario is defined by combination of: a) replacement rate (v1 = rate fixed at 20%; v2= variable rate) and b) selection criterion: (TBV), estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP or LASSO) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO). In these scenarios, selection for a trait influenced by 1,000 QTL was simulated (trait A).

Regardless of the scenario of genetic architecture, GS outperformed BLUP in terms of genetic progress (Figure 6).

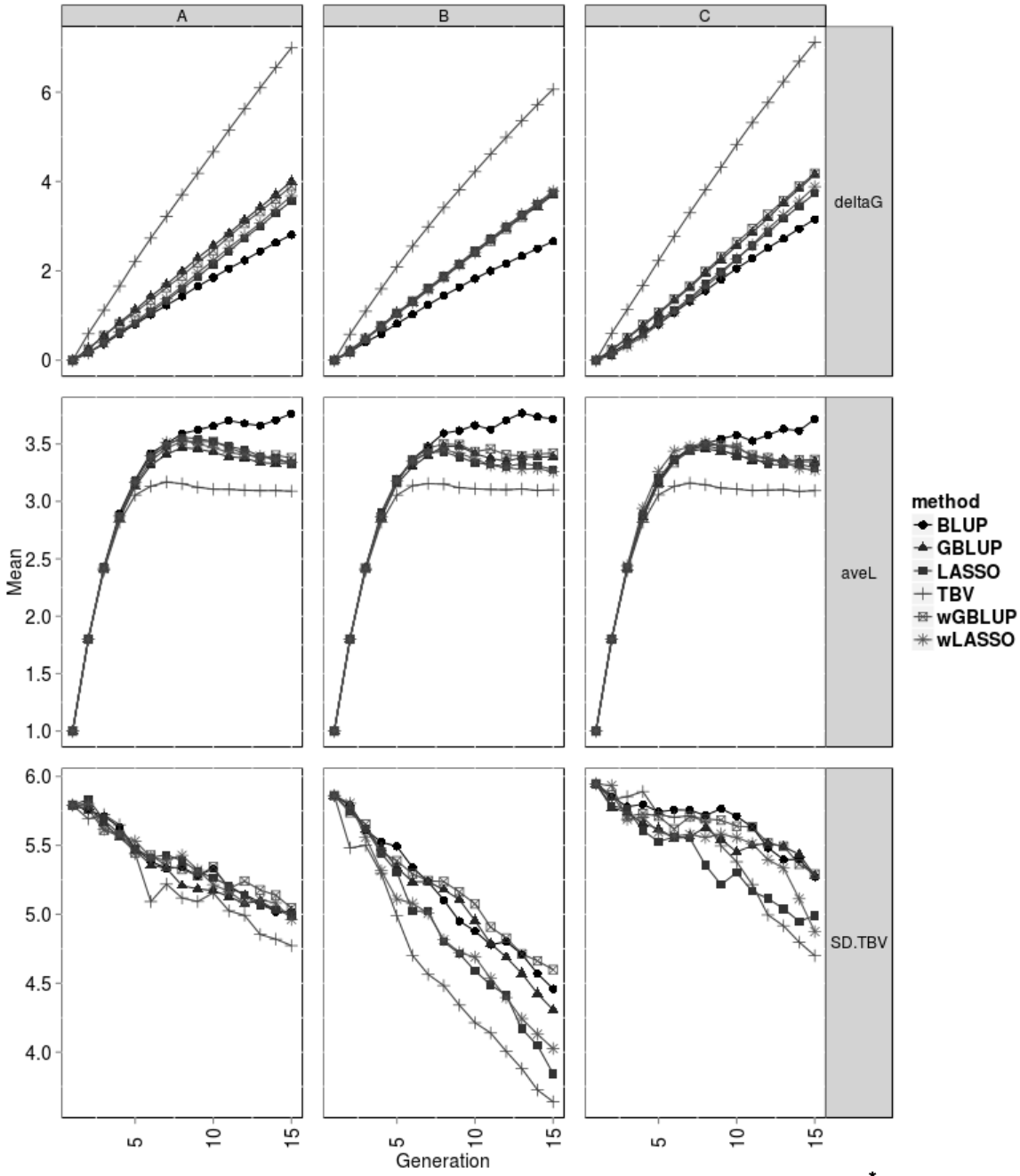


Figure 6. Trends of genetic progress and genetic variability[*] for different scenarios[**] of selection for meat quality traits (A, B and C)[***]

[*]deltaG = accumulated genetic gain (in units of additive SD of the base population); aveL = average generation interval; SD.TBV = SD of true breeding values (TBV); Averages of 10 replicates plotted for each statistic and generation. [**]Each scenario is defined by a selection criterion: TBV, EBV (BLUP), genomic selection using GBLUP or LASSO (GBLUP or LASSO) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO). [***]trait A(1000 QTL), trait B (100 QTL), trait C(1,000 QTL, 5 had lager effect, explaining 7% of the additive variance). A fixed replacement rate of 20% was adopted (v1).

For trait B (100 QTL), the similarity between different GS methods in terms of genetic progress was evident, while a slightly difference between them was observed for traits A and C (Figure 6). When GS applied, a trend of reduction in the generation interval (aveL) was observed, especially after generation 10. In terms of deltaG, GBLUP outperformed LASSO trend, for traits A and C (Figure 6).

Genetic diversity and inbreeding:

For all simulated traits, BLUP selection resulted in more pronounced increase in inbreeding coefficients (IBD and Fped) (Figure 7), while the selection for TBV resulted in the smallest inbreeding incidence. For trait A, LASSO and wLASSO resulted in more inbreeding than the other genomic selection methods. For traits B and C, lower inbreeding incidence was verified for wGBLUP and wLASSO, when compared to GBLUP and LASSO, respectively (Figure 7).

Contrasts between averages of scenarios, at generation 15

Replacement strategy:

When the replacement strategies were contrasted (v1 vs v2), significantly larger genetic progress was verified for scenarios with variable replacement rate (Table 3), confirming the same trend verified in the scheme mimicking selection for female reproduction. Such contrast also suggested more inbreeding incidence (Fped, IBD and HML) under variable replacement rate, when a same selection criterion was considered (Table 4).

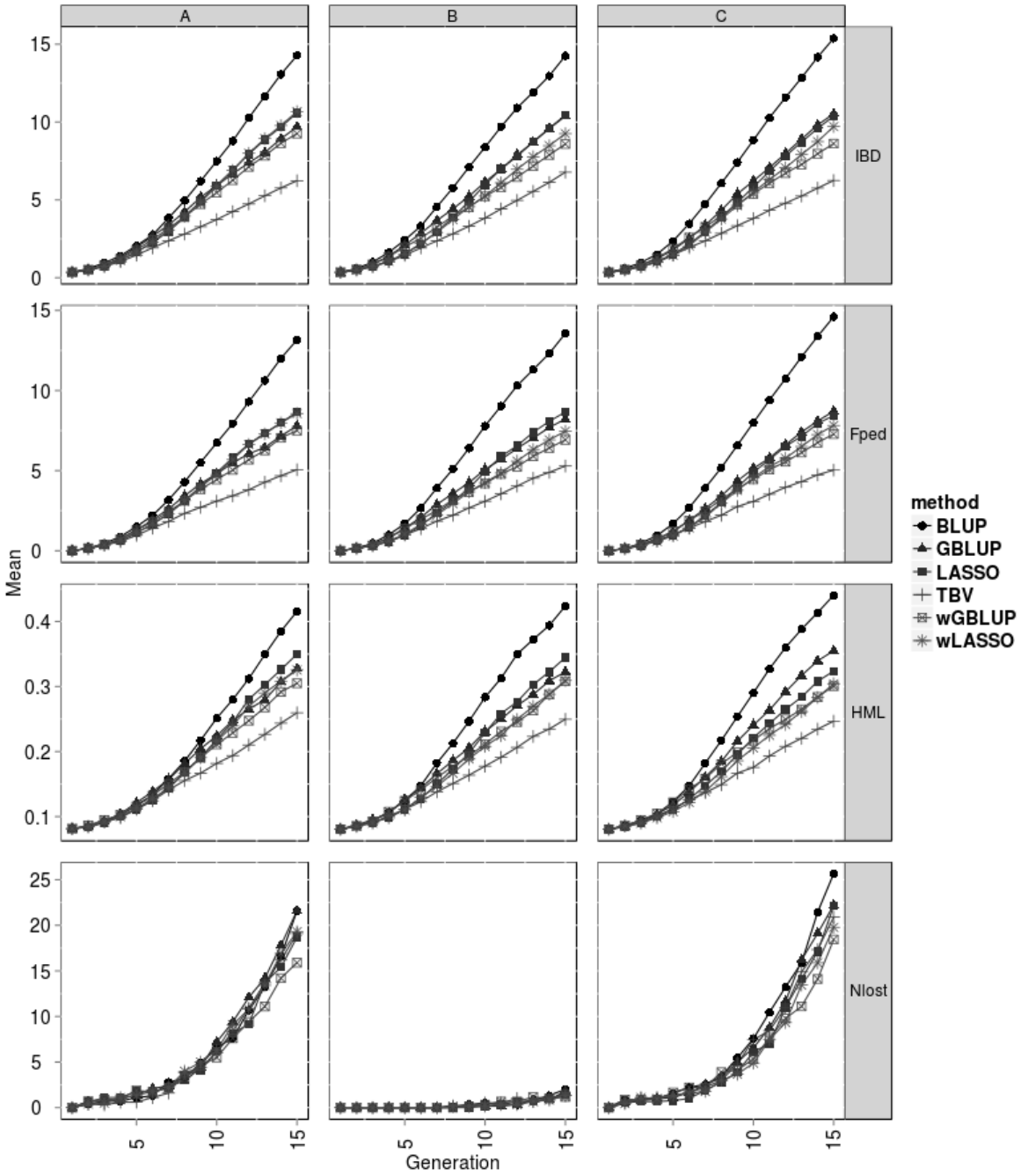


Figure 7. Trends for genetic diversity* under different scenarios of selection for meat quality traits (A, B and C)*****

* IBD = proportion of genome that is homozygous due to identity-by-descent, in %; Fped = average of individual pedigree-based inbreeding coefficients, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Nlost= accumulated number of favorable alleles lost. Averages of 10 replicates are plotted for each statistic and generation.

**Selection criteria: true breeding value (TBV), estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP or LASSO) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO).

***Simulated traits differed in genetic architecture: trait A(1000 QTL), trait B (100 QTL), trait C(1,000 QTL, five of which had larger effect and explained 7% of the additive variance). A fixed replacement rate of 20% was adopted (v1).

Table 3. Genetic progress and genetic variability[1] under different scenarios[2] of selection for meat quality at generation 15, for different replacement rates

		deltaG			SD.TBV			aveL		
Repl	Method	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
v1	BLUP	2.81	0.13	i	0.50	0.01	ab	3.76	0.07	a
v1	GBLUP	4.01	0.07	efg	0.50	0.01	ab	3.34	0.02	b
v1	LASSO	3.56	0.12	h	0.50	0.01	ab	3.33	0.02	b
v1	wGBLUP	3.89	0.14	fgh	0.50	0.01	a	3.38	0.03	b
v1	wLASSO	3.68	0.09	gh	0.50	0.01	abc	3.34	0.03	b
v1	TBV	6.99	0.08	b	0.48	0.01	c	3.09	0.01	c
v2	BLUP	4.13	0.20	ef	0.50	0.01	ab	2.16	0.04	d
v2	GBLUP	5.11	0.10	c	0.48	0.01	bc	1.86	0.01	e
v2	LASSO	4.65	0.18	d	0.49	0.01	abc	1.88	0.04	e
v2	TBV	8.61	0.08	a	0.42	0.01	d	1.74	0.01	f
v2	wGBLUP	5.25	0.11	c	0.50	0.01	ab	1.89	0.02	e
v2	wLASSO	4.27	0.19	e	0.49	0.01	abc	1.83	0.01	e

[1] deltaG = accumulated genetic gain (in units of additive SD of the base population); SD.TBV = standard deviation of true breeding values; aveL = average generation interval; ȳ = average of 10 replicates, SE= standard error. *For a same statistic, averages with the same letter do not differ (adjusted p-value > 0.05).

[2] Each scenario is defined by combination of: a) Repl = replacement strategy (fixed rate = 20%, v1) or variable rate (v2) and b) selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO). The selection for a trait influenced by 1,000 QTL was simulated (trait A).

Table 4. Inbreeding incidence[1] under different scenarios[2] of selection for meat quality at generation 15, for different replacement rates

		IBD(%)			HML(%)			Fped(%)			Nlost		
Repl	Scenario	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
v1	BLUP	14.02	0.06	a	0.41	0.06	b	12.89	0.07	a	21.6	1.72	de
v1	GBLUP	9.64	0.03	c	0.33	0.04	def	7.76	0.03	bc	21.6	1.16	de
v1	LASSO	10.36	0.07	bc	0.34	0.06	cdef	8.51	0.06	bc	18.7	1.13	ef
v1	TBV	6.18	0.04	d	0.26	0.04	g	5.03	0.04	d	19.1	1.20	ef
v1	wGBLUP	9.03	0.07	c	0.30	0.04	f	7.37	0.06	c	15.9	1.40	f
v1	wLASSO	10.33	0.09	bc	0.32	0.06	ef	8.31	0.08	bc	19.3	2.30	ef
v2	BLUP	16.25	0.06	a	0.49	0.06	a	15.00	0.06	a	36.4	1.54	a
v2	GBLUP	10.32	0.07	bc	0.36	0.07	bcd	7.88	0.07	bc	29.0	2.59	bc
v2	LASSO	11.77	0.09	b	0.41	0.08	b	8.90	0.09	b	30.1	2.95	b
v2	TBV	4.92	0.03	e	0.22	0.02	h	3.63	0.03	e	16.4	1.11	f
v2	wGBLUP	10.20	0.06	bc	0.37	0.04	bc	7.63	0.07	bc	24.6	1.78	d
v2	wLASSO	11.67	0.11	b	0.35	0.07	cde	8.66	0.09	b	25.6	1.80	cd

[1] IBD = proportion of genome that is homozygous due to identity-by-descent, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Fped = average of individual pedigree-based inbreeding coefficients, in %; Nlost= accumulated number of favorable alleles lost. ȳ = average of 10 replicates, SE= standard error. *For a same statistic, averages with the same letter do not differ (adjusted p-value > 0.05).

[2] Each scenario is defined by combination of: a) Repl = replacement strategy (fixed rate = 20%, v1) or variable rate (v2) and b) selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO). The selection for a trait influenced by 1,000 QTL was simulated (trait A).

Along the generations, there was a trend of reduction in the difference between these two replacement strategies with respect to inbreeding levels (data not shown). At generation 15, there was no difference between replacement strategies in terms of averages of IBD and Fped, under all scenarios of GS and BLUP, while significantly larger number of favorable alleles lost was verified under variable replacement rate (v2, Table 4). On the other hand, significantly larger homozygosis at HML loci was verified under fixed replacement rate, for scenarios with selection based on BLUP, LASSO and wGBLUP (Table 4).

-BLUP vs genomic selection (GS):

Regardless of genetic architecture scenario, most genomic selection scenarios resulted in genetic progress significantly greater than that verified for correspondent scenarios under BLUP selection. When 100 QTL were simulated (trait B), the advantage of GS over BLUP in terms of genetic progress varied from +30% (generation 5 and 10, data not shown) to +40% (generation 15, Table 5) .

When LASSO was used to estimate marker effects (LASSO or wLASSO) and 1,000 QTL were simulated, no significant benefit over BLUP was verified in terms of genetic progress in the first 5 generations (data not shown), regardless of whether major QTL were simulated (trait C) or not (trait A). This result is possibly related to the small size of the reference population in such period of the simulation.

The genetic gain accumulated until generation 15 for selection based on LASSO and wLASSO was significantly larger than that verified for BLUP, with increases between 20% and 30% being estimated, in the case of the traits A and C (Table 5).

