

# Accuracy of genomic prediction of maternal traits in pigs using Bayesian variable selection methods

Maria V. Kjetså<sup>1</sup>  | Arne B. Gjuvslund<sup>2</sup>  | Øyvind Nordbø<sup>2</sup>  | Eli Grindflek<sup>2</sup> | Theo Meuwissen<sup>1</sup>

<sup>1</sup>Norwegian University of Life Sciences, Faculty of Biosciences, Ås, Norway

<sup>2</sup>Norsvin SA, Hamar, Norway

## Correspondence

Maria V. Kjetså, Norwegian University of Life Sciences, Faculty of Biosciences, Ås, Norway.

Email: [maria.kjetsa@nmbu.no](mailto:maria.kjetsa@nmbu.no)

## Funding information

Norges Forskningsråd, Grant/Award Number: 255297

## Abstract

The aim of this study was to compare three methods of genomic prediction: GBLUP, BayesC and BayesGC for genomic prediction of six maternal traits in Landrace sows using a panel of 660 K SNPs. The effects of different priors for the Bayesian methods were also investigated. GBLUP does not take the genetic architecture into account as all SNPs are assumed to have equally sized effects and relies heavily on the relationships between the animals for accurate predictions. Bayesian approaches rely on both fitting SNPs that describe relationships between animals in addition to fitting single SNP effects directly. Both the relationship between the animals and single SNP effects are important for accurate predictions. Maternal traits in sows are often more difficult to record and have lower heritabilities. BayesGC was generally the method with the higher accuracy, although its accuracy was for some traits matched by that of GBLUP and for others by that of BayesC. For piglet mortality within 3 weeks, BayesGC achieved up to 9.2% higher accuracy. For many of the traits, however, the methods did not show significant differences in accuracies.

## KEYWORDS

Bayesian genomic prediction, genomic prediction accuracy, genomic selection, maternal traits

## 1 | INTRODUCTION

Genomic prediction (GP) (Meuwissen et al., 2001) is a method to predict breeding values (GEBVs) in animal and plant breeding. GP predicts the GEBVs by using a reference population of animals with both phenotype and marker information to estimate marker effects. Meuwissen et al. proposed three methods for genomic prediction: two Bayesian variable selection methods (BayesA and B) and a linear marker effects model estimating marker effects from single-nucleotide polymorphisms (SNPs) using best

linear unbiased prediction (BLUP), referred to as SNP-BLUP. An alternative method of SNP-BLUP is to use a marker-derived genomic relationship matrix (often called a **G**-matrix) as a covariance matrix when solving mixed-model equations (MME) (VanRaden, 2008) referred to as genomic best linear unbiased predictions (GBLUP). The two methods, SNP-BLUP and GBLUP, are mathematically equivalent (Strandén & Garrick, 2009; VanRaden, 2008).

All markers are assumed to have equal weight in the prediction for the linear models, while Bayesian methods try to differentiate SNPs relative to their importance.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Animal Breeding and Genetics* published by John Wiley & Sons Ltd.

Markers associated with causal mutations get a higher relative weight, and markers not linked to causal loci are down-weighted, thus only giving weights to the most important SNPs (Meuwissen et al., 2001; Verbyla et al., 2010). Several alternative Bayesian variable selection methods are proposed, often referred to as the “Bayesian Alphabet” (Gianola et al., 2009). Differences between the methods are the prior distributions used for the estimation of SNP effects. For example, BayesA uses one t-distribution for SNP effects, while BayesB has a mixture of a t-distribution with probability  $\pi$ , and a null effect with probability  $1-\pi$ . BayesC (Habier et al., 2011) is similar to BayesB, as both have a mixture distribution prior, where one has a null effect. However, BayesC uses a normal distribution instead of a t-distribution and assumes a common variance for all SNPs, while BayesB assumes SNP-specific variances. BayesR uses four normal distributions, where one of them has a null effect (Erbe et al., 2012). The recently proposed BayesGC method (Meuwissen et al., 2021), fits a polygenic effect through a **G**-matrix in addition to a BayesC term. Hence, BayesGC fits many SNPs with a small effect through the **G**-matrix and a group of SNPs selected by the model with more significant effects through the BayesC term.