Table 5. Genetic progress and genetic variability[1] under different scenarios[2] of selection for meat quality traits[3] at generation 15

		deltaG			SD.TBV			aveL		
Trait	Scenario	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
A	BLUP	2.81	0.13	e	0.50	0.01	a	3.76	0.07	a
	GBLUP	4.01	0.07	b	0.50	0.01	a	3.34	0.02	b
	LASSO	3.56	0.12	d	0.50	0.01	a	3.33	0.02	b
	TBV	6.99	0.08	a	0.48	0.01	b	3.09	0.01	c
	wGBLUP	3.89	0.14	bc	0.50	0.01	a	3.38	0.03	b
	wLASSO	3.68	0.09	cd	0.50	0.01	ab	3.34	0.03	b
B	BLUP	2.66	0.11	c	0.45	0.03	ab	3.72	0.06	a
	GBLUP	3.70	0.21	b	0.43	0.03	b	3.39	0.04	b
	LASSO	3.74	0.21	b	0.38	0.02	cd	3.28	0.03	c
	TBV	6.06	0.28	a	0.36	0.02	d	3.10	0.01	d
	wGBLUP	3.74	0.17	b	0.46	0.02	a	3.42	0.03	b
	wLASSO	3.78	0.23	b	0.40	0.02	c	3.26	0.02	c
C	BLUP	3.15	0.13	d	0.53	0.02	a	3.72	0.04	a
	GBLUP	4.16	0.18	b	0.53	0.02	a	3.35	0.02	bc
	LASSO	3.73	0.11	c	0.50	0.02	ab	3.29	0.03	cd
	TBV	7.11	0.23	a	0.47	0.01	b	3.09	0.01	e
	wGBLUP	4.18	0.21	b	0.53	0.01	a	3.37	0.02	b
	wLASSO	3.88	0.21	bc	0.49	0.01	b	3.27	0.02	d

[1]deltaG = accumulated genetic gain (in units of additive SD of the base population); SD.TBV = SD of true breeding values(TBV); aveL = average generation interval; ȳ = average of 10 replicates, SE= standard error. *For a same statistic and trait, averages with the same letter do not differ (adjusted p-value > 0.05).[2]Each scenario is defined by a selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP or LASSO) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO). [3]trait A(1000 QTL), trait B (100 QTL), trait C (1,000 QTL, five of which had lager effect and explained 7% of the additive variance). A fixed replacement rate of 20% was adopted.

As a general rule, for all scenarios of genetic architecture, lower levels of inbreeding were obtained under genomic selection compared to BLUP, although these differences were not significant in the first generations simulated (data not shown). In generation 15, the averages of the pedigree-based inbreeding coefficients were between 34% (LASSO) and 42% (wGBLUP) lower under genomic selection scenarios, when compared to BLUP, in the scenario in which more polygenic background was simulated (trait A, Table 6).

Table 6. Inbreeding incidence[1] under different scenarios[2] of selection for meat quality traits[3] at generation 15

		IBD(%)			HML(%)			Fped(%)			Nlost		
Trait	Method	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
A	BLUP	14.02	0.06	a	0.41	0.06	a	12.89	0.07	a	21.60	1.72	a
	GBLUP	9.64	0.03	b	0.33	0.04	bc	7.76	0.03	bc	21.60	1.16	a
	LASSO	10.36	0.07	b	0.34	0.06	b	8.51	0.06	b	18.70	1.13	ab
	TBV	6.18	0.04	c	0.26	0.04	d	5.03	0.04	d	19.10	1.20	ab
	wGBLUP	9.03	0.07	b	0.30	0.04	c	7.37	0.06	c	15.90	1.40	b
	wLASSO	10.33	0.09	b	0.32	0.06	bc	8.31	0.08	bc	19.30	2.30	ab
B	BLUP	13.81	0.08	a	0.41	0.08	a	13.11	0.09	a	3.00	0.76	a
	GBLUP	10.21	0.06	b	0.32	0.05	b	8.09	0.06	bc	2.40	0.48	ab
	LASSO	10.08	0.09	bc	0.34	0.06	b	8.39	0.08	b	2.10	0.59	ab
	TBV	6.73	0.04	d	0.25	0.03	c	5.25	0.04	d	1.40	0.48	b
	wGBLUP	8.54	0.04	c	0.31	0.04	b	6.88	0.03	c	1.50	0.34	b
	wLASSO	8.96	0.08	bc	0.30	0.07	b	7.29	0.07	bc	1.90	0.59	ab
C	BLUP	15.24	0.04	a	0.44	0.04	a	14.52	0.04	a	25.67	1.55	a
	GBLUP	10.36	0.06	b	0.35	0.06	b	8.55	0.07	b	22.22	1.61	ab
	LASSO	10.24	0.06	b	0.32	0.05	bc	8.33	0.06	bc	22.11	1.20	abc
	TBV	6.21	0.04	d	0.25	0.04	d	5.01	0.04	d	20.89	1.51	bc
	wGBLUP	8.45	0.06	c	0.30	0.03	c	7.14	0.07	c	18.44	1.31	c
	wLASSO	9.41	0.08	bc	0.30	0.07	c	7.54	0.09	bc	19.78	2.00	bc

[1] IBD = proportion of genome that is homozygous due to identity-by-descent, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Fped = average of individual pedigree-based inbreeding coefficients, in %. Nlost= accumulated number of favorable alleles lost. ȳ = average of 10 replicates, SE= standard error. *For a same statistic and trait, averages with the same letter do not differ (adjusted p-value > 0.05).
[2] Each scenario is defined by a selection criterion: (TBV), estimated breeding value (BLUP), genomic estimated breeding value using GBLUP or LASSO (GBLUP or LASSO) and GBLUP or LASSO giving more weight to low-frequency favorable alleles (wGBLUP or wLASSO).
[3]Trait A(1000 QTL), trait B (100 QTL), trait C(1,000 QTL, five of which had lager effect and explained 7% of the additive variance). A fixed replacement rate of 20% was adopted.

For trait A, over all GS methods, the averages of Fped, HML and IBD were 38%, 21% and 30% lower, respectively, than those verified for BLUP (Table 6). The advantage of GS methods over BLUP in terms of inbreeding was even greater when a smaller number of QTL (trait B) or major QTL were simulated (trait C) (Table 6). For trait B, the averages of Fped, HML and IBD across GS methods were 42%, 23% and 32% lower, respectively, than those of BLUP (Table 6), while when major QTL were simulated, the correspondent figures were 46%, 28% and 37% lower, respectively (Table 6).

-*Variable selection methods vs GBLUP:*

For the more polygenic trait (trait A), the selection based on LASSO prediction resulted in genetic progress between 16% and 11% lower than GBLUP, being these differences significant (Table 5). At generation 15, LASSO also resulted in inbreeding

coefficients larger than GBLUP (+7% and +10% for IBD and Fped, respectively), although such differences were not significant (Table 6).

When 100 QTL were simulated (trait B), there was no significant difference between LASSO and GBLUP in terms of genetic progress (Table 5). In the first generations, the averages for deltaG were slightly lower for LASSO, what could suggest that this method was more sensitive to the small size of the reference population. As more animals were included in the reference population, the genetic progress was slightly larger for LASSO when compared to GBLUP (Table 5). For this same trait, there was no significant difference between LASSO and GBLUP in terms of inbreeding incidence (Table 6), although the variability of true breeding values was significantly lower (-11%) for LASSO in the same period (Table 5).

Conversely to what would be expected, when major QTL were simulated (trait C), LASSO selection resulted in genetic gain significantly lower than GBLUP (difference about 10%) (Table 5). In this situation, while inbreeding incidence was significantly lower for LASSO in the first generations simulated, such differences tended to diminish along the selection process and there was no difference between LASSO and GBLUP in terms of inbreeding incidence at generation 15 (Table 6).

-Weighting on low frequency favorable alleles (wGBLUP and wLASSO):

Regardless of the simulated genetic architecture, there was no significant difference between GBLUP and wGBLUP in terms of genetic gain, what was also verified when LASSO and wLASSO were contrasted (Table 5).

For the more polygenic trait, although the averages of Fped and IBD were slightly lower for wGBLUP and wLASSO when compared to GBLUP and LASSO, respectively, such differences were not significant (trait A, Table 6). In this situation, the only benefit of attributing more weight to low-frequency favorable alleles was verified for wGBLUP with respect to the number of favorable alleles (Nlost 26% lower at generation 15, when compared to GBLUP).

When 100 QTL were simulated, inbreeding incidence was significantly lower under wGBLUP in the last generations, so that the average for IBD was 16% smaller, when compared to GBLUP at generation 15 and SD_TBV was 7% larger for wGBLUP in the same situation (trait B, Table 6). No significant difference was verified between wLASSO and LASSO for all statistics related to inbreeding incidence,

although the averages of Fped, IBD and HML were slightly lower for wLASSO (trait B, Table 6).

When major QTL were simulated (trait C), wGBLUP resulted in significantly lower inbreeding incidence when compared to GBLUP in the last generation (Table 6). The averages of Fped, HML and IBD were 16.5%, 14.5% e 18.4% lower for wGBLUP, respectively, while the number of favorable alleles lost was 17% lower. There was no significant difference between LASSO and wLASSO in terms of inbreeding incidence, even though wLASSO resulted in slightly lower averages of Fped and IBD.

Scheme E3 (Selection for a growth trait)

Trends verified for the different statistics along the generations

-Genetic progress and genetic diversity:

In Figure 8, the averages for statistics related to genetic progress and inbreeding are presented for the scheme of selection for a growth trait. It must be noticed that genomic predictions were based on a amount of phenotypic information between 2-fold and 4-fold smaller than that employed to compute BLUP predictions.

Given the simulated heritability and the smaller amount of information available for genomic prediction, larger rates of genetic progress were observed under BLUP selection, what also resulted in smaller generation intervals for BLUP, especially after generation 7 (Figure 8).

After generation 10, some advantage of GS over BLUP could be observed with respect to inbreeding incidence. As verified for the other selection schemes, the selection based on TBV resulted in the greatest genetic progress and the smaller levels of inbreeding, among all scenarios compared. (Figure 8).

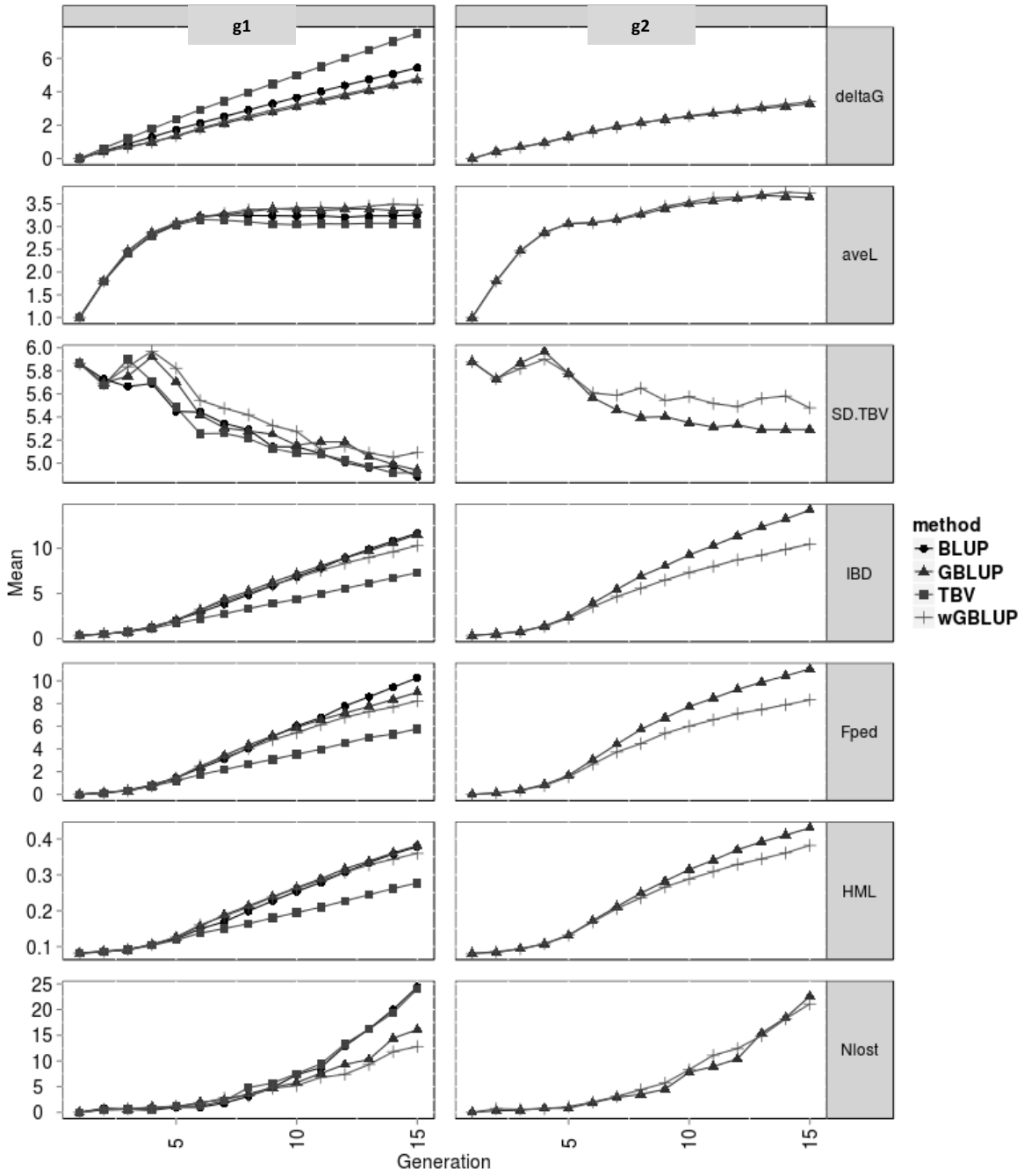


Figure 8. Trends for genetic progress and genetic diversity* under different scenarios of selection for a growth trait**

* deltaG = accumulated genetic gain (in units of additive SD in the base population); aveL = average generation interval; SD.TBV = standard deviation of true breeding values (TBV); IBD = proportion of genome in homozygosis due to identity-by-descent, in %; Fped = average of pedigree-based inbreeding coefficients, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Nlost= accumulated number of favorable alleles lost. Averages of 10 replicates are plotted. **Each scenario is defined by a selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP). For GBLUP and WGBLUP, the reference population was updated with information of extreme 10% animals (g1) or top 20% (g2) for phenotypic value. A trait influenced by 1,000 QTL, expressed in both sexes before selection, was simulated.

Contrasts between averages of scenarios, at generation 15

-BLUP vs genomic selection (GS)

For all generations considered, genomic selection resulted in genetic progress significantly lower than BLUP (Table 7). When extreme animals were genotyped to update the reference population (GBLUP and wGBLUP, g1), the difference from BLUP in terms of genetic progress tended to diminish along the selection process, so that at generations 5, 10 and 15, the accumulated genetic gain under these strategies was 20%, 15% and 12% lower, respectively, than that verified under BLUP selection (data not shown).

Table 7. Genetic progress and genetic variability[1] under different scenarios[2] of selection for a growth trait at generation 15

	deltaG			SD_TBV			aveL		
Scenario	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
BLUP	5.44	0.11	b	0.49	0.01	c	3.25	0.02	c
GBLUP/g1	4.70	0.09	c	0.49	0.01	c	3.36	0.02	c
GBLUP/g2	3.29	0.10	d	0.53	0.01	ab	3.64	0.06	a
TBV	7.51	0.09	a	0.49	0.01	c	3.07	0.00	d
wGBLUP/g1	4.76	0.09	c	0.51	0.01	bc	3.48	0.04	b
wGBLUP/g2	3.42	0.18	d	0.55	0.01	a	3.73	0.06	a

[1] deltaG = accumulated genetic gain (in units of additive SD of the base population); SD_TBV = standard deviation of true breeding values (TBV); aveL = average generation interval; ȳ = average of 10 replicates, SE= standard error. *For a same statistic and generation, averages with the same letter do not differ (adjusted p-value > 0.05).

[2] Each scenario is defined by a selection criterion: TBV, estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP). For GBLUP and WGBLUP, the reference population was updated with information of extreme 10% animals (g1) or top 20% (g2) for phenotypic value. A polygenic trait, influenced by 1,000 QTL, and expressed in both sexes before selection was simulated.

When the reference population was updated through genotyping of the top 20% animals for phenotypic value (GBLUP and wGBLUP, g2), the disadvantage over BLUP tended to increase along the generations, so that accumulated genetic gain under such GS was about 40% lower in the last generation (Table 7).

Under genotyping of the top 20% animals (g2), the average generation intervals were significantly larger than those verified under BLUP or GS with genotyping of extreme animals (g1). In terms of inbreeding incidence, smaller averages of Fped and IBD were obtained for wGBLUP, when compared to BLUP (Table 8).

Table 8. Inbreeding incidence[1] under different scenarios[2] of selection for a growth trait at generation 15

	IBD(%)			HML(%)			Fped(%)			Nlost		
Scenario	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*	ȳ	SE	*
BLUP	11.41	0.07	ab	0.37	0.05	ab	10.06	0.07	a	24.40	1.75	a
GBLUP/ g1	11.48	0.03	ab	0.38	0.03	ab	8.94	0.04	ab	16.10	1.55	b
GBLUP/ g2	13.45	0.11	a	0.42	0.06	a	10.39	0.12	a	22.60	1.79	a
TBV	7.28	0.02	c	0.28	0.03	c	5.76	0.02	c	24.00	1.45	a
wGBLUP/ g1	10.24	0.03	b	0.36	0.04	b	8.18	0.04	b	12.80	1.16	b
wGBLUP/ g2	10.18	0.08	b	0.37	0.07	ab	8.16	0.07	b	21.10	1.19	a

[1]IBD = proportion of genome that is homozygous due to identity-by-descent, in %; HML = homozygosis for alleles potentially associated to deleterious mutations, in %. Fped = average of individual pedigree-based inbreeding coefficients, in %. Nlost= accumulated number of favorable alleles lost. ȳ = average of 10 replicates, SE= standard error. *For a same statistic and generation, averages with the same letter do not differ (adjusted p-value > 0.05).