In this paper, we look at the genomic prediction of maternal traits in Landrace pigs, which are considered complex traits with a low to moderate heritability and explore the effect of the genetic architecture on the prediction accuracy. Specifically, we look at the traits; the total number of born piglets (**TNB**), number of stillborn piglets (**STB**), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (**M3W**), total litter weight at 3 weeks (**LW3W**), sow shoulder lesions (**SHL**) and the sow's body condition score (**BCS**). These maternal traits were included in the breeding goal for Topigs Norsvin at the time of recording (Eriksen, 2018).

Maternal traits in pigs are related to the sow's ability to produce and raise offspring. Maternal traits are essential for efficiency in pig production, the economy and animal welfare (Ocepek & Andersen, 2017). An ideal sow produces a litter of piglets corresponding to the number of functional teats available, and all the piglets born survive until weaning. Furthermore, the piglets should grow evenly, and the sow should not spend all her resources on the litter, implying that she maintains a good body condition score and does not develop shoulder lesions.

Simulation studies have shown great potential for using genomic prediction methods to predict maternal traits in pigs (Lillehammer et al., 2011, 2013). Although few studies have reported genomic prediction accuracies for maternal traits in pigs (Tan et al., 2017), there are very few that have reported prediction accuracies for Bayesian genomic prediction methods for maternal traits. Some have looked

**TABLE 1** Number ( $n$ ) of animals with records for each trait and partition into the reference and validation population, and mean ( $m$ ) number of parity records in each trait

Trait <sup>a</sup>	Total, $n$	Reference, $n$	Validation, $n$	Parity, $m$
TNB	15,690	14,513	1177	2.5
STB	15,690	14,513	1177	2.5
M3W	10,611	9466	1145	1.7
LW3W	10,804	9656	1148	1.7
SHL	15,084	13,934	1150	2.2
BCS	15,084	13,933	1151	2.2

<sup>a</sup>Total number born (TNB), number of stillborn piglets (STB), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (M3W), total litter weight at 3 weeks (LW3W), sow shoulder lesions (SHL) and the sow's body condition score (BCS).

at Bayesian methods in growth and reproduction traits (Song et al., 2017) and slaughter traits (Salek Ardestani et al., 2021). Although the basis of inheritance and breeding is the same across livestock species, their differences in breeding structure, genetic architecture and trait biology make it important to study the different prediction methods across the species (Samorè & Fontanesi, 2016).

This study aimed to determine the prediction accuracy of six maternal traits in Landrace sows using a panel of 660k SNP markers and a large reference population (9–15 thousand reference animals). The study also compares three methods of genomic prediction: GBLUP (VanRaden, 2008), BayesC (Habier et al., 2011) and BayesGC (Meuwissen et al., 2021).

## 2 | MATERIALS AND METHODS

### 2.1 | Phenotypic data

The phenotypic data consisted of records from 15,703 unique individual Landrace sows with at least one record for one of the six traits; the total number of born piglets (**TNB**), number of stillborn piglets (**STB**), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (**M3W**), total litter weight at 3 weeks (**LW3W**), sow shoulder lesions (**SHL**) and the sow's body condition score (**BCS**). Of the 15,703 sows, 10,306 had records for all six traits. The traits were recorded between 2008 and 2019. Each of the different traits had between 10,611–15,690 phenotypic records (see Table 1).

Yield Deviations (VanRaden & Wiggans, 1991) for the six traits were derived from the commercial breeding value evaluations from Topigs Norsvin. There were multiple records for each trait, as we had one YD for each parity. The maximum number of parities recorded for each

trait was 6, and the mean number of parities recorded for each trait is shown in Table 1. Because the software used for the Bayesian variable selection models (Meuwissen et al., 2021) could not handle multiple records per animal, we used the average YD for each sow, with a weighting of each record corresponding to the effective number of records calculated by the formula  $\frac{n \cdot (1 + \lambda)}{(n + \lambda)}$  where  $\lambda$  is  $\sigma_e^2 / \sigma_{pe}^2$  and  $n$  is the number of records for each individual,  $\sigma_e^2$  is the residual variance and  $\sigma_{pe}^2$  is the permanent environmental variance ( $\lambda$  was obtained from Topigs Norsvin's breeding value evaluation).

## 2.2 | Genotype data

The sows were genotyped with varying SNP densities and imputed to a 660 K-genotype density. Of the 15,703 sows, 526 were genotyped on a 10 K chip (GGP Porcine LD), and the rest were genotyped on medium density chips: Illumina PorcineSNP60 (60 K) and two Illumina GeneSeek custom chips (80 K and 50 K). All genotypes were imputed using Fimpute v2.2 (Sargolzaei et al., 2014), first to the 50 K chip, and then to the 660 K Axiom Porcine Genotyping Array with reference genotypes from 467 Landrace animals. After quality control, the 660 K High-Density genotype data had a total of 429,403 SNPs with MAF > 0.01.

## 2.3 | Validation and reference data

The ~1000 youngest sows were used for validation of the predictions, in order to imitate a typical genomic breeding program where one wishes to predict the breeding values of young animals before they have their own recorded traits. This was done by masking their phenotypic records in the analysis and using them for validation. The number of validation sows was between 1145 and 1177 (Table 1). The rest of the animals was used as the reference data with both phenotypic and genotype records. Our smallest reference dataset consisted of 9466 animals for the trait M3W and the largest of 14,513 animals for traits TNB and STB (Table 1).

## 2.4 | Prediction accuracy and regression coefficients

The accuracy of prediction for all methods was estimated as:

$$r_{\text{pred}} = \frac{\text{cor}(\text{GEBV}, \text{YD})}{\sqrt{h^2}},$$

and the bias (coefficient of regression) was the calculated slope ( $b$ ) of the linear regression  $Y = a + bX$ , where  $a$  is the intercept,  $Y$  is the yield deviation (YD) and  $X$  is the genomic estimated breeding value (GEBV) of the sows in the validation datasets, estimated with only marker information and not phenotypic records.  $h^2$  is the heritability of the trait and was estimated on the full dataset.

## 2.5 | Variance components and GBLUP

We estimated variance components for each trait using the pedigree relationship matrix and the DMUAI package from the DMU software (Madsen & Jensen, 2013). The variance components were estimated on the full dataset (i.e. both the reference and validation animals). The model for the variance component estimations was as follows:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

where  $\mathbf{y}$  is a vector of the average YD of a sow,  $\mathbf{1}$  is a vector of ones corresponding to the size of  $\mathbf{y}$ ,  $\mu$  is the mean,  $\mathbf{Z}$  is a design matrix linking individuals to the phenotype,  $\mathbf{u}$  is the random effect of the individual animal ( $\mathbf{u} \sim N(0, \mathbf{A}\sigma_u^2)$ ), where  $\mathbf{A}$  is the pedigree relationship matrix and  $\mathbf{e}$  is the residual effect ( $\mathbf{e} \sim N[0, \mathbf{D}\sigma_e^2]$ ), where  $\mathbf{D}$  is a diagonal matrix where the diagonals are the inverses of the effective number of records. The same model was used for the GBLUP analyses except that the individual animal effect was modelled as ( $\mathbf{u} \sim N[0, \mathbf{G}\sigma_u^2]$ ). The variance components used were from the above pedigree-based estimates. The  $\mathbf{G}$ -matrix was calculated using the VanRaden method 1 (VanRaden, 2007).

## 2.6 | BayesC

The model for BayesC (Habier et al., 2011) was:

$$\mathbf{y} = \mathbf{1}\mu + \sum_i \mathbf{I}_i \mathbf{x}_i s_i + \mathbf{e}.$$

where  $\mathbf{y}$  is a vector of Yield Deviations,  $\mathbf{1}$  is a vector of ones,  $\mu$  is the overall mean,  $\mathbf{x}_i$  is a vector of genotypes for SNP  $i$  containing  $-2p_i$  for homozygote individuals,  $1-2p_i$  for heterozygotes and  $2-2p_i$  for the alternative homozygote genotype with  $p_i$  being the allele frequency of SNP  $i$ , and  $\mathbf{I}_i$  is an indicator of whether the SNP  $i$  is in the model in a particular MCMC cycle or not (0/1), where the prior probability of  $\mathbf{I}_i$  being equal to 1 is denoted by  $\pi$  (values in Table 2),  $s_i$  is the SNP effect, where if the SNP  $i$  is in the model:  $s_i \sim N(0, \sigma_m^2)$ ,  $\mathbf{e}$  is the residual with variance  $\mathbf{e} \sim N(0, \mathbf{D}\sigma_e^2)$ , where  $\mathbf{D}$  is a diagonal matrix where the diagonals are the inverses of the effective number of records and  $\sigma_e^2$  is the residual

**TABLE 2**  $\pi$  values used for BayesC and BayesGC methods at different fractions of total genetic variance explained by a single-fitted SNP (Fr)

Fr	BayesGC_10 <sup>a</sup>	BayesGC_50 <sup>b</sup>	BayesGC_90 <sup>c</sup>	BayesC
1/100	0.00002	0.00012	0.00021	0.0002
1/500	0.00012	0.00058	0.00105	0.0012
1/1000	0.00023	0.00116	0.00210	0.0023
1/5000	0.00116	0.00582	0.01048	0.0116
1/10,000	0.00233	0.01164	0.02096	0.0233

<sup>a</sup>BayesGC\_10 is Bayes\_GC with 10% marker variance and 90% polygenic variance.