[2] Each scenario is defined by a selection criterion: true breeding value (TBV), estimated breeding value (BLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP). For GBLUP and WGBLUP, the reference population was updated with information of extreme 10% animals (g1) or top 20% (g2) for phenotypic value. A polygenic trait, influenced by 1,000 QTL and expressed in both sexes before selection was simulated.

In the scenarios with selection based on GBLUP and update of reference population after genotyping superior animals (GBLUP/g2), the inbreeding coefficients were consistently larger than those of the remainder scenarios, while selection based on GBLUP with genotyping of extreme animals did not provided any significant benefit in terms of inbreeding incidence when compared to BLUP (Table 8).

At generation 15, wGBLUP resulted in lower inbreeding than BLUP, regardless of the strategy to update the reference population, so that averages of Fped and IBD were 19% and 10% lower, respectively. GBLUP under genotyping of extreme animals also resulted in loss of favorable alleles significantly smaller when compared to BLUP.

-Strategy to update the reference population

For both GBLUP and wGBLUP, the update of the reference population based on genotyping of extreme animals resulted in larger genetic progress than genotyping only the top 20% animals, so that differences between these strategies were significant in the last generations (+40% at generation 15, Table 7). For selection based on GBLUP, inbreeding coefficients and number of alleles lost were significantly lower under genotyping of extreme animals, so that averages of Fped and IBD were about 14% lower at generation 15 (Table 8).

-Weighting on low frequency alleles

For all generations considered, attributing more weight to low frequency favorable alleles (wGBLUP) did not result in benefit in terms of genetic progress, when compared to GBLUP (Table 7). Under update of reference population through genotyping of extreme animals, wGBLUP and GBLUP did not differ in terms of inbreeding incidence, for all generations considered, although average inbreeding coefficients were slightly lower for wGBLUP (IBD about 10% lower at generation 15) (Table 8). The loss of favorable alleles was significantly smaller (-20%) for wGBLUP, when compared to GBLUP. Under genotyping of superior animals, the advantage of wGBLUP over GBLUP in terms of inbreeding incidence was larger, so that at generation 15, the averages of IBD and Fped were between 18% and 25% lower for wGBLUP.

Association between different estimators of inbreeding

The average of correlations between different estimators of inbreeding and the individual averages of homozygosis due to identity-by-descent (IBD) were calculated considering the animals born at each generation (Figure 9). Aiming to reduce the influence of progeny of founder animals on such statistics, only the results from generation 5 onward are presented. Due to monitoring of founder alleles for a large number of loci (N=3064), individual IBD estimates are expected to be very close to the real levels of homozygosis due to identity-by-descent.

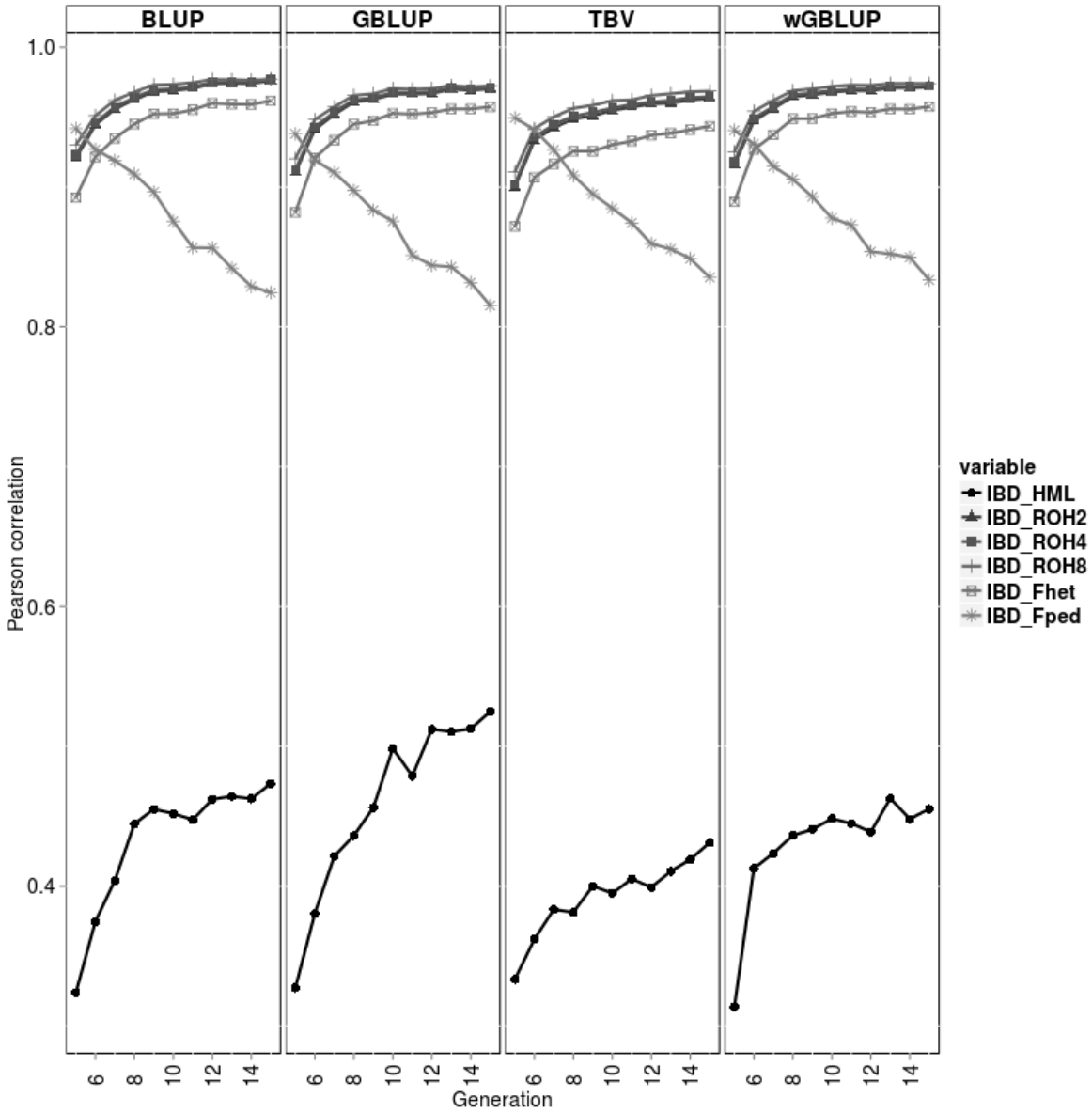


Figure 9. Correlations* between inbreeding estimators and homozygosis due to identity-by-descent (IBD) after 5 to 15 generations of selection for a simulated growth trait**

*Average of 10 replicates. **IBD = individual inbreeding coefficients estimated after monitoring founder alleles for 3,064 loci evenly spaced along the genome. HML = homozygosis for alleles potentially associated to deleterious mutations. ROH2, ROH4 and ROH8 = individual inbreeding coefficients based on the proportion of the genome located in runs of homozygosity (ROH) of length ≥ 2Mb, 4Mb and 8Mb, respectively. Fhet = individual inbreeding coefficients based on SNP homozygosity excess. Fped = individual inbreeding coefficients based on pedigree information.

Except for individual homozygosis at HML loci, all estimators of inbreeding were strongly correlated (> 0.80) to individual IBD estimates (Figure 9). The trends for correlations between statistics were consistent across selection criteria. One

important result is related to the opposite trend verified between pedigree-based and marker-based estimators with respect to their correlation with individual IBD estimates. While such correlation tended to increase along the generations for marker-based estimators, correspondent figures for pedigree-based estimator tended to decrease in the same period (Figure 9).

The correlations between estimates based on HML and IBD were lower (between 0.35 e 0.50) than those verified for the remainder estimators, highlighting the difference between homozygosis due to identity-by-descent and homozygosis at rare variants. All estimators based on ROH exhibited very similar association to individual IBD estimates and were highly correlated to each other (Figure 9). From generation 7 onward, F_{ROH} estimates were the most strongly correlated to homozygosis due to identity-by-descent, being that the correlation between Fhet and IBD was slightly lower than the correlation between F_{ROH} and IBD.

The marker-based inbreeding estimators were also strongly correlated to the pedigree-based inbreeding coefficients (Fped) (Figure 10), but there was a trend of decrease in such correlations along the selection process. The extent of association to Fped was also very similar for all estimators based on ROH, while the estimates of Fhet had slightly lower correlation with Fped than observed for all estimates of F_{ROH}.

While Fped was strongly correlated to IBD, the association between pedigree-based inbreeding coefficients and homozygosis for alleles potentially associated to deleterious mutations (HML) was much weaker ($r<0.40$) (data not shown). The marker-based estimators exhibited stronger association to HML, so that correlations between them and individual homozygosis for rare alleles at HML loci were above 0.50 in some scenarios (data not shown).

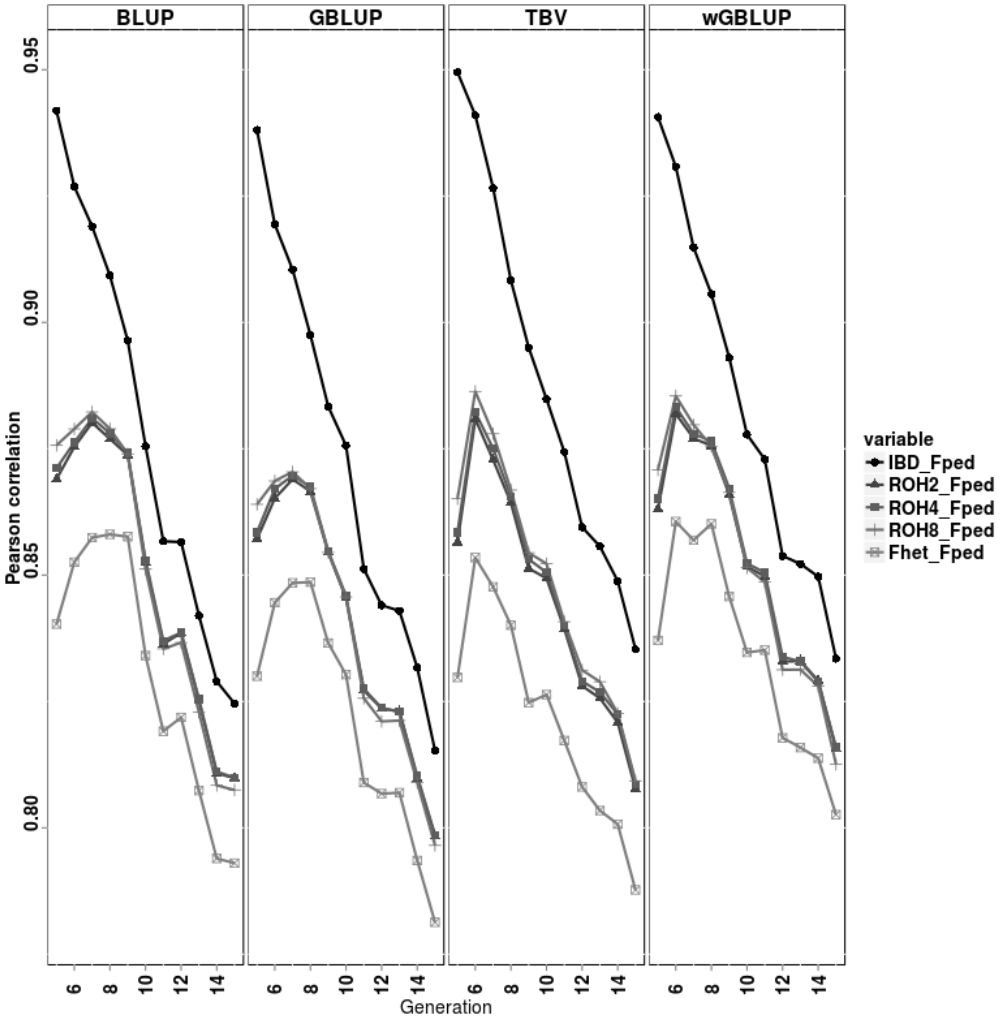


Figure 10. Correlations* between marker-based inbreeding and pedigree-based inbreeding after 5 to15 generations of selection for a simulated growth trait**

* Average of 10 replicates. **IBD = individual inbreeding coefficients estimated after monitoring founder alleles for 3,064 loci evenly spaced along the genome. HML = homozygosis for alleles potentially associated to deleterious mutations. ROH2, ROH4 and ROH8 = individual inbreeding coefficients based on the proportion of the genome located in runs of homozygosity (ROH) of length ≥ 2Mb, 4Mb and 8Mb, respectively. Fhet = individual inbreeding coefficients based on SNP homozygosity excess. Fped = individual inbreeding coefficients based on pedigree information.

Discussion

Long-term consequences of genomic selection(GS)

The relative benefit of genomic selection when compared to traditional selection based on BLUP varied in function of the selection schemes. For the schemes simulating selection for female reproduction (E1) and meat quality (E2), genomic selection allowed considerable gains in terms of genetic progress,

regardless of the time horizon considered. In these situations, the genetic gain accumulated until the last generation was up to 40% greater for GS when compared to the genetic gain achieved with selection based on BLUP.

When selection for a growth trait in beef cattle was simulated, genomic selection lead to smaller genetic gain than BLUP (between 20% and 10% lower at generations 5 and 15, respectively). This result can be explained by the larger amount of information that was simulated to be available for BLUP prediction and by the nature of the simulated trait (moderate-to-high heritability and expressed in both sexes before selection), a situation for which it is not expected large benefit for the use of genomic selection when compared to BLUP [38].

Due to computational limitations in the step of estimation of marker effects, it was not possible to simulate the inclusion of a larger proportion of genotyped animals at each generation in order to update the reference population, so that the number of phenotypic records available was between 2-fold and 4-fold smaller for GS prediction when compared to BLUP, and this possibly contributed to worsen the circumstances in which genomic selection was applied.

The relative benefit of GS over BLUP was already subject of previous studies, in which larger advantage for the application of GS was envisaged under selection for hard-to-measure traits and/or traits expressed late in life (e.g. [19], [20], [39]). However, the simulation of repeated cycles of genomic selection considering continuous update of the reference population, similarly as already occur with dairy cattle schemes, allowed to investigate other consequences of application of this technology, which hardly could be assessed by deterministic predictions, especially with respect to inbreeding incidence and genetic diversity.

As a general rule, for schemes in which greater genetic progress was achieved using genomic selection, smaller pedigree-based inbreeding coefficients were also estimated, which were relatively consistent with the smaller levels of homozygosis due to identity-by-descent tracked by the simulation of IBD loci. In such situation (schemes E1 and E2), the relative benefit of GS in terms of inbreeding incidence increased along the generations, so that inbreeding levels estimated using pedigree or IBD information were at least 25% smaller under GS in the last generation, when compared to similar figures estimated under BLUP selection.

Similarly, in these situations, GS also allowed significant benefits in terms of diminishing the loss of favorable alleles as well as reducing homozygosis for alleles potentially associated to deleterious mutations.

Such results partially confirm the theoretical expectations found in [9] and [40], who suggested that lower levels of inbreeding could be achieved through the application of genomic selection, as a result of the lower emphasis on family under this strategy, due to the possibility of obtaining better estimates of the Mendelian sampling (MS) term.

However, under the scheme E3, selection based on GBLUP did not result in lower levels of inbreeding when compared to BLUP. In this situation, only when predictions considered more emphasis on lower-frequency favorable alleles (wGBLUP) some benefit was observed over BLUP in the last generations simulated. For this selection scheme, the phenotypic information available before selection allowed estimation of MS with greater accuracy, what can contribute to reduce the emphasis on family information under BLUP.

The results verified under different scenarios of genetic architecture (scheme E2) suggest that larger advantage of GS over BLUP in terms of inbreeding incidence can be expected when the trait under selection had less polygenic background. The largest benefit for GS over BLUP in terms of inbreeding levels was observed for the scenarios where a smaller number of QTL or QTL of larger effect were simulated. Nevertheless, while [24] also reported lower incidence of inbreeding for GS when compared to BLUP, these authors did not find any evidence of association between genetic architecture and inbreeding incidence.

When different replacement strategies were compared, larger genetic progress was verified for scenarios without restriction on the replacement rate, strategy that in practice meant that a larger proportion of younger animals were selected. For both schemes in which replacement strategies were compared (E1 and E2), a larger replacement rate resulted in larger benefit for BLUP scenarios, where the accumulated genetic gain was about 40% larger compared to a situation when a fixed replacement rate of 20% was applied. Under genomic selection scenarios, the increase in genetic gain attributed to the allowance of variable replacement rate was about 30%.

Conversely, the inbreeding coefficients were larger for most of the scenarios in which a larger replacement rate was applied, and noticeably larger for selection based on BLUP. In the present study, for the sake of simplicity, a specific age for animals to be considered available for reproduction was not simulated, what could allow to compute generation intervals in a more standard manner. The larger reduction in generation intervals under the scenarios with no constraint on the replacement rate would imply in a larger increase in inbreeding levels, if the rate of inbreeding was considered in an annual basis. A reduction in generation intervals was also verified when GS was compared to BLUP in most situations and regardless of the replacement strategy.

According to [41], selection schemes in which the annual increase in inbreeding is high would require strategies such as optimum contribution selection (OCS), which can be more effective to manage genetic diversity. In the present study, although a considerable advantage was associated to a larger replacement rate in terms of genetic progress, the sharp reduction in the generation intervals verified under this strategy and the consequent increase in inbreeding rates draw attention for the need to design alternative GS schemes aiming to prevent severe increases in inbreeding levels, possibly considering OCS.

The genomic information was assumed to be available for all selection candidates in the present study. Although the strategic use of imputation, with genotyping of all selection candidates from a extremely-low density has been suggested as a cost-effective strategy in swine breeding [31], this can be judged as a scenario still hard to be met in some real beef cattle populations, given the costs involved in genotyping. Previous studies on long-term consequences of GS also had such assumption about all selection candidates being genotyped (e.g. [12], [24]). Such assumption was judged to be more suitable to allow testing the hypotheses investigated in the present study and making possible to compare its results to those of studies carried out under similar conditions.

The results verified for the growth trait (E3) suggested that different genotyping strategies can affect decisively the genetic progress in the long-term, so that genotyping extreme animals seems to be a more advantageous strategy, as also pointed out by some previous studies (e.g. [42], [43]). Since the main objective of the

present study was not to test genotyping strategies, a few scenarios were simulated and thus results with respect to this point are not totally conclusive. A further study should be carried out to investigate long-term consequences of genotyping strategies in more detail.

Performance of genomic selection methods under different scenarios of genetic architecture

For the scheme of selection for meat quality traits (E2), where different scenarios of genetic architecture were simulated, no significant benefit over GBLUP was found for variable selection methods in all time horizons considered. Especially for the scenario where 100 QTL were simulated, it would be expected larger accuracy for the predictions obtained with LASSO, similarly as pointed out in [34] and [35].

The LASSO method, such as implemented in this study, was possibly more sensitive to the small size of the reference population size in the first generations simulated. It is known that LASSO method selects at most as many variables as there are observations in the reference population [32]. Given the level of linkage disequilibrium simulated in this study, it can be expected that, in the first generations (in which the reference population size was small), a number of markers larger than the size of the reference population could be associated to at least one QTL, what would imply in selection of a sub-optimal set of markers and thus smaller accuracy for LASSO prediction. The consideration of a small weight to the ridge penalty in the LASSO predictions in the present study was intended to alleviate this problem, but it did not seem to be effective.

Such hypothesis could be corroborated by the fact that, as the reference population increased, the genetic gain accumulated under LASSO prediction was slightly larger than that verified under GBLUP, when 100 QTL were simulated. Even though the difference was significant, this result would suggest some benefit of LASSO prediction in this scenario, also drawing attention for the limitations of using LASSO for genomic prediction based on small reference populations.

No significant differences between LASSO and GBLUP were found in terms of inbreeding incidence, although for the trait under more polygenic control (trait A) the inbreeding levels were slightly larger for LASSO (+10%). Bastiaansen et al. [24] also

reported that larger pedigree-based inbreeding coefficients were obtained under two variable selection methods (Bayesian regression and partial least squares), when compared to GBLUP (coefficients about 10% larger), although differences between methods in terms of genomic inbreeding were not reported.

Sonesson et al. [12] simulated a trait influenced by 1,000 QTL, so that phenotypic information was available on full-sibs of the selection candidates. In this situation, using a variable selection method (Bayes B) to compute genomic predictions did not result in larger genetic progress when compared to GBLUP, under truncation selection, while it resulted in larger IBD coefficient (+12% in the last generation simulated). According to these authors, this result was associated to the fact that variable selection methods applies larger selective pressure nearby QTL of larger effect, increasing the levels of homozygosity in these regions (phenomenon known as hitchhiking effect, [44]) and thus increasing IBD coefficients. In the present study, the larger estimates of IBD verified for LASSO in the last generations of a scenario with comparable genetic architecture (E2, trait A) would corroborate such hypothesis.

Many statistical methods can be employed for genomic prediction (e.g. Bayesian regression methods, non-parametric methods, penalized methods). A comprehensive review of genomic prediction methods can be found in [45]. For the sake of simplicity, only two penalized methods (GBLUP and LASSO) were investigated in the present study. Future studies could investigate long-term consequences of GS based on other genomic prediction methods, applying a similar framework to that employed in the present study.

Impact of weighting on low-frequency favorable alleles

There was no conclusive evidence in favor of the hypothesis that attributing more weight to favorable alleles of lower frequency (as applied in the definition of wGBLUP and wLASSO) would enhance long-term genetic gain, although in most of the scenarios using this weighting strategy tended to produce lower levels of inbreeding and reduced the loss of favorable alleles. When traits under less polygenic control were simulated, it was verified slightly larger genetic gain for wGBLUP, when compared to GBLUP at generation 15, but such differences were not significant.

Conversely, Jannink [16], verified that greater long-term genetic gain was achieved under adoption of a selection criterion similar to wGBLUP. Some of the explanations for such divergence could be related to the genetic architecture of the simulated trait. Jannink [16] simulated 100 QTL explaining equal proportion of the additive variance, in a way that QTL of lower MAF had larger allele substitution effect. This constraint was not applied in the present study, what could have reduced the relative importance of losing favorable QTL alleles that did not contribute much to the additive variance.

Moreover, [16] simulated bi-allelic QTL with same properties of the markers, what theoretically would suggest higher LD between QTL and markers than in a situation in which they had different properties, as is the case of the present study. The effective population size (Ne) also plays an major role in the loss of alleles due to drift, together with the selection coefficient of a given allele [8], but there is no information in [16] that would allow to evaluate the impact of eventual differences in terms of Ne when compared to the present study.

In addition, due to the quadratic trend observed for Nlost, it could be expected that more pronounced differences between methods (e.g. GBLUP vs wGBLUP) in terms of loss of favorable alleles could observed if more generations were simulated. Another point that is worth to mention is that the function employed to derive wGBLUP, attributing more weight to low-frequency favorable alleles, could have been ineffective to ensure optimal weighting in the simulated population. Further studies are needed to investigate other alternative strategies to compute criteria similar to wGBLUP.

In the present study, effects of inbreeding depression were not simulated, what means that larger levels of inbreeding in a given scenario did not reflect in worsen phenotypic performance. Because inbreeding coefficients under wGBLUP (and wLASSO) tended to be smaller, the benefit of weighting on low-frequency favorable alleles could be larger in the case of traits affected by inbreeding depression. Thus, future studies could investigate the consequences of selection based on criteria similar to wGBLUP also considering some form of discount rate for losses due to inbreeding incidence, similarly as in [46].

While monitoring IBD loci is feasible in simulations, such approach cannot be employed in real applications. However, SNP marker information could be employed to design strategies aiming more effective management of genetic diversity, what theoretically would even allow to deal with local patterns of homozygosis. For instance, Pryce et al. [46] simulated mate selection strategies using information based on runs of homozygosity aiming to control inbreeding incidence.

Inbreeding estimators and genetic diversity

Previous studies with real data have investigated the association between Fped and F_{ROH} in real cattle populations (e.g. [14], [36]). Despite this, the knowledge about the association between such estimators and true levels of homozygosis due to IBD was still limited and could be assessed through simulation. At our knowledge, the present study was the first to investigate the use of inbreeding coefficients based on runs-of-homozygosity (F_{ROH}) to estimate identity-by-descent coefficients in a simulated livestock population undergoing selection. This strategy was also compared to other alternatives based either on pedigree (Fped) or marker information (Fhet).

The results of the present study evidenced that marker-based estimators (Fhet and F_{ROH}) are superior to pedigree-based information (Fped) to estimate individual coefficients of homozygosis due to identity-by-descent (IBD), given the stronger correlations verified between these marker-based estimators and individual estimates of IBD computed after tracking founder alleles for more than 3,000 loci.

Bouquet et al. [47] also verified stronger association between IBD coefficient and marker-based estimates of inbreeding, when compared to Fped, after simulating two lines with Ne equal to 1,000 (selected and non-selected). Although no marker-based estimator is directly comparable to those of the present study, these authors reported a reduction of 7% in the correlation between Fped and IBD in the selected line, as well as larger correlations between IBD and marker-based estimators, under selection. Such results agree with the trend of reduction in the degree of the association between Fped and IBD along the selection process.

One of the main benefits of marker-based inbreeding estimators is the possibility to identify identical-by-descent segments due to common ancestors at much more generations in the past than it is possible using only pedigree information

([48], [36]). Keller et. al [27] also verified by simulation that F_{ROH} would allow more statistical power to study inbreeding depression when compared to other estimators of inbreeding, including Fped and Fhet.

Given the larger empirical support for the partial dominance hypothesis to explain inbreeding depression [49], Keller et. al [27] also correlated different estimators of inbreeding to estimates at HML loci, simulated similarly as in the present study. Although the simulations considered parameters particular to human populations, such authors also reported larger correlations between marker-based estimators of inbreeding and homozygosis at HML loci. Under Ne=100, larger correlations were found between HML and F_{ROH} (r~0.60) than between HML and Fhet (~0.45) and between HML and Fped (~0.25). In the present study, while the correlations between HML and Fped were greater (0.35-0.40), such values were, at least, 20% lower than those verified for the correlation between HML and the marker-based estimators.

The association between the marker-based estimators and Fped was also high in the present study (>0.8), with the estimated correlations being slightly larger for F_{ROH}. By analyzing real data of bovine populations undergoing selection, [36] and [14] also reported moderate to strong correlations between Fped and F_{ROH}, with values of 0.70 and 0.50-0.72, respectively. As in the present study, the authors of [14] also verified stronger association between Fped and F_{ROH} when compared to that between Fped and Fhet.

The estimated correlations between Fped and F_{ROH} were also slightly larger in the present study when compared to those reported by [46], who obtained figures between 0.64-0.71 for this statistic, depending on the minimum length considered to define a ROH, after simulating a dairy cattle population. According to Ferenkacovic et [14], the strong correlations between Fped and F_{ROH} would indicate that F_{ROH} is a good estimator of the inbreeding detectable through pedigree, while could provide additional information about more remote common ancestors.

Usually, inbreeding coefficients have been estimated based only on pedigree information [27], situation in which the precision of the estimates is dependent on the quality and completeness of pedigree records [50]. Under this approach, the

estimated coefficients are computed relatively to a base population, in practice determined by the availability of pedigree information.

In bovine populations, problems related to missing information or wrong assignment of parents are recurrent. Wiggans et al. [51] reported rates between 4% and 14% of inconsistency in the informed sires of the animals genotyped in their study. In the case of beef cattle, using multiple sire breeding is a common management practice [52], what implies that the sire of a large proportion of the animals born is not known. For instance, in the study of Shiotsuki et al. [53], about 37% of the animals with phenotypic records were product of multiple sire breeding.

The results of the present study suggest that marker-based estimators of inbreeding are superior to pedigree-based estimators when the aim is to estimate homozygosis due to identity-by-descent at the level of individuals, being that the estimator based on ROH outperformed the remainder for this task. It must be emphasized that the advantage of using marker information for this purpose could be even larger in the case of real populations, since no pedigree errors or missing information were simulated.

The extent of association between IBD and HML ($r < 0.50$) in the present study would indicate that accurate estimators of homozygosis due to identity-by-descent could not be enough to track all variation in loci potentially associated to deleterious mutations, under the assumption that the genetic basis of such mutations is consistent with that which was simulated, i.e. homozygosis at rare alleles (MAF <5%). While there is no consensus on the role of rare variants on diseases with complex inheritance [54], there is some support for the assumption that deleterious alleles rarely would reach frequencies larger than 5% [27]. Under such assumption, the present results also suggest that SNP marker information can be a more effective tool to be considered in the design of mating strategies aiming to minimize the probability of deleterious mutations and inbreeding depression.

Conclusions

While genomic selection allowed genetic progress up to 40% greater than BLUP under selection for female reproduction and meat quality traits, no benefit was verified under selection for a beef cattle growth trait.

The simulation of repeated cycles of genomic selection and updating of the reference population allowed more detailed investigation on the long-term consequences of such strategy in terms of inbreeding incidence and genetic diversity. There was evidence that larger advantage for genomic selection compared to BLUP, in terms of controlling inbreeding rates, are expected when the selected trait is under less polygenic background.

Using a genomic selection criterion in which markers having favorable alleles at low frequency were more heavily weighted did not provide benefit in terms of genetic progress in both short-term and long-term. The present results are suggestive that this strategy can contribute to reduce inbreeding rates and loss of favorable alleles.

Estimation of inbreeding using marker information based on the concept of runs of homozygosity outperformed the remainder strategies to estimate homozygosis due to identity-by-descent.

References

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

2. Schaeffer LR: **Strategy for applying genome-wide selection in dairy cattle.** *J Anim Breed Genet* 2006, **123**:218-223.

3. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME: **Invited review: Genomic selection in dairy cattle: Progress and challenges.** *J Dairy Sci* 2009, **92**:433-443.

4. Luan T, Woolliams JA, Lien S, Kent M, Svendsen M, Meuwissen THE: **The accuracy of genomic selection in Norwegian Red cattle assessed by cross-validation.** *Genetics* 2009, **183:**1119-1126.

5. Ostersen T, Christensen OF, Henryon M, Nielsen B, Su G, Madsen P: **Deregressed EBV as response variable yield reliable genomic predictions for pigs.** *Genet Sel Evol* 2011, **43**:38.

6. Daetwyler HD, Swan AA, Van der Werf JH, Hayes BJ: **Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation.** *Genet Sel Evol* 2012*,* **44**:33.

7. Saatchi M, Schnabel RD, Rolf MM, Taylor JF, Garrick DJ: **Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle.** *Genet Sel Evol* 2012, **44**:38.

8. Bijma P: **Long-term genomic improvement – new challenges for population genetics.** *J Anim Breed Genet* 2012*,* **129**:1-2.

9. Daetwyler HD, Villanueva B, Bijma P, Wolliams JA: **Inbreeding in genome-wide selection.** *J Anim Breed Genet* 2007, **124**:369–376.

10. Meuwissen THE: **Maximizing the response of selection with a predefined rate of inbreeding.** *J Anim Sci* 1997, **75**:934-940.

11. Carvalheiro R, Queiroz SA, Kinghorn B: **Optimum contribution selection using differential evolution.** *Rev Bras Zootec* 2010, **39**:1429-1436.

12. Sonesson AK, Wooliams JA, Meuwiseen THE: **Genomic selection requires control of inbreeding.** *Genet Sel Evol* 2012, **44**:27.

13. Villanueva B, Pong-Wong R, Fernandez J, Toro MA: **Benefits from marker-assisted selection under an additive polygenic genetic model.** *J Anim Sci* 2005, **83**:1747–1752.

14. Ferencakovic M, Hamzic E, Gredler B, Solberg TR, Klemetsdal G, Curik I, Sölkner J: **Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations.** *J Anim Breed Genet* 2013, **130**: 286-293.

15. Goddard ME**: Genomic selection: Prediction of accuracy and maximisation of long term response.** *Genetica* 2009**, 136**: 245-257.

16. Jannink J-L: **Dynamics of long-term genomic selection.** *Genet Sel Evol* 2010*,* **42**:35.

17. Wiggans GR, Sonstegard TS, VanRaden PM, Matukumalli LK, Schnabel RD, Taylor JF, Chesnais JP, Schenkel FS, Van Tassell CP: **Genomic Evaluations in the United States and Canada: A Collaboration.** In *International Committee on Animal Recording (ICAR*), June 22–26, 2009. Munich: ICAR Tech Ser; 2009:347-353.

18. VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, Sonstegard TS, Connor EE, Winters M, Van Kaam J, Van Doormaal BJ, Faust MA, Doak GA: **Genomic imputation and evaluation using high density Holstein genotypes.** *J Dairy Sci* 2013, **96**:668-678.

19. Muir WM: **Comparison of genomic and traditional BLUP estimated breeding value accuracy and selection response under alternative trait and genomic parameters.** *J Anim Breed Genet* 2007, 124:342–355.

20. Goddard ME, Hayes BJ, Meuwissen THE: **Genomic selection in livestock populations.** *Gen Res Camb* 2010, **92**: 413-421.

21. Brito FV, Braccini Neto J, Sargolzaei M, Cobuci JA, Schenkel FS: **Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle.** *BMC Genet* 2011, **12**:80.