<sup>b</sup>BayesGC\_50 has 50% marker variance and 50% polygenic variance.

<sup>c</sup>BayesGC\_90 has 90% marker variance and 10% polygenic variance.

variance estimated from the variance component estimations (Table 5). The Markov Chain Monte Carlo (MCMC) chain was run for 20,000 Gibbs cycles using 4000 burn-in cycles, in two distinct chains.

We used the same variance components as for the GBLUP analyses; however, the total genetic variance  $\sigma_u^2$  was partitioned. In the following, we describe how the total genetic variance  $\sigma_u^2$  (see Table 4) is partitioned over the fitted SNPs for the Bayes C method:

$$\sigma_m^2 = \frac{Fr \cdot \sigma_u^2}{\overline{HET}},$$

where  $\sigma_m^2$  is the genetic variance explained by a single SNP,

Fr is the fraction of the total genetic variance explained by a single fitted SNP, i.e. 1/1000 when we assume each SNP explains 1/1000th of the genetic variance. We test different values of Fr, namely 1/100, 1/500, 1/1000, 1/5000 and 1/10,000, respectively.

$$\overline{HET} = \text{average heterozygosity} = \frac{2 \sum p_i (1 - p_i)}{N_{loci}},$$

where  $p_i$  is the allele frequency of locus  $i$  and  $N_{loci}$  is the total number of loci.

For a Bayes C model, this would mean using a prior probability of fitting an SNP of:

$$\pi_c = \frac{1/Fr}{N_{loci}}.$$

Such that the total genetic variance is  $\sigma_u^2 = \pi_c \cdot N_{loci} \cdot \overline{HET} \cdot \sigma_m^2$ .

## 2.7 | BayesGC

The BayesGC model is as follows:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \sum_i I_i \mathbf{x}_i s_i + \mathbf{e},$$

where  $\mathbf{y}$  is a vector of the Yield Deviations,  $\mathbf{1}$  is a vector of ones,  $\mu$  is the overall mean,  $\mathbf{Z}$  is a design matrix that links individuals to the  $\mathbf{y}$ ,  $\mathbf{u}$  is a vector of random polygenic effects with variance  $V(\mathbf{u}) = \mathbf{G}\sigma_{pol}^2$ ,  $\mathbf{x}_i$  is the vector of genotypes for SNP  $i$  coded as for BayesC.  $I_i$  is an indicator of whether SNP  $i$  is in the model in an MCMC cycle or not (0/1) and the prior probability of  $I_i$  being equal to 1 is  $\pi$  (listed in Table 2),  $s_i$  is the SNP effect, where if the SNP  $i$  is in the model:  $s_i \sim N(0, \sigma_m^2)$ ,  $\mathbf{e}$  is the residual with variance  $\mathbf{e} \sim N(0, \mathbf{D}\sigma_e^2)$  where  $\mathbf{D}$  is a diagonal matrix where the diagonals are the inverses of the effective number of records and  $\sigma_e^2$  is the residual variance estimated from the variance component estimations (Table 5). The MCMC chain was run for 4000 burn-in cycles and a total of 20,000 Gibbs cycles for two independent chains. The EBVs from the two Gibbs chains for both BayesC and BayesGC had a correlation of  $>0.9999$ , and thus, the EBVs were assumed to be converged, and the results presented for both BayesC and BayesGC are the average of two Gibbs chains.

The BayesGC model basically fits the previous two models (GBLUP and BayesC) simultaneously, i.e. it fits a polygenic and a BayesC term. The polygenic effect is fitted using the genomic relationship matrix ( $\mathbf{G}$ ) as in the GBLUP model. The BayesC term assumes SNPs to have normally distributed effects with probability ( $\pi$ ) or an effect of 0 with probability ( $1 - \pi$ ).

In the following, we describe how the total genetic variance  $\sigma_u^2$  is partitioned over the fitted SNPs and the polygenic effect, following the method described in Kjetså et al. (2020). For BayesGC, we need an assumption on the fraction of the variance that is explained by the individually fitted SNPs in the BayesC term of the model. In addition, the total genetic variance  $\sigma_u^2$  should not be affected by the partitioning of the variance across the SNPs and the polygenic effect. Let  $q$  be the fraction of  $\sigma_u^2$  explained by the BayesC term, then the variance explained by the polygenic effect is  $\sigma_{pol}^2 = (1 - q)\sigma_u^2$ . Hence,

$$\sigma_u^2 = \sigma_{pol}^2 + q \cdot \pi \cdot loci \cdot \overline{HET} \cdot \sigma_m^2,$$

It follows that  $\pi_{gc} = q \cdot \pi_c$ ,



where  $\pi_{gc}$  is the  $\pi$  value used for the BayesGC method. Four different values of  $q$  were tested for BayesGC,  $q = 0.1, 0.5$  and  $0.9$  corresponding to the BayesC term with fitted marker effects explaining 10%, 50% or 90% of the total genetic variance (denoted BayesGC\_10, BayesGC\_50, BayesGC\_90, respectively), with the rest of the variance  $1 - q$  explained by the polygenic effect through the G-matrix. The values of  $\sigma_m^2$  used are shown in Table 3 and the values of  $\sigma_{pol}^2$  are shown in Table 4.

### 3 | RESULTS

The heritabilities of the traits ranged from 0.09 (M3W) to 0.34 (SHL) (Table 5). M3W had the lowest heritability of 0.09, followed by STB and TNB with moderate heritabilities of 0.13 and 0.19. LW3W, BCS and SHL had the highest heritabilities with 0.31, 0.31 and 0.34, respectively (Table 5). For the trait total number born (TNB), the highest accuracy was achieved at 0.610–0.611 for GBLUP and BayesGC\_10 (Table 6) and the method giving the lowest prediction accuracy is BayesC (Fr 1/100) (Figure 1), which achieved an accuracy of 0.515 for TNB. For all the Bayesian methods (BayesGC\_10, BayesGC\_50, BayesGC\_90 and BayesC), fitting more SNPs (Fr = 1/10000) gave the highest accuracy of prediction for the trait TNB. The accuracy of prediction for Stillborn (STB; Figure 2), is lower than the other traits.

For STB, there were also minor, but no significant differences in accuracy between the methods, with the highest accuracy achieved by BayesGC\_10 (Fr 1/5000) and GBLUP at 0.318 and the lowest accuracy for STB was BayesC (Fr 1/100) at 0.272 (see Figure 2). M3W (Figure 3) is the trait with the largest differences between the methods. GBLUP and BayesC (Fr 1/500) had an accuracy of 0.441 and 0.464, respectively, while the highest accuracy from the BayesGC methods was achieved by BayesGC\_50 (Fr 1/100) with an accuracy of 0.484 (Table 6), making a difference of 9.8% between GBLUP and BayesGC. However, the difference was not significant. For the trait LW3W (Figure 4) all the

methods had high accuracies of 0.717, 0.722 and 0.718 for the methods GBLUP, BayesGC\_90 (Fr 1/10,000) and BayesC (Fr 1/10,000), respectively.

Shoulder Lesions (SHL) (Figure 5) showed a prediction accuracy of 0.406 and 0.409 for GBLUP and BayesC while the highest accuracy for BayesGC was 0.418 for BayesGC\_50 (Fr1/100). Trait BCS (Figure 6) also had minor differences between the methods and obtained the highest accuracy from the BayesC and BayesGC\_90 (Fr 1/10,000) methods with an accuracy of 0.518 for both methods, while GBLUP obtained an accuracy of 0.511.

The regression coefficients for the method yielding the highest accuracy were also the regression coefficient closest to 1 for all the traits except L3W3. L3W3 was the only trait with a regression coefficient above 1, indicating the variance of the GEBV being slightly lower compared to the yield deviations. TNB was the trait with a regression coefficient closest to 1 with a regression coefficient of 0.966 while SHL was the trait with a regression coefficient furthest from 1 at 0.436 (see Table 6), implying that the variance of the GEBV for SHL was inflated.