22. Carvalheiro R, McEwan JC, Neves HHR, Utsunomiya YT, O'Brien AMP, Boison, SA, Sölkner J, Schenkel FS, Van Tassell CP, Sonstegard TS, Garcia JF: **Accuracy of genomic predictions in Nellore beef cattle.** In *Proceedings of the 20th AAABG Conference*. 2013, Napier, New Zealand.

23. Sargolzaei M, Schenkel FS: **QMSim: a large-scale genome simulator for livestock.** *Bioinformatics* 2009*,* **25**:680-681.

24. Bastiaansen JWM, Coster A, Calus MPL, Van Arendonk JAM, Bovenhuis H: **Long-term response to genomic selection: effects of estimation method and reference population structure for different genetic architectures.** *Genet Sel Evol* 2012, **44**:3.

25. Snelling WM, Chiu R, Schein JE, Hobbs M, Abbey CA, Adelson DL, Aerts J, Bennett GL, Bosdet IE, Boussaha M, Brauning R, Caetano AR, Costa MM, Crawford AM, Dalrymple BP, Eggen A, Van der Wind AE, Floriot S, Gautier M, Gill CA, Green RD, Holt R, Jann O, Jones SJM, Kappes SM, Keele JW, Jong PJ, Larkin M, Lewin HA, McEwan JC, McKay S, Marra MA, Mathewson CA, Matukumalli LK, Moore SS, Murdoch B, Nicholas FW, Osoegawa K, Roy A, Salih H, Schibler L, Schnabel RD, Silveri L, Skow LC, Smith TPL, Sonstegard TS, Taylor JF, Tellam R, Van Tassell CP, Williams JL, Womack JE, Wye NH, Yang G, Zhao S. for the International Bovine BAC Mapping Consortium: **A physical map of the bovine genome.** *Genome Biol* 2007, **8**:R165.

26. Sargolzaei M, Schenkel FS, Jansen GB, Schaeffer LR: **Extent of linkage disequilibrium in Holstein Cattle in North America.** *J Dairy Sci* 2008, **91**:2106-2117.

27. Keller MC, Visscher PM, Goddard ME: **Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data.** *Genetics* 2011, **89**:237–249.

28. R Development Core Team: **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria, 2013. ISBN 3-900051-07-0 [http://www.R-project.org]

29. Misztal I: **BLUPF90 - a flexible mixed model program in Fortran 90.** [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90.pdf]

30. Neves HHR, Carvalheiro R, Queiroz SA: **A comparison of statistical methods for genomic selection in a mice population.** *BMC Genet* 2012, **13**:100.

31. Huang Y, Hickey JM, Cleveland MA: **Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost.** *Gen Sel Evol* 2012, **44**:25.

32. Ogutu JO, Schulz-Streeck T, Piepho HP: **Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions.** *BMC Proceed* 2012, **6**:S.10.

33. Tibshirani R: **Regression shrinkage and selection via the lasso.** *J R Stat Soc B* 1996, **58**:267-288.

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

35. Clark SA, Hickey JM, Van der Werf JHJ: **Different models of genetic variation and their effect on genomic evaluation.** *Genet Sel Evol* 2011, **43**:18.

36. Purfield D, Berry D, Mcparland S, Bradley D: **Runs of homozygosity and population history in cattle.** *BMC Genet* 2012, **13**:70.

37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC: **PLINK: a toolset for whole-genome association and population-based linkage analysis.** *Am J Hum Genet* 2007, **81**: 559-575.

38. Van der Werf JHJ: **Potential benefit of genomic selection in sheep.** *Proc. Assoc. Adv. Anim. Breed.Genet.* 2009, **18**:38–41.

39. Dekkers JCM: **Application of Genomics Tools to Animal Breeding.** *Curr Genom* 2012, 13:207-212.

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

41. Pryce JE, Daetwyler HD: **Designing dairy cattle breeding schemes under genomic selection: a review of international research.** *Anim Prod Sci* 2012, **52**:107-114.

42. Boligon AA, Long N, Albuquerque LG, Weigel KA, Gianola D, Rosa GJM: **Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection.** *J Anim Sci* 2012, **90**: 4716-4722.

43. Jiménez-Montero JA, González-Recio O, Alenda R: **Genotyping strategies for genomic selection in small dairy cattle populations.** *Animal* 2012, **6**:1216–1224.

44. Smith MJ, Haigh J: **The hitch-hiking effect of a favorable gene.** *Genet Res* 1974, **23**:23-35.

45. de los Campos G, Hickey J, Pong-Wong R, Daetwyler HD, Calus MPL: **Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding.** *Genetics* 2013, **193**:327-345.

46. Pryce JE, Hayes BJ, Goddard ME: **Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information.** *J Dairy Sci* 2012, **95**: 377-388.

47. Bouquet A, Sillanpää M, Juga J: **Estimating inbreeding using dense marker panels and pedigree information.** In *Maataloustieteen Päivät*, Helsinki, 2012.

48. Mcquillan R, Leutenegger A.-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, Macleod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H, Wilson JF: **Runs of homozygosity in European populations.** *Am J Hum Genet* 2008**, 83**:359–372.

49. Charlesworth D, Willis JH: **The genetics of inbreeding depression.** *Nat Rev Genet* 2009, **10**:783-796.

50. Li MH, Strandén I, Tiirikka T, Sevón-Aimonen ML, Kantanen J: **A comparison of approaches to estimate the inbreeding coefficient and pairwise relatedness**

using genomic and pedigree data in a sheep population. *PLOS ONE* 2011: **6**:e26256.

51. Wiggans GR, Cooper TA, VanRaden PM, Olson KM, Tooker ME: **Use of the Illumina Bovine3K BeadChip in dairy genomic evaluation.** *J Dairy Sci* 2012, **95**:1552–1558.

52. VanEenennaam AL: **How might DNA - based information generate value in the beef cattle sector?** In *ICAR Annual Meeting*, Cork, Ireland, 2012.

53. Shiotsuki L, Cardoso FF, Silva JAIIV, Rosa GJM, Albuquerque LG: **Evaluation of an average numerator relationship matrix model and a Bayesian hierarchical model for growth traits in Nellore cattle with uncertain paternity.** *Liv Sci* 2012**, 144**: 89–95.

54. Gibson G: **Rare and common variants: twenty arguments**. *Nat Rev Genet* 2012, **13**: 135-145.

CHAPTER 5 - Long-term consequences of selective genotyping strategies in a beef cattle population undergoing selection

Abstract

Background

In many applications of genomic selection not all animals can be genotyped, what opens the possibility to investigate different strategies to decide the animals which will be genotyped (selective genotyping), aiming to enhance cost-effectiveness of genomic selection. The aim of this study was to evaluate the impact of different genotyping strategies to update the reference population in an application of genomic selection (GS) in beef cattle.

Methods

Forward-in-time simulation was employed to mimic a population with pattern of linkage disequilibrium (LD) close to that verified in real beef cattle populations. Different scenarios of genomic selection (GS) and traditional BLUP selection were simulated for 15 generations, mimicking selection for a trait in which own performance information was available before the selection of the parents of the next generation took place. The GS scenarios were defined by a combination of genotyping strategy and statistical method (single-step or multi-step genomic BLUP).

Results

The use of the single-step approach provided more accurate and less deflated predictions than a multiple-step approach using EBVs to estimate marker effects. The single-step approach also presented advantage over multi-step prediction in terms of inbreeding incidence and accumulated genetic gain. In the long term, there was no clear advantage for a particular genotyping strategy under investigation, although, under the single-step approach, scenarios with genotyping of superior males presented larger accumulated genetic gain until the last generation simulated.

Conclusions

Conversely to suggested by previous studies, there was no evidence that would disqualify updating the reference population trough only genotyping superior animals for EBV, under the single-step approach. The single-step approach is more appealing for genomic prediction under selective genotyping. Further studies should

compare the single-step approach to multi-step prediction, considering other alternatives of blending and response variable not investigated in the present study.

Background

Selective genotyping strategies were the subject of recent simulation studies on genomic selection (GS) and could be employed to improve cost-effectiveness of GS applications (e.g. [1], [2], [3]). For beef cattle breeding programs, cost-effective genotyping strategies would be even more important, because in this situation the added value of GS is generally lower than that envisaged with the application of this technology in dairy cattle schemes [4].

The most common genotyping strategy in dairy cattle has been genotyping individuals with highly-accurate pseudo-phenotypes (e.g. EBVs or deregressed proofs) to compose the reference population, what lead to proven bulls being genotyped in the first applications of GS ([5], [6]). The reasoning of such strategy is that the reliability of pseudo-phenotypes employed for estimation of marker effects plays an important role in accuracy of genomic predictions [7]. However, the number of highly informative individuals available is lower for beef cattle breeding programs, what would imply in the need for a larger number of individuals being genotyped to achieve accuracies similar to those verified for genomic predictions obtained from highly accurate proofs.

Although is well-established the idea that GS can increase the genetic progress for traits in which small improvement has been achieved using traditional selection ([8], [9], [10]), the large-scale application of GS in beef cattle breeding programs would also have to cope with selection for regular traits, which are part of the current selection criteria (e.g. growth-related traits and carcass composition), especially because these traits are expected to have larger influence on the economic return in the short term.

The assumption that all selection candidates are genotyped was made in some studies that addressed long-term consequences of GS (e.g. [11], [12]). For beef cattle, given the small benefit envisaged for applying GS to improve regular traits and the extra cost associated to genotyping, one cannot expect that all

selection candidates would be genotyped , what allows employing selective genotyping to improve cost-effectiveness of GS.

As a general rule, the results of simulations on selective genotyping have considered only the first generation subsequent to marker effect estimation when comparing the results of each strategy applied ([1], [2], [3]). While such approach gives insight about short-term consequences of genotyping strategies, the knowledge about their long-term consequences is important to enhance the sustainability of GS schemes. Because genomic prediction equations are routinely updated in real applications of GS (e.g. [13], [14]), it is also desirable that alternative genotyping strategies are compared under such a scenario of prediction.

Besides the potential benefits of selective genotyping, this strategy can lead to problems like the need for methods allowing the comparison among selection candidates with different amounts of information (e.g. genotyped and non-genotyped selection candidates), as well as the possibility of introducing bias in genomic predictions ([15], [16]).

For Brazilian beef cattle breeding programs, genotyping superior animals can be a very likely scenario in the near future, especially in the case of the top 20% males born in commercial breeding programs, which are candidates to an official certificate (CEIP) that attests their genetic superiority and enables special conditions for them to be commercialized as young replacement bulls. Among these top animals, some are chosen to be progeny tested and the possibility of obtaining EBVs of larger accuracy through the use of genomic prediction can be an important motivation for genotyping, because the added value in this situation could pay off the genotyping costs.

On the other hand, some drawbacks of genotyping superior animals have been identified. Boligon et al. [2] reported smaller accuracy of genomic predictions obtained under genotyping of superior animals compared to a situation in which extreme animals were genotyped. Jiménez-Montero et al. [3] also alerted to the fact that genotyping only superior animals could produce genomic predictions with larger bias and mean squared error (MSE), when compared to genotyping extreme animals.

Regarding to the problem of comparing genotyped and non-genotyped animals, the single-step approach (SS) [17] has been pointed out as an optimal strategy to combine all information available [18] through the consideration of relationships obtained after combining all pedigree and genomic information available.

Despite this, most practical applications of genomic selection currently involves multiple steps (multi-step, MS): traditional genetic evaluation, computing pseudo-phenotypes, estimation of marker effects , genomic prediction and blending of genomic to traditional proofs [19]. According to this author, the complexity and the several approximations involved in the multi-step approach would make it prone to errors and contribute to reduce the prediction accuracy, while the single-step approach would allow genomic prediction in a simple, fast and accurate way. Vitezica et al. [16] also provided a formal proof that, under a well-formed relationship matrix, the single-step approach is more accurate and less biased than the multi-step approach.

This study was carried out to compare different selective genotyping strategies in a simulated beef cattle population undergoing selection, in both short and long-term, and also to evaluate the quality of genomic predictions obtained using either the single-step or the multi-step approach under the different genotyping strategies.

Methods

Population structure and simulated genome

The population structure and the properties of the simulated genome were identical to those of the simulation described in the previous chapter, so that simulation parameters were defined similarly as in [10], aiming to mimic the extent of linkage disequilibrium estimated using data of real beef cattle populations. The target marker density was that verified after commercial SNP panels designed for cattle, containing about 50,000 markers, were submitted to quality control procedures (e.g. [20], [6]).

Simulated scenarios

The simulation routine mimicked selection for a trait with polygenic background, affected by 1000 loci (QTL) and whose heritability was equal to 0.25. The QTL effects were drawn from a gamma distribution with shape parameter equal

to 0.4 and the number of QTL alleles in each locus varied from to two to four. The true breeding value (TBV) of each animal was obtained as the sum of the QTL allele effects over all loci and the phenotypes were simulated by adding a random residual to the TBV, so that the target heritability was met.

It was assumed that the phenotypes were available on animals of both sexes. For selection candidates of a given generation, own performance information was available before the selection of the parents of the next generation took place, similarly as occurs with some beef cattle production traits (for instance weight traits and carcass composition evaluated trough visual scores).

Ten populations (replicates) were generated using the same simulation parameters, what allowed that identical base populations (G0) were employed to simulate all scenarios, similarly as applied by [11]. For each replicate, 15 generations were simulated, so that different selection strategies were applied in each scenario. The mating among G0 animals produced the 1,000 animals from generation G1. In order to produce the generation **i** ($2 \leq i \leq 15$), the selection criterion employed in each scenario was used to rank the animals available for reproduction, i.e. animals from generation **(i-1)** or their parents, so that the top 1,000 females and the top 40 males were selected. This step was succeeded by the random mating of the selected animals and the production of their offspring, comprising 1,000 individuals (i.e. each selected dam and sire produced one and 25 offspring per generation, respectively).

In the scenarios under genomic selection, the initial reference population was composed by 1,000 animals of the generations G0 and G1 (500 animals randomly selected from each generation). Starting from generation 2 (G2), different genotyping strategies were applied in each genomic selection scenario to update the reference population, so that marker effects were re-estimated at each generation.

The simulated scenarios are described in Table 1 and were defined aiming to investigate the influence of the following factors on genomic predictions: genotyping strategy (only top animals, extreme animals, random sampling and sampling within-sire family), criterion to rank candidates for genotyping (EBV or adjusted phenotype) and genetic evaluation methodology (single-step or multi-step). In the scenarios in which EBVs were used to rank the animals for genotyping, before the update of the reference population and the production of the next generation, a traditional genetic

evaluation was carried out using BLUP, which already included own performance records of all candidates for genotyping.

In this study, alternative genotyping strategies were investigated. Genotyping upper extreme and lower extreme animals as well as genotyping top animals or randomly chosen animals were already investigated in previous studies (e.g. [2], [3]). Given the possibility of greater economic appeal for genotyping males candidates to CEIP (top20M) and the better results obtained in previous studies for genotyping extreme rather than only superior animals, the first alternative genotyping strategy was designed aiming to increase predictive ability of genotyping of the top% 20 males by combining this strategy to the genotyping of extreme animals (bottom 20% females, top20M_bot20F) (Table 1). This choice for bottom females was due to the fact that female genotypes can be useful for other applications like imputation and detection of lethal recessive haplotypes, for instance.

Because genomic predictions for selection candidates more closely related to the reference population are expected to be more accurate [21], another alternative strategy was designed aiming to update the reference population with a set of animals that could provide more representative sampling of the half-sibs families at each generation (WS20MF). Another alternative strategy combined genotyping of the top 20% males and the random sampling of 20% of the females within each sire-family (top20M_WS20F).

Table 1. Description of the simulated scenarios

Scenario	Genotyping strategy[1]	Genetic evaluation[2]
BLUP	-	BLUP
top20_YD	top 20% of each sex, yield deviation (YD)	ssGBLUP
RND20	20% of each sex, random	
top20	top 20% of each sex, EBV	
top20M_bot20F	top 20% males / bottom 20% females, EBV	
top20M_WS20F	top 20% males EBV / 20% females randomly sampled within sire-families	
WS20MF	20% of each sex randomly sampled within sire-families	
ALL_ms1	all genotyped	msGBLUP1
top20_ms1	top 20% of each sex, EBV	msGBLUP1
top20_ms2	top 20% of each sex, EBV	msGBLUP2
top20_ms3	top 20% of each sex, EBV	msGBLUP3
top20_ms4	top 20% of each sex, EBV	msGBLUP4
top20M_bot20F_ms2	top 20% males / bottom 20% females, EBV	msGBLUP2

[1]Starting from generation 2, a fixed proportion of animals was genotyped to update the reference population used to estimate marker effects. For each sex, the percentages refer to the proportion of genotyped animals among all born in a given generation. YD = adjusted phenotype (yield deviation), simulated as the sum of true breeding value and random residual (heritability = 0.25); EBV= breeding value estimated using BLUP (i.e. considering pedigree and phenotypic information).