### 4 | DISCUSSION

#### 4.1 | Genomic prediction methods

We have compared GBLUP, BayesGC and BayesC for six maternal traits with different priors for BayesC and BayesGC. In general, for all traits, one of the BayesGC methods yielded the highest prediction accuracy (Table 6), although its accuracy was often, but not always, matched by GBLUP and for one trait (BCS) by BayesC. This implies that fitting a combination of individual SNPs with large effects and a polygenic effect often yielded the highest prediction accuracy; however, the differences were not significant. The traits M3W and SHL yielded a 9.8% and 3.0% increase in accuracy when moving from GBLUP to BayesGC\_50 (Fr1/100) (Table 6). The trait BCS had a somewhat increased accuracy of prediction

Fr	$\sigma_m^2$					
	TNB <sup>a</sup>	STB <sup>a</sup>	M3W <sup>a</sup>	LW3W <sup>a</sup>	SHL <sup>a</sup>	BCS <sup>a</sup>
1/100	0.03196	0.00509	0.00518	0.53802	0.00259	0.00274
1/500	0.00639	0.00102	0.00104	0.10760	0.00052	0.00055
1/1000	0.00320	0.00051	0.00052	0.05380	0.00026	0.00027
1/5000	0.00064	0.00010	0.00010	0.01076	0.00005	0.00005
1/10,000	0.00032	0.00005	0.00005	0.00538	0.00003	0.00003

TABLE 3 Priors of variance of a single marker ( $\sigma_m^2$ ) used in the BayesC and BayesGC methods under the different priors for the fraction of total genetic variance explained by a single-fitted SNP (Fr) where  $\sigma_m^2 = \frac{Fr \cdot \sigma_{HET}^2}{HET}$  for each trait

<sup>a</sup>Total number born (TNB), number of stillborn piglets (STB), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (M3W), total litter weight at 3 weeks (LW3W), sow shoulder lesions (SHL) and the sow's body condition score (BCS).

**TABLE 4** Priors for variance attributed to the polygenic effect for the different traits for the different BayesGC methods

	Trait	BayesGC_10 <sup>a</sup>	BayesGC_50 <sup>b</sup>	BayesGC_90 <sup>c</sup>
$\sigma_{pol}^2$	TNB	0.944	0.525	0.105
	STB	0.150	0.084	0.017
	M3W	0.153	0.085	0.017
	LW3W	15.89	8.830	1.766
	SHL	0.077	0.043	0.009
	BCS	0.081	0.045	0.009

<sup>a</sup>BayesGC\_10 is BayesGC with 10% marker variance and 90% polygenic variance.

<sup>b</sup>BayesGC\_50 has 50% marker variance and 50% polygenic variance.

<sup>c</sup>BayesGC\_90 has 90% marker variance and 10% polygenic variance.

**TABLE 5** The estimated total genetic variance ( $\sigma_u^2$ ), residual variance ( $\sigma_e^2$ ) and heritabilities ( $h^2$ ) for the six maternal traits

Trait <sup>a</sup>	$\sigma_u^2$	$\sigma_e^2$	$h^2$
TNB	1.049	4.490	0.189
STB	0.167	1.125	0.130
M3W	0.170	1.791	0.087
LW3W	17.66	39.99	0.306
SHL	0.085	0.163	0.343
BCS	0.090	0.203	0.307

<sup>a</sup>total number born (TNB), number of stillborn piglets (STB), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (M3W), total litter weight at 3 weeks (LW3W), sow shoulder lesions (SHL) and the sow's body condition score (BCS).

(1.4% higher than GBLUP) when fitting either BayesC (Fr 1/5000) or BayesGC\_50 (Fr 1/5000). The trait LW3W had a 0.7% higher accuracy for BayesGC\_90 (Fr 1/10,000) than GBLUP. The traits TNB and STB showed no benefit of fitting Bayesian variable selection methods compared with GBLUP.

A limited increase in accuracy when going from GBLUP to BayesGC could be because the accuracy of prediction for the trait using GBLUP already is quite high. Our reference population was quite large (9–15,000 animals). A reference population of 7–11,000 animals was sufficient to obtain GEBV prediction accuracies comparable to the EBVs obtained with progeny testing for Japanese Black cattle (Takeda et al., 2021). TNB and LW3W with an accuracy of ~0.6 and ~0.7 for GBLUP, respectively, might not have as much potential for increasing their accuracy as M3W, with a much lower general accuracy of prediction (~0.44 for GBLUP). However, the trait STB showing the least benefit of fitting a Bayesian model also has the lowest general prediction accuracy of ~0.3. This could mean that there are other factors impacting the possible prediction accuracy of STB. For example, there could be fewer or no major QTL for the trait STB, lower linkage disequilibrium between markers and QTL, or low minor allele frequency of QTL for STB.

## 4.2 | Genetic architecture

The accuracy of GP depends on the proportion of genetic variance captured by the markers, the size of the reference population, the additive genetic relationship between the animals in the reference and the validation population, the heritability of the trait, the number of independent QTL and the effective number of chromosome segments (Daetwyler et al., 2010, 2013; Habier et al., 2007, 2010; Wientjes et al., 2013). Most individuals have records for all traits in our current data, which implies that the genetic relationships between the reference and validation populations are approximately the same over the six traits. However, some individuals have missing records for some traits, resulting in reduced reference population size. M3W and LW3W have ~10 K reference animals, while the other traits have ~15 K reference populations. LW3W, SHL and BCS have the highest heritability, implying more informative reference data (Table 1). Thus, it seems that the main differences between the traits in our study are the genetic architectures of the traits, i.e. how much genetic variance is captured by the markers and the size and number of major QTLs present for each trait.

While the traits are all considered to be complex and polygenic, some of the traits might have major genes and SNPs in close linkage disequilibrium that explain a substantial part of the genetic variance. However, if there happen to be many SNPs with substantial LD to a major gene, e.g. due to high genetic drift in the region, the GBLUP method may still perform well, since it can use many SNPs to explain the major gene effect. Also, for some traits, genomic predictions may have been over larger genetic distances, i.e. reduced relationships between reference and validation animals, which favours variable selection genomic prediction methods since they focus on SNPs that are in close LD with the QTL (Meuwissen & Goddard, 2010; Solberg et al., 2009).

The QTL database (Pig QTLdb; Hu et al., 2022) for each trait shows that there were 228 detected QTL for “Total number born alive” (TNB) and 138 QTL for “Number

Trait <sup>a</sup>	Method	Fr	Accuracy	SE	b
TNB	BayesGC_10 <sup>b</sup>	1/10,000	0.610	0.07	0.97
	BayesC	1/10,000	0.607	0.07	0.95
	GBLUP	—	0.611	0.07	0.98
STB	BayesGC_10 <sup>b</sup>	1/5000	0.318	0.08	0.50
	BayesC	1/5000	0.311	0.08	0.49
	GBLUP	—	0.318	0.08	0.50
M3W	BayesGC_50 <sup>c</sup>	1/100	0.484	0.10	0.79
	BayesC	1/500	0.464	0.10	0.74
	GBLUP	—	0.441	0.10	0.74
LW3W	BayesGC_90 <sup>d</sup>	1/10,000	0.722	0.05	1.17
	BayesC	1/10,000	0.718	0.05	1.04
	GBLUP	—	0.717	0.05	1.04
SHL	BayesGC_50 <sup>c</sup>	1/100	0.418	0.05	0.44
	BayesC	1/5000	0.409	0.05	0.41
	GBLUP	—	0.406	0.05	0.41
BCS	BayesGC_50 <sup>c</sup>	1/5000	0.518	0.05	0.70
	BayesC	1/5000	0.518	0.05	0.70
	GBLUP	—	0.511	0.05	0.70

<sup>a</sup>Total number born (TNB), number of stillborn piglets (STB), piglet mortality within 3 weeks, i.e. number of piglets dead after birth and until 3 weeks (M3W), total litter weight at 3 weeks (LW3W), sow shoulder lesions (SHL) and the sow's body condition score (BCS).

<sup>b</sup>BayesGC\_10 is BayesGC with 10% marker variance and 90% polygenic variance.

<sup>c</sup>BayesGC\_50 has 50% marker variance and 50% polygenic variance.

<sup>d</sup>BayesGC\_90 has 90% marker variance and 10% polygenic variance.

of stillborn" (STB). The trait "Piglet mortality within 3 weeks" (M3W) did not exist in the database. However, 10 QTL were found for the trait "Piglet Mortality". There was also no trait in the database defined as "Total litter weight within 3 weeks" (LW3W), but 1 QTL was listed for "Total litter weight at weaning" (He et al., 2021). There was also no QTL listed for Body Condition Score (BCS) or Shoulder Lesions (SHL). However, the published QTL listed in the database does not only reflect the genetic architecture of the traits but serve also as indicators for which traits that are more or less investigated.