[2]ssGBLUP: best linear unbiased predictor using a relationship matrix combining pedigree and genomic information in a single-step [17]. msGBLUP1: best linear unbiased predictor using a genomic relationship matrix (GBLUP) implemented in multiple steps, using Sullivan's blending method [22]. msGBLUP2: similar to msGBLUP1, but with an additional step to adjust EBVs for the same scale/location of EBV. msGBLUP3: multi-step GBLUP, the sum of marker effects (DGV) was adjusted for the location/scale of the EBVs and employed as selection criterion (i.e. no blending). msGBLUP4: similar to msGBLUP3, but without adjustment for EBV scale/location. (More details are presented later in the Methods section)

Regarding to the genomic evaluations carried out in each scenario, starting from generation 1, the selection of the parents of the next generation was based on predictions of the genetic merit obtained using either of the following methods:

-msGBLUP: best linear unbiased predictor using a genomic relationship matrix, implemented in multiple steps. At each generation, the first step comprised a regular genetic evaluation using BLUP, considering all information available (phenotypes and pedigree), aiming to generate pseudo-phenotypes (EBVs) for the genotyped animals. The second step comprised fitting a GBLUP model using EBVs of the genotyped animals as response variables. The predictions obtained from this model will be regarded to as direct genomic values (DGV). The GBLUP model can be described by:

$$\mathbf{y}=\mathbf{1}_n\boldsymbol{\mu} + \mathbf{Z}\mathbf{g} +\mathbf{e}, \text{ [1]}$$

in which **y** is a vector of pseudo-phenotypes, **μ** is a location parameter common to all observations, $\mathbf{1}_n$ is a vector of 1's, **Z** is an incidence matrix relating DGVs to **y**, **g** is a vector of DGVs and **e** is a vector of random residuals. It was assumed that $\mathbf{g} \sim N(0,\mathbf{G}^*\sigma^2 g)$ and $\mathbf{e} \sim N(0,\mathbf{D}\sigma^2 e)$, where **G*** is combined relationship matrix and **D** is a diagonal matrix of weights defined to take differences in EBVs' accuracies into account. The **G*** matrix was obtained as **G*=**(1-w)***G+**w***A**, where **G** is the genomic relationship matrix and **A** is the numerator relationship matrix (pedigree based*)*, both of order equal to the number of genotyped animals and w=0.20. In a previous study, after testing different values for w (ranging from 0 to 0.40), Gao et al. [23] found that w=0.20 provided the best compromise in terms of reliability and scale of genomic predictions, what motivated the use of this value for w in the present study. **G** was obtained following VanRaden [24], so that **G**=**MM'**/Σ[2pi(1-pi)] where **M** is an incidence matrix of additive marker effects, where the i-th column contains elements equal to (0-2pi), (1-2pi) and (2-2pi) depending on whether the genotype contains, respectively, 0, 1 or 2 copies of the reference allele for the i-th marker and whose allele frequency is pi. This method was implemented using the gebv software [25].

The next step was carried out to obtain genomic predictions for the genotyped selection candidates, so that this information was more comparable to the information available on the non-genotyped candidates (i.e. EBV). For this, the Sullivan's blending method [22] was used to combine the information of EBV and DGV of the genotyped animals, resulting in an estimate that will be regarded to as genomic estimated breeding values (GEBV) hereafter.

The last step consisted in the calculation of a selection criterion (Isel) that was used to rank all selection candidates, regardless of whether they were genotyped or not. For all non-genotyped animals, the value of Isel was equal to their respective EBVs. Given the possibility of biased genomic predictions obtained using selective genotyping reported in previous studies, we tested different strategies to compute Isel for the genotyped animals.

The first alternative to compute Isel consisted in assuming this selection criterion to be equal to the value of GEBV (Isel = GEBV) (msGBLUP1). The second consisted in computing the intercept (b0) and the slope (b1) of the regression of EBV on GEBV, considering all genotyped animals. In this situation (msGBLUP2), the value of Isel for the genotyped animals was computed as Isel= b0+b1*GEBV. The other alternative procedures were similar to msGBLUP1 (msGBLUP4) and msGBLUP2 (msGBLUP3), except by replacing GEBV by DGV.

-**ssGBLUP**: best linear unbiased predictor using a relationship matrix (**H**) combining pedigree and genomic information, in a single-step (Legarra et al., 2009). According to Aguilar et al (2010), the inverse of **H** ($\mathbf{H^{-1}}$) can be obtained as:

$$\boldsymbol{H^{-1}} = \boldsymbol{A^{-1}} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{G^{-1}} - \boldsymbol{A_{22}^{-1}} \end{bmatrix}$$

where $\mathbf{A^{-1}}$ is the inverse of the numerator relationship matrix, $\mathbf{G^{-1}}$ é is the inverse of the genomic relationship matrix (**G**), obtained as in VanRaden et al. [24] and, $\boldsymbol{A_{22}^{-1}}$ is the inverse of the numerator relationship matrix among the genotyped animals.

This formulation can be understood as a projection of the genetic merit from the genotyped to the non-genotyped animals using pedigree relationships [17]. Thus, replacing $\mathbf{A^{-1}}$ by $\mathbf{H^{-1}}$ in regular genetic evaluation software would allow prediction of genetic merit of all animals simultaneously, regardless of whether they were genotyped or not.

An important detail involved in the computation of $\mathbf{H^{-1}}$ is to ensure that both genomic (**G**) and pedigree-based ($\mathbf{A_{22}}$) relationships are compatible. For this, **G** was re-scaled so that the average of off-diagonal elements and the average diagonal of **G** were equal to their respective counterparts in $\mathbf{A_{22}}$ [16]. In the present study, this method was implemented using BLUPF90 [26]. At each generation, ssGBLUP was fitted to allow that all information available (genotypes, phenotypes and pedigree) was employed to predict the genetic merit of the selection candidates.

In order to establish a baseline for the results obtained using either ssGBLUP or msGBLUP, a scenario under traditional genetic evaluation was simulated by using EBVs obtained by BLUP to rank all selection candidates at each generation. As a way to evaluate the relative benefit of selective genotyping contrasted to genotyping all selection candidates, another scenario consisted in genotyping all animals born in generation 2 or later. Due to limitations to invert **G** in the last generations of the simulation under ssGBLUP, this methods was only implemented using msGBLUP.

Analyses of the results

For each generation obtained after mating selected candidates (**i,** 2≤ i ≤15) the following statistics were computed for each scenario:

-deltaG: accumulated genetic gain, computed as the difference between the average TBV at generation **i** and the average TBV in the respective base population (G0), standardized by the SD of TBVs in the base population.

-acc: accuracy of prediction, computed as the Pearson's correlation between predicted genetic merit and TBV, considering all selection candidates, i.e. animals from generation (**i**-1) or their parents.

-b0 and b1: respectively, the intercept and the slope of the regression of TBV on predicted genetic merit, considering all selection candidates. Such statistics aim to measure the extent of bias (b0) and the scale (b1) of the predictions in each scenario. Values of b0 and b1 closer to 0 and 1, respectively, indicate less biased predictions and with scale close to that of the true breeding value (TBV).

-aveL: average generation coefficient of the parents of the animals from generation **i**, computed as a proxy for generation interval in each scenario.

-IBD: inbreeding coefficient estimated by genomic identity-by-descent, similarly as in [12]. Founder alleles were monitored for 3,064 SNP markers evenly distributed along the genome and removed from the panel available for genomic prediction. At each locus, homozygosis was computed as fj = $\sum f_{kj}^2$, where $\mathbf{f_{kj}}$ is the frequency of the founder allele **k** at locus **j** in the generation **i**. The IBD coefficient was calculated as the average of fj over all monitored loci.

At each generation, all statistics were computed separately for each replicate (population). Aiming to objectively enhance comparisons among scenarios, results of each statistic were analyzed by fitting the following linear model:

$$s_{ij} = c_i + p_j + e_{ij},$$

where $\mathbf{s_{ij}}$ is the j-th observation for statistic **s** in the i-th scenario, $\mathbf{c_i}$ is the effect of the i-th scenario on s, $\mathbf{p_j}$ is the effect of the j-th base population and **eij** ~ $N(0,\sigma^2 e)$. The assumptions of normally distributed and homoscedastic residuals were checked using Shapiro-Wilk and Breusch-Pagan tests, respectively. For each statistic and generation, least squares means of each scenario were contrasted using t tests, adjusted for false discovery rate (α=5%).

Results

As a general rule, in the first three generations of selection, the highest accuracy of prediction was obtained using the single-step method (ssGBLUP), which was, on average, 6%, 3.5% and 1.5% more accurate than BLUP in the generations 2, 3 and 4, respectively (Table 2), while genomic selection under multi-step GBLUP provided worse accuracy than BLUP in most of the scenarios in this period. Such differences in accuracy of prediction also resulted in similar differences in terms of accumulated genetic gain.

In this period, genotyping of the top 20% animals for yield deviation and prediction under ssGBLUP produced the best results in terms of genetic progress, with gains about 7%, 4% and 3.5% larger than those verified under BLUP selection (Table 2). Although the genetic progress was slightly larger in this scenario when compared to other ssGBLUP scenarios, neither of these differences was significant.

The multi-step genomic prediction consistently resulted in genetic gain quite similar and even slightly worse than BLUP, between generations 2 and 4. The worse result in terms of genetic progress was verified for the scenario where the estimated merit of genotyped candidates was based on the DGV generated using multi-step GBLUP, without adjustment for the scale of EBV (top20_ms4, Table 2).

At generation 2, the worse result for scenario top20_ms4 cannot be attributed to the genotyping strategy, since the scenarios started to differ in terms of selective genotyping only after the animals born in this generation were genotyped, what means that this result can be attributed to the genetic evaluation method, i.e. genomic prediction under multi-step GBLUP without blending nor adjustment for the scale of EBV (msGBLUP4).

Table 2. Statistics[1] used to assess the quality of genetic predictions, genetic progress and inbreeding under different scenarios[2] in generations (gen) 2 to 4

		acc		b0		b1		deltaG		IBD	
gen	Scenario	ȳ	l	ȳ	s	ȳ	s	ȳ	l	ȳ	l
	BLUP	0.63	c	-0.01	0.02	1.12	0.02	0.83	c	0.9	cd
	RND20	0.67	a	-0.01	0.02	1.12	0.02	0.87	ab	0.9	d
	top20	0.67	a	-0.01	0.02	1.12	0.02	0.87	ab	0.9	d
	top20_YD	0.67	a	-0.01	0.02	1.12	0.02	0.88	a	0.9	d
	top20M_bot20F	0.67	a	-0.01	0.02	1.12	0.02	0.87	ab	0.9	d
	top20M_WS20F	0.67	a	-0.01	0.02	1.12	0.02	0.87	ab	0.9	d
2	WS_20MF	0.67	a	-0.01	0.02	1.12	0.02	0.87	ab	0.9	d
	ALL_ms1	0.63	c	-0.01	0.02	1.27	0.02	0.83	c	0.9	bc
	top20_ms1	0.63	c	-0.01	0.02	1.27	0.02	0.83	c	0.9	bc
	top20_ms2	0.63	b	-0.01	0.02	1.14	0.02	0.82	c	0.9	b
	top20_ms3	0.62	d	-0.01	0.02	1.17	0.02	0.85	b	1.0	a
	top20_ms4	0.60	e	-0.01	0.02	1.32	0.02	0.80	d	0.9	b,d
	top20M_bot20F_ms2	0.63	b	-0.01	0.02	1.14	0.02	0.82	c	0.9	b
	BLUP	0.59	b	0.01	0.02	1.01	0.02	1.38	b,e	1.5	d
	RND20	0.61	a	0.04	0.02	1.03	0.02	1.40	a,d	1.4	d
	top20	0.61	a	0.06	0.02	1.04	0.02	1.43	ab	1.4	d
	top20_YD	0.62	a	0.06	0.02	1.04	0.01	1.44	a	1.4	d
	top20M_bot20F	0.61	a	0.05	0.02	1.02	0.02	1.43	ab	1.4	d
	top20M_WS20F	0.61	a	0.05	0.02	1.03	0.02	1.43	ab	1.4	d
3	WS_20MF	0.61	a	0.04	0.02	1.03	0.02	1.41	a,c	1.4	d
	ALL_ms1	0.59	b	-0.02	0.02	1.14	0.02	1.34	ef	1.7	bc
	top20_ms1	0.58	b	0.00	0.02	1.07	0.02	1.33	f	1.7	bc
	top20_ms2	0.59	b	0.02	0.02	0.99	0.02	1.37	c,f	1.6	c
	top20_ms3	0.59	b	0.03	0.02	0.99	0.02	1.35	d,f	1.9	a
	top20_ms4	0.58	b	-0.01	0.02	1.15	0.02	1.33	ef	1.8	ab
	top20M_bot20F_ms2	0.59	b	0.01	0.02	0.99	0.02	1.38	b,f	1.6	c
	BLUP	0.61	a,d	0.02	0.02	0.97	0.01	1.88	a,d	2.2	d,f
	RND20	0.62	a	0.09	0.02	1.01	0.01	1.88	a,d	2.1	f
	top20	0.62	a	0.11	0.02	1.01	0.01	1.91	a,c	2.1	f
	top20_YD	0.62	a	0.11	0.02	1.00	0.01	1.94	a	2.1	f
	top20M_bot20F	0.62	ab	0.11	0.02	0.97	0.01	1.93	ab	2.1	ef
	top20M_WS20F	0.62	ab	0.11	0.02	0.99	0.01	1.93	ab	2.2	d,f
4	WS_20MF	0.61	a,c	0.11	0.02	0.98	0.01	1.90	a,c	2.2	d,f
	ALL_ms1	0.60	b,e	-0.01	0.02	1.08	0.02	1.81	de	2.6	bc
	top20_ms1	0.59	de	0.02	0.02	1.01	0.02	1.85	c,e	2.6	bc
	top20_ms2	0.59	e	0.06	0.02	0.93	0.02	1.85	c,e	2.4	cd
	top20_ms3	0.59	e	0.07	0.02	0.93	0.01	1.86	b,d	3.0	a
	top20_ms4	0.58	e	0.03	0.02	1.02	0.02	1.78	e	2.8	ab
	top20M_bot20F_ms2	0.60	c,e	0.04	0.03	0.95	0.02	1.87	a,d	2.4	c,e

[1]acc, b0 and b1= accuracy, bias and scale of predictions, respectively; deltaG=accumulated genetic gain (in units of SD of true breeding values of the base population); ȳ and s = average and standard error (10 replicates). IBD = average coefficient of genomic inbreeding (measured at IBD loci), in %. For a same statistic and generation, averages followed by the same letter (l) do not differ (adjusted p-value > 0.05). More than two consecutive letters in the same cell are presented as an interval (e.g. a,d = abcd). [2]Each scenario was defined by a combination of genotyping strategy and genetic evaluation method (as described in Material and Methods, Table 1).

One of the main reasons for the worse performance of msGBLUP seems to be associated to the deflation of genomic predictions under msGBLUP4, which was significantly more pronounced in this method, especially in the first two generations considered.

Between generations 2 and 4, the single-step method resulted in predictions with estimates of slope (b1) very close to 1, as well as to those verified for BLUP (Table 2), what was also verified in scenarios under multi-step prediction after adjustment for the scale of EBV (i.e. msGBLUP2 and msGBLUP3). On the other hand, starting from generation 3, ssGBLUP resulted in predictions with intercept (b0) slightly greater than 0 and superior to the estimates verified for BLUP, without important differences between genotyping strategies with respect to this statistic.

In these first three generations, although the IBD coefficients were relatively low for all scenarios, selection using multi-step GBLUP resulted in genomic inbreeding coefficients slightly larger than BLUP, while single-step prediction resulted in IBD coefficients up to 7.5% lower than BLUP.

At generation 5, the accuracy of prediction of selection candidates was only slightly higher for ssGBLUP compared to BLUP (at most 1.5%, without significant differences) (Table 3). Possibly as consequence of the larger accuracies in the first generations, the accumulated genetic gain until generation 5 was, on average, 3.6% larger for ssGBLUP compared to BLUP, being that the genotyping strategies did not differ significantly in this case.

The multi-step GBLUP resulted in smaller prediction accuracy than BLUP in generation 5, regardless of the genotyping strategy or use of scale adjustment and blending to EBV, although most of these differences were not significant (Table 3). The worse result was verified for the scenario top20_ms4, for which the accuracy and the accumulated genetic gain were significantly smaller (about 5%) than under selection using BLUP.

Table 3. Statistics[1] used to assess the quality of genetic prediction, genetic progress and inbreeding under different scenarios[2] in generations(gen) 5, 10 and 15

		acc		b0		b1		deltaG		IBD	
gen	Scenario	ȳ	l	ȳ	s	ȳ	s	ȳ	l	ȳ	l
5	BLUP	0.62	a,c	0.00	0.02	0.99	0.02	2.34	c,f	3.2	c,e
	RND20	0.63	ab	0.13	0.02	0.99	0.01	2.42	a,d	2.9	ef
	top20	0.63	a	0.12	0.02	1.01	0.01	2.42	a,d	2.8	f
	top20_YD	0.63	a	0.14	0.02	1.00	0.01	2.44	a,c	2.9	ef
	top20M_bot20F	0.62	a,c	0.16	0.02	0.97	0.01	2.46	a	2.9	d,f
	top20M_WS20F	0.63	a	0.14	0.02	1.00	0.01	2.45	ab	2.8	f
	WS_20MF	0.62	ab	0.15	0.02	0.97	0.01	2.39	a,e	3.0	d,f
	ALL_ms1	0.60	b,d	-0.05	0.03	1.09	0.01	2.27	fg	3.6	bc
	top20_ms1	0.61	a,c	0.00	0.03	1.02	0.01	2.33	d,g	3.6	bc
	top20_ms2	0.60	cd	0.04	0.03	0.95	0.02	2.33	d,f	3.3	cd
	top20_ms3	0.61	a,c	0.05	0.02	0.95	0.01	2.30	e,g	4.1	a
	top20_ms4	0.59	d	0.02	0.03	1.02	0.02	2.23	g	4.0	ab
	top20M_bot20F_ms2	0.62	a,c	0.02	0.02	0.97	0.01	2.35	b,f	3.5	c
10	BLUP	0.58	e	0.13	0.03	0.94	0.01	4.60	de	7.1	d
	RND20	0.62	b,d	0.30	0.04	0.95	0.02	4.75	a,d	6.5	de
	top20	0.63	ab	0.21	0.04	0.96	0.02	4.83	a,c	6.8	de
	top20_YD	0.62	b,d	0.20	0.03	0.97	0.01	4.72	a,e	6.1	e
	top20M_bot20F	0.63	ab	0.34	0.04	0.95	0.02	4.91	a	6.6	de
	top20M_WS20F	0.63	ab	0.23	0.03	0.98	0.02	4.89	ab	6.6	de
	WS_20MF	0.63	a,c	0.21	0.04	0.99	0.02	4.79	a,d	7.4	cd
	ALL_ms1	0.65	a	-0.27	0.05	1.14	0.02	4.66	c,e	8.3	bc
	top20_ms1	0.60	c,e	-0.04	0.05	1.03	0.02	4.62	de	8.5	ab
	top20_ms2	0.62	b,d	0.03	0.06	0.99	0.02	4.70	b,e	8.5	a,c
	top20_ms3	0.61	b,d	0.03	0.05	0.99	0.02	4.64	c,e	9.0	ab
	top20_ms4	0.62	b,d	-0.15	0.04	1.09	0.02	4.55	e	9.5	a
	top20M_bot20F_ms2	0.60	de	-0.06	0.04	1.02	0.02	4.66	c,e	8.4	a,c
15	BLUP	0.60	de	0.15	0.04	0.95	0.01	6.80	de	11.1	de
	RND20	0.62	a,d	0.37	0.06	0.95	0.02	6.98	cd	10.6	e
	top20	0.64	a	0.32	0.05	0.94	0.01	7.21	a,c	10.6	e
	top20_YD	0.63	ab	0.36	0.05	0.92	0.01	6.99	b,d	10.0	e
	top20M_bot20F	0.61	c,e	0.54	0.04	0.91	0.01	7.22	ab	10.3	e
	top20M_WS20F	0.63	a,c	0.43	0.05	0.92	0.01	7.23	a	10.8	de
	WS_20MF	0.61	c,e	0.48	0.06	0.91	0.02	7.01	a,d	12.2	cd
	ALL_ms1	0.64	ab	-0.27	0.06	1.09	0.02	6.83	de	13.1	c
	top20_ms1	0.61	c,e	-0.08	0.04	1.03	0.01	6.73	e	13.8	a,c
	top20_ms2	0.61	b,d	0.08	0.05	0.98	0.01	6.92	de	13.6	a,c
	top20_ms3	0.60	c,e	0.13	0.06	0.96	0.01	6.82	de	14.9	ab
	top20_ms4	0.60	de	-0.06	0.08	1.03	0.02	6.69	e	15.4	a
	top20M_bot20F_ms2	0.59	e	0.08	0.04	0.98	0.01	6.89	de	13.5	bc

[1]acc, b0 and b1= accuracy, bias and scale of genetic predictions, respectively; deltaG=accumulated genetic gain (in units of SD of true breeding values in the base population); IBD = average coefficient of genomic inbreeding (measured at IBD loci), in %. For a same statistic and generation, averages followed by the same letter (l) do not differ (adjusted p-value > 0.05). More than two consecutive letters in the same cell are presented as an interval (e.g. a,d = abcd).

[2]Each scenario was defined by a combination of genotyping strategy and genetic evaluation method (as described in Material and Methods, Table 1).

In terms of scale of predictions, multi-step evaluation also resulted in slightly greater departure from the results verified for BLUP when estimates of b1 in generation 5 were considered, especially for the scenario with all candidates being genotyped. The estimates of b0 were considerably larger for scenarios under prediction based on ssGBLUP, with estimates of b0 averaging 0.14, compared to 0 for BLUP and at most 0.05 for msGBLUP (Table 3). In this generation, the greatest estimates of IBD were verified for scenarios with multi-step evaluation and selection of the top 20% animals for DGV (regardless of the scale adjustment), which resulted in estimates about 25% larger than those verified for BLUP, while all genotyping strategies under ssGBLUP resulted in smaller increase in inbreeding (IBD coefficient about 11% lower than in BLUP).

At generation 10, ssGBLUP resulted in accuracy about 8% higher than BLUP, while the multi-step approach with all candidates genotyped (ALL_ms1) resulted in the greatest prediction accuracy (12% higher than BLUP, although not significantly different from ssGBLUP). However, due to the smaller genetic progress in the previous generations, genotyping all candidates provided accumulated gain only 1.3% larger than BLUP after 9 generations, while ssGBLUP for scenarios genotyping of the top 20% males for EBV (top20, top20M_bot20F and top20M_WS20F) provided the largest gains (between 4.9% and 6.6% larger than BLUP). Judged by the accumulated gain, the multi-step approach with genotyping of the top20% males combined with selection based on GEBV after scale adjustment (top20_ms2) was more beneficial than genotyping all candidates and using selection based on GEBV without scale adjustment (Table 3).

As a general rule, at generation 10, the single-step approach and the multi-step approach using scale adjustment resulted in predictions with slope (b1) closer to 1, when compared to BLUP, while the multi-step approach without adjustment resulted in deflated predictions (b1 > 1). Estimates of b0 considerably larger than 0 were verified under ssGBLUP, with the larger values (about 0.30) being observed for scenarios with genotyping of extreme animals or random sampling (top20M_bot20F and RND20), while the smallest value was observed for ALL_ms1 (about -0.30).

Because the same initial reference population was applied to all scenarios under genomic selection, significant differences between genotyping strategies could

be verified only when longer time horizons were considered. Under selection based on prediction using the single-step method, the genetic gain accumulated until the last generation, was significantly larger in the scenarios with genotyping of the top 20% males for EBV (top20, top20M_bot20F and top20M_WS20F) resulting in accumulated genetic gain up to 4% larger when compared to the other genotyping strategies and up to 6.5% larger than that verified under selection based on BLUP (Table 3).

In the last generation (generation 15), the best scenarios under multi-step prediction method only resulted in marginal increase in accumulated genetic gain when compared to BLUP, so that the largest gain over BLUP was about 1.9%, for the scenario with genotyping of the top 20% animals for EBV and adjustment of GEBV for the scale of EBV (top20_ms2) and this difference was not significant.

In this generation, the multi-step prediction without adjustment for the scale of EBV (ms1 and ms4) also resulted in a trend of deflation of predictions, while the estimated slope under all scenarios including BLUP was slightly lower than 1. As a general, for single-step prediction, there was no great difference between genotyping strategies in terms of scale, so that the estimated slope was close to that estimated under BLUP. The multi-step prediction with scale adjustment (ms2 and ms3) resulted in estimated slope close to 1 (between 0.96 and 0.98).

At generation 15, the estimated intercept (b0) was considerably larger for scenarios under selection based on single-step prediction compared to BLUP, while under multi-step prediction, only the scenario with all candidates genotyped resulted in estimated b0 with larger departure from 0 (-0.27), when compared to BLUP.

For both G10 and G15, the results regarding to genomic inbreeding confirmed the findings of the previous generations, so that multi-step prediction resulted in IBD coefficients significantly larger than BLUP, with the worse results verified for the scenario with genotyping of the top 20% animals for DGV (top20_ms4). Conversely, the single-step method resulted in IBD coefficients between 4% and 14% percent lower than BLUP in this period, although this difference was not significant (Table 3).

While the accuracies of different genotyping strategies were quite similar when all selection candidates were considered, larger differences between them were found when accuracies were computed considering genotyped and non-genotyped

animals separately (Figure 1). In the case of accuracy considering only the genotyped animals, there was a clearer trend of increase in accuracy along the generations for most strategies, as would be expected since more animals were included in the reference population.

Under the multi-step approach, for genotyped animals, it could be observed a clear advantage for genotyping extreme animals (topM_bot20F_ms2) over genotyping superior animals, which was even larger when genomic predictions were updated through genotyping superior animals and also no adjustment scale nor blending were applied (top20_ms4) (Figure 1). Although the genotyping strategies under the multiple step approach do not the predictions for non-genotyped animals from the same generation, different sampling strategies imply in slight differences in the set of non-genotyped animals (e.g. in mean and variance), what could partially explain the larger accuracies for non-genotyped candidates under genotyping of extreme animals, while the selection decisions under each scenario could also contribute to affect such differences in the long-term.

Notice that for the scenario with all animal being genotyped starting from generation 2 (ALL_ms1), some non-genotyped animals (from generations 0 and 1) figured among the selection candidates, reason why the figures for acc.cand and acc.geno were slightly different (Figure 1).

Under the single-step approach, genotyping extreme animals also resulted in a slight increase in the prediction accuracy for genotyped animals along most of the generations considered (Figure 1), while the worse results were verified for strategies based on random sampling both across all candidates (RND20) or within sire families (WS_20MF). The previous comment on the effect of each genotyping strategy on the makeup of the genotyped and non-genotyped sets of animals would also apply in this situation, although under the single-step approach the different genotyping strategies impact directly on the predictions of all animals, since the genetic values of non-genotyped animals are, a priori, conditioned on genetic values of the genotyped animals.

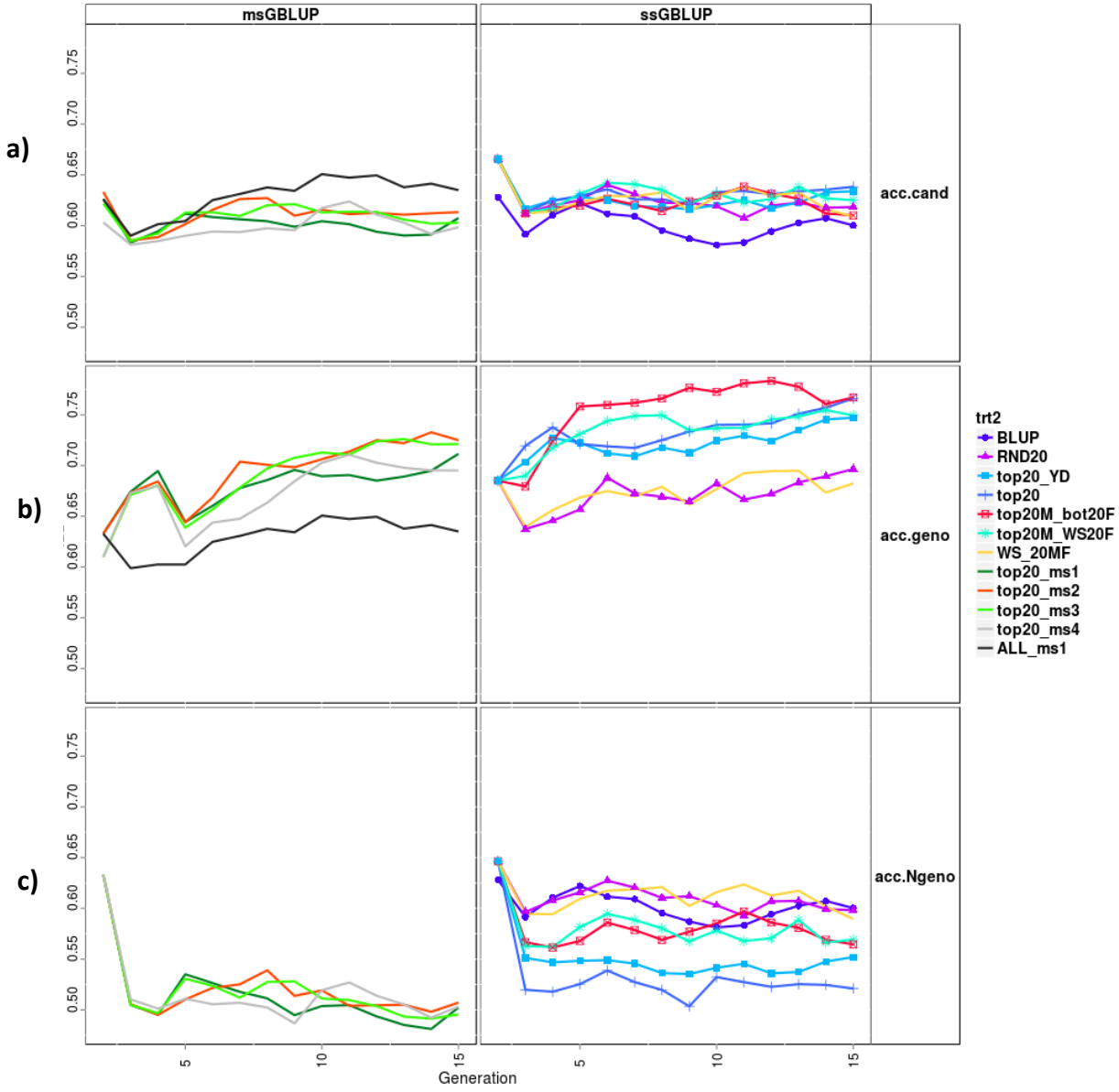


Figure 1. Accuracy of genetic prediction[1] under different scenarios[2]

[1] Pearson's correlation between true and estimated genetic merit of selection candidates. Average of ten replicates. The values were calculated considering either: a) all selection candidates in a given generation (acc.cand), b) only the genotyped candidates (acc.geno) or c) only the non-genotyped candidates (acc.Ngeno).

[2]Each scenario was defined by a combination of genotyping strategy and genetic evaluation method (as described in Material and Methods, Table 1).

The single-step approach resulted in predictions with very similar scale (estimated slope), regardless of genotyping strategy employed to update the reference population and of whether candidates were genotyped or not (Figure 2).

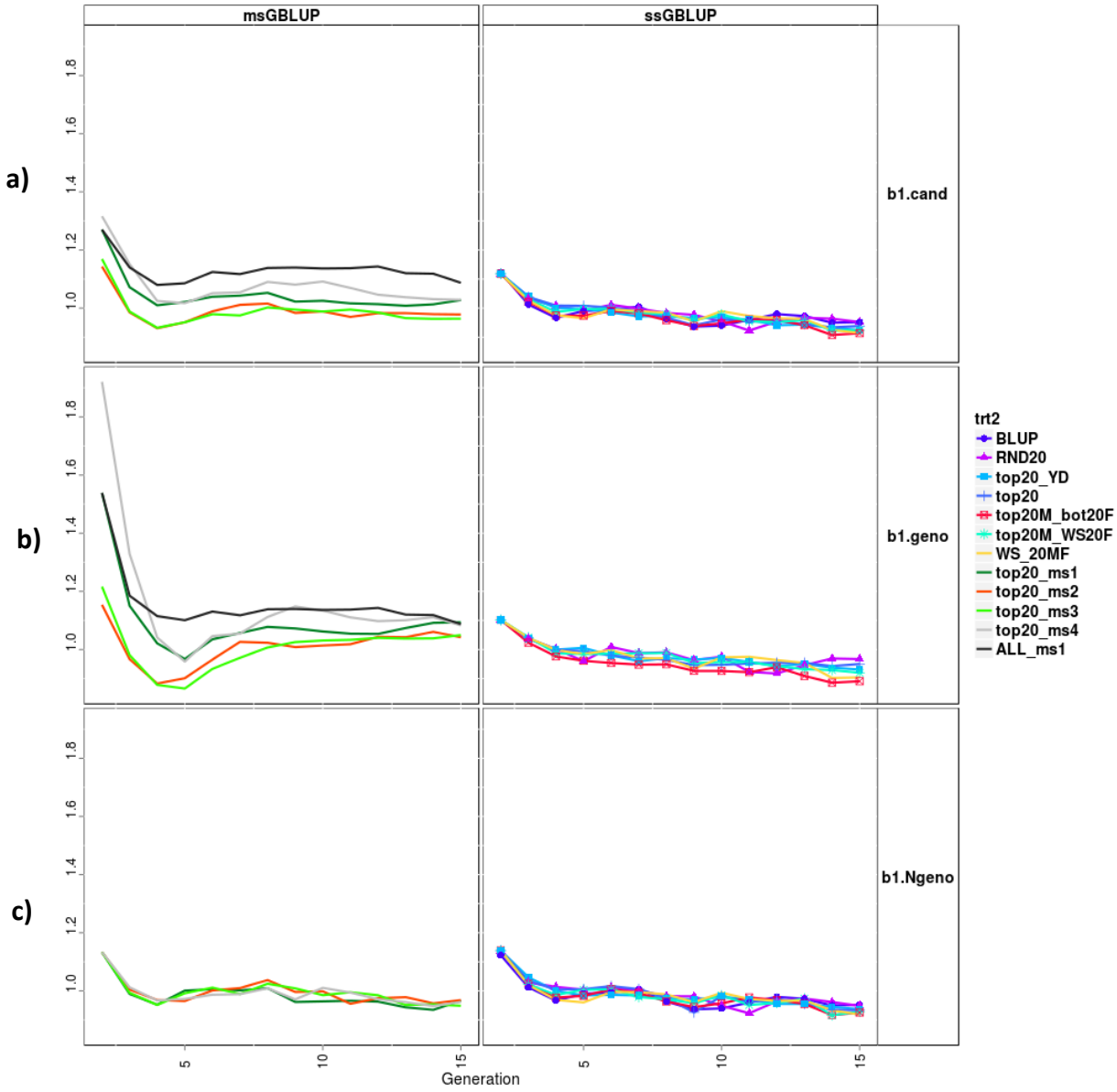


Figure 2. Scale of genetic predictions[1] under different scenarios[2]

[1] Estimated slope of the regression of true breeding value on the estimated genetic merit of selection candidates in each scenario. Average of ten replicates. The values were calculated considering either: a) all selection candidates in a given generation (b1.cand), b) only the genotyped candidates (b1.geno) or c) only the non-genotyped candidates (b1.Ngeno).

[2]Each scenario was defined by a combination of genotyping strategy and genetic evaluation method (as described in Material and Methods, Table 1).

The multi-step approach led to larger coefficients of scale for the genotyped animals when compared to non-genotyped animals, especially in the first two generations after update of the reference population, where much larger coefficients were estimated for the scenarios with genotyping superior animals only and with no scale adjustment (top_ms1 and top_ms4), indicating that genetic predictions were excessively regressed for the genotyped animas (Figure 2). For genotyped animals, after generation 8, the estimated slopes for all genotyping strategies under the multi-step approach were smaller (closer to 1) and remained within a narrower range, up to generation 15.

Discussion

While the small gains observed for using genomic selection over BLUP in the simulated scenarios of this study could create some doubt about the cost-effectiveness of GS in such situations, the simulated design was intended to enable comparisons among genotyping strategies under circumstances where the main problem was to evaluate the best strategy to update the reference population considering both short-term and long-term perspectives.

In such a situation, given the evidence that selection could lead to biased genomic predictions ([15], [16]) and the need to compare genotyped to non-genotyped selection candidates, it is important that the genotyping strategies allow more accurate genetic predictions for genotyped animals without introducing any bias that would lead to unfair comparisons between genotyped and non-genotyped animals.

When genotyping strategies were compared aiming to define the best reference population to predict the genetic merit of genotyped candidates from the next generation, there was some evidence that genotyping extreme animals for yield deviation would result in genomic predictions considerably more accurate and less biased than a reference population composed only by superior animals ([2], [3]). Such evidences could generate some concern about genotyping only elite animals in beef cattle breeding programs, as discussed earlier for the case of candidates to CEIP.

In the present study, because a larger proportion of selection candidates from each generation is non-genotyped (80%) and also due to the fact that all selective genotyping strategies are only employed to update the same previously established reference population, the differences were not expected to be so pronounced as in previous studies ([2], [3]), although they could provide some insight on the design of approaches to update reference populations in a suitable way for beef cattle schemes .

Under the single-step approach, the differences between genotyping strategies were not too pronounced in terms of accuracy and realized genetic gains, even if only genotyped animals were compared. According to Vitezica et al. [16], since the genetic values of non-genotyped animals are, a priori, conditioned on the genetic values of genotyped animals under the single-step approach [17], biases arisen from selection could be alleviated, and even more when, as in the present study, the set up of the relationship matrix was preceded by the adjustment proposed by [16] to allow more compatibility between pedigree-based and genomic relationships. Under such an approach, only in the last generations simulated it was possible to envisage a small advantage in terms of genetic gain (4%) in the scenarios in which reference population was updated trough genotyping of superior animals.

When only genotyped animals were considered, the results for accuracy under the multi-step prediction were more consistent with previous findings ([2], [3]), also suggesting more advantage of genotyping extreme animals over the strategy of only genotyping superior animals. Anyway, such differences did not reflect in the long-term genetic gain (when the true genetic merit of animals born each generation was evaluated), since they did not account exactly for the selection decisions practiced in the different scenarios. As an overall result, no important differences were identified between strategies in terms of genetic gain in the last generations, although a more limited number of genotyping scenarios was compared in the multi-step approach.

For the multi-step approach, especially in the short-term, the prediction of genetic merit of genotyped selection candidates without any adjustment for the scale of EBVs resulted in predictions more regressed when compared to the predictions obtained after adjustment, what could lead to more unfair comparisons between genotyped and non-genotyped animals in the former case. As a general rule,

scenarios using scale adjustment presented coefficients of inflation of genomic predictions much closer to those verified under single-step prediction and BLUP.

The blending strategy investigated [22] tended to result in a slightly advantage in terms of accumulated genetic gain and inbreeding incidence, although no significant differences were found to support the recommendation of such strategy. Another blending strategies were proposed (e.g. [5]) and were not investigated in this study for the sake of simplicity. A future study on genotyping strategies in beef cattle could investigate the long-term consequences of different blending strategies under the multi-step approach.

There is some criticism about the use of EBV to estimate marker effects. Due to the fact that EBVs are regressed toward the mean depending upon accuracy [27], excessively shrunk estimates of the marker effects could be obtained, thus generating deflated genomic predictions (i.e. too regressed when compared to the true breeding value). Garrick et al. [28] highlighted other potential problems of this approach, among which the problem of heterogeneity of variances resulting from EBVs being based on different amounts of information, and proposed using deregressed EBVs (dEBV) to estimate marker effects. While the present implementation of the multi-step approach involved weights proportional to EBV accuracies, in the scenarios without scale adjustment it was clear that such approach still resulted in predictions too regressed for the genotyped animals.

Because the non-genotyped animals were selected based on EBV under the multi-step approach and in order to avoid the inclusion of an additional factor in the simulation (response variable), we opted to conduct this study using EBV as response variable in msGBLUP. Future studies on selective genotyping strategies should investigate other alternatives for the response variable in msGBLUP (e.g. dEBV), what could also require to study strategies to allow proper comparison between predictions of genotyped and non-genotyped animals.

At our knowledge, no previous study evaluated consequences of selective genotyping strategies on inbreeding incidence. For a simulated trait of moderate heritability, in which own performance information was available before selection on candidates of both sexes, it was verified that the single-step approach exhibited significant advantage over BLUP in the long-term with respect to inbreeding

incidence, estimated after monitoring identity-by-descent in a large number of loci, being that no important differences were found between genotyping strategies under this two-step approach. For scenarios under multiple-step prediction, when genomic predictions were obtained with the same method (after blending and scale adjustment), no differences were verified between genotyping superior or extreme animals for EBV in terms of inbreeding incidence, while all scenarios of multi-step prediction resulted in considerably larger coefficients than BLUP.

Conclusions

Under comparable selective genotyping strategies to update the reference population, the use of the single-step approach provided more accurate and less deflated predictions than a multiple-step approach using EBVs to estimate marker effects. The single-step approach also presented advantage over multi-step prediction in terms of inbreeding incidence and accumulated genetic gain.

In the long term, there was no evidence of clear advantage for a particular genotyping strategy under investigation, although, under the single-step approach, scenarios with genotyping of superior males presented larger accumulated genetic gain until the last generation simulated. Thus, under the single-step approach, there was no evidence that would disqualify updating the reference population trough only genotyping superior animals for EBV.

References

1. Ehsani A, Janss L, Christensen OF: **Effects of selective genotyping on genomic prediction.** In *World Congress on Genetic Applied to Livestock Production*, 2–7 August 2010, Leipzig. Abstract 444.

2. Boligon AA, Long N, Albuquerque LG, Weigel KA, Gianola D, Rosa GJM: **Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection.** *J Anim Sci* 2012**, 90**: 4716-4722.

3. Jiménez-Montero JA, González-Recio O, Alenda R: **Genotyping strategies for genomic selection in small dairy cattle populations.** *Animal* 2012, **6**:1216–1224.

4. Carvalheiro R, McEwan JC, Neves HHR, Utsunomiya YT, Pérez O'Brien AM, Boison SA, Sölkner J, Schenkel FS, Van Tassell CP, Sonstegard TS, Garcia JF:

Accuracy of genomic predictions in Nellore beef cattle. In *Proceedings of the 20th AAABG Conference*. 2013, Napier, New Zealand.

5. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS: **Invited review: Reliability of genomic predictions for North American Holstein bulls.** *J Dairy Sci* 2009, **92**:16–24.

6. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME: **Invited review: Genomic selection in dairy cattle: Progress and challenges.** *J Dairy Sci* 2009, **92**:433-443.

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

8. Muir WM: **Comparison of genomic and traditional BLUP estimated breeding value accuracy and selection response under alternative trait and genomic parameters.** *J Anim Breed Genet* 2007, 124:342–355.

9. Goddard ME, Hayes BJ, Meuwissen THE: **Genomic selection in livestock populations.** *Gen Res Camb* 2010, **92**: 413-421.

10. Brito FV, Braccini Neto J, Sargolzaei M, Cobuci JA, Schenkel FS: **Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle.** *BMC Genet* 2011, **12**:80.

11. Bastiaansen JWM, Coster A, Calus MPL, Van Arendonk JAM, Bovenhuis H: **Long-term response to genomic selection: effects of estimation method and reference population structure for different genetic architectures.** *Genet Sel Evol* 2012, **44**:3.

12. Sonesson AK, Wooliams JA, Meuwiseen THE: **Genomic selection requires control of inbreeding.** *Genet Sel Evol* 2012, **44**:27.

13. Wiggans GR, Sonstegard TS, VanRaden PM, Matukumalli LK, Schnabel RD, Taylor JF, Chesnais JP, Schenkel FS, Van Tassell CP: **Genomic Evaluations in the United States and Canada: A Collaboration.** In *International Committee on Animal Recording (ICAR*), June 22–26, 2009. Munich: ICAR Tech Ser; 2009:347-353.

14. VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, Sonstegard TS, Connor EE, Winters M, Van Kaam J, Van Doormaal BJ, Faust MA, Doak GA: **Genomic imputation and evaluation using high density Holstein genotypes.** *J Dairy Sci* 2013, **96**:668-678.

15. Patry C, Ducrocq V: **Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle.** *J Dairy Sci* 2011, **94**:1011–1020.

16. Vitezica ZG, Aguilar I, Misztal I, Legarra A: **Bias in genomic predictions for populations under selection.** *Genet Res* 2011, **93**:357-366.

17. Legarra A, Aguilar I, Misztal I: **A relationship matrix including full pedigree and genomic information.** *J Dairy Sci* 2009, **92**:4656-4663.

18. Forni S, Aguilar I, Misztal I: **Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information.** *Genet Sel Evol* 2011, **43**:1.

19. Misztal I: **Finding the balance between quantitative and genomics issues – a case of a single-step genomic evaluation (ssGBLUP)** In *9 th Biennial Symposium of the Brazilian Society of Animal Breeding*, June 20-22, 2012, João Pessoa. [http://sbmaonline.org.br/anais/ix/palestras/pdf/Palestra10.pdf]

20. Sargolzaei M, Schenkel FS, Jansen GB, Schaeffer LR: **Extent of linkage disequilibrium in Holstein Cattle in North America.** *J Dairy Sci* 2008, **91**:2106-2117.

21. Clark SA, Hickey JM, Daetwyler HD, Van Der Werf JHJ: **The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes.** *Genet Sel Evol* 2012, **44**:4.

22. Sullivan P: **Options for Combining Direct Genomic and Progeny-Test Results.** In *Canadian Dairy Network Genetic Evaluation Board Meeting,* June, 2009. [http://cgil.uoguelph.ca/dcbgc/Agenda0910/OptionsCombining.pdf]

23. Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS, Su G: **Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population.** *Genet Sel Evol* 2012, **44**: 8.

24. VanRaden P: **Efficient methods to compute genomic predictions.** *J Dairy Sci* 2008, **91**:4414 – 4423.

25. Sargolzaei M, Schenkel FS, VanRaden PM. **gebv: Genomic breeding value estimator for livestock**. *In Technical Report to the Dairy Cattle Breeding and Genetics Committee***.** Guelph: University of Guelph; 2009.

26. Misztal I: **BLUPF90 - a flexible mixed model program in Fortran 90.** 2012. [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90.pdf]

27. Daetwyler HD, Calus MP, Pong-Wong R, de Los Campos G, Hickey JM: **Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking.** *Genetics* 2013, **193**:347-365.

28. Garrick DJ, Taylor JF, Fernando RL: **Deregressing estimated breeding values and weighting information for genomic regression analyses.** *Genet Sel Evol* 2009**, 41**:55.

CHAPTER 6 - Final considerations

Different studies were carried out using publicly available or simulated data, aiming to investigate alternative strategies to improve the efficiency of genomic selection (GS) in animal breeding programs. Some final considerations will be made in this chapter.

In the study presented in Chapter 2, it was verified that close relationships between reference population and testing animals were decisive to achieve greater predictive ability of GS, even in a population with relatively high extent of linkage disequilibrium. Similar results have been reported for application of GS in other species as well as using simulated data. Although the extent at which different factors contribute to predictive ability of GS still deserves further investigation, from a practical perspective, these findings suggest the need of dynamic training in applications of GS in livestock populations, so that the animals selected for reproduction at each selection cycle are routinely incorporated to the reference population, what would enhance predictions for young animals.

The results of Chapter 2 also evidenced a re-ranking of statistical methods used for genomic prediction as a function of the trait analyzed. This finding reinforces the importance of testing different statistical methods before application of GS, so that the most suitable methods for each trait can be identified and incorporated to routine evaluations.

Based on the results of Chapter 3, the use of deregressed EBV (dEBV) can be suggested as the predictand in multi-step genomic prediction, for both model training and for validation. However, in a situation in which selection candidates consist of genotyped and non-genotyped animals, it is very important that genomic predictions are in the same scale of the traditional predictions (EBV) of the non-genotyped candidates. In some situations, the EBV of non-genotyped animals can be more regressed than genomic predictions derived using dEBV as predictand and thus both types of proofs would not be comparable. While an ad-hoc scale adjustment of genomic predictions was investigated in Chapter 5, aiming to enable comparison among genotyped and non-genotyped candidates, previous studies proposed other

strategies that address the same problem, e.g. single-step genomic evaluation, blending strategies based on selection index theory and the inclusion of genomic predictions as pseudo-records in bivariate models. These different strategies can be more or less appealing, depending on aspects like reliability, computational feasibility and ease to incorporate into routine evaluations. Further studies are needed to explore this question in detail. A future simulation study could compare these alternative strategies, aiming to identify a more consistent framework to compare genotyped and non-genotyped selection candidates.

The results of Chapter 4 highlight opportunities for using genomic selection in beef cattle, especially to improve meat quality and female reproduction traits. A stochastic simulation was carried out, considering the use of a dynamic training population and giving more support to predict long-term consequences of GS in beef cattle populations. It was shown that, even when high-quality pedigree information is available, the use of genomic information can be a more valuable tool to monitor inbreeding incidence and to design strategies to maintain genetic diversity in the long-term.

Long-term consequences of selective genotyping strategies to update a pre-established reference population were investigated in Chapter 5. No clear advantage was verified for a particular genotyping strategy in the long-term, although a slightly benefit was found in terms of genetic progress for scenarios in which superior males were genotyped, when using single-step prediction. The impact of selective genotyping can be more pronounced when a strategy is employed to define the whole reference population. This could be the case of meat quality traits, where phenotyping requires slaughtering animals and an indicator trait can be used to define the genotyped animals. Under dynamic training, especially if phenotypes are available on selection candidates, it can be more important to choose the genotyped animals aiming to ensure that parents or close relatives of the selection candidates will be in the reference population, given the influence of relatedness to the reference population on the accuracy of genomic predictions. Future studies could investigate this question in more detail, by simulating selective genotyping in different selection schemes and for different traits.