QTL markers identified by GWAS on sequence data may be included in 50k marker panels for genomic prediction. In Holstein cattle (Brøndum et al., 2015), this method showed increased reliability of genomic prediction, especially when the QTL is included as a separate variance component, as it allows for extra emphasis on the QTL.

If large QTL included in the prediction model can help increase the prediction accuracy, why not just include the QTL directly in the linear model? This, however, requires a two-step approach, where one first finds the QTL associated with the trait and then includes them in the genomic prediction model. The BayesGC method fits both the polygenic trait and the important SNPs in one analysis. Both

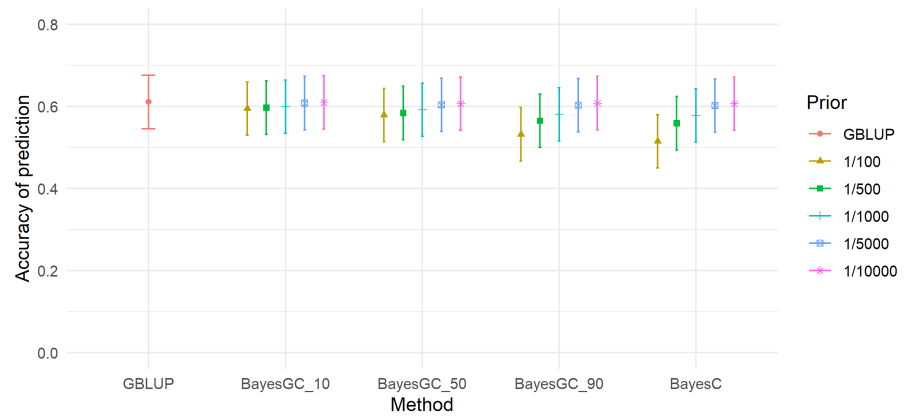
**TABLE 6** Accuracy, standard error (SE) and regression coefficients (b) for each trait from the method and fraction of the total genetic variance explained by a single-fitted SNP (Fr) yielding the highest accuracy for each trait

approaches do, however, show that there is room for improvement in prediction accuracy by including important SNPs with higher emphasis in a genomic prediction model. Bayesian variable selection methods also have the potential to find the functional SNPs to include in a linear model (Meuwissen et al., 2021; van den Berg et al., 2013).

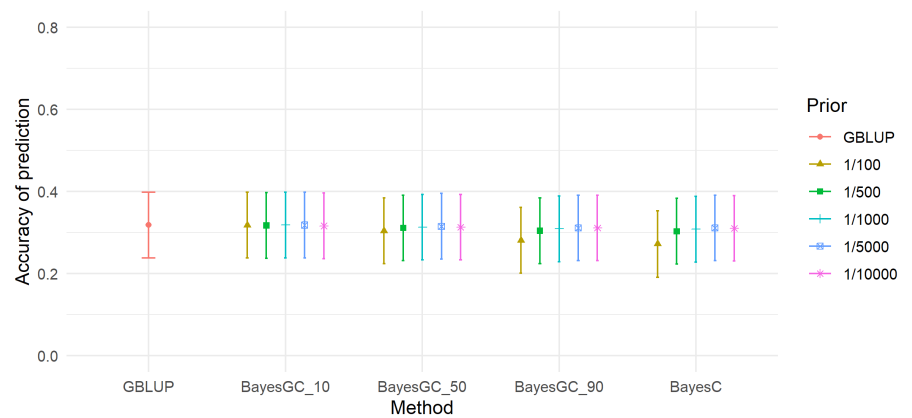
### 4.3 | Prior distributions

Bayesian variable selection methods use priors, which need to be carefully chosen or hyper-parameters of the priors estimated as part of the prediction method. The latter would extend the number of MCMC cycles substantially, as these hyper-parameters converge much slower to their equilibrium distribution than GEBVs. In our study we tried a range of different priors, varying both the number of SNPs to be included in the model through Fr, the emphasis of each SNP through the variance explained by markers ( $\sigma_m^2$ ) and the ratio between variance explained by markers and variance explained by the polygenic effect ( $\sigma_{pol}^2$ ), where with BayesGC\_10, 10% of the total genetic variance is fitted with markers ( $\sigma_m^2$ ) and 90% with the polygenic effect ( $\sigma_{pol}^2$ ), BayesGC\_50 the variance is split 50/50 and with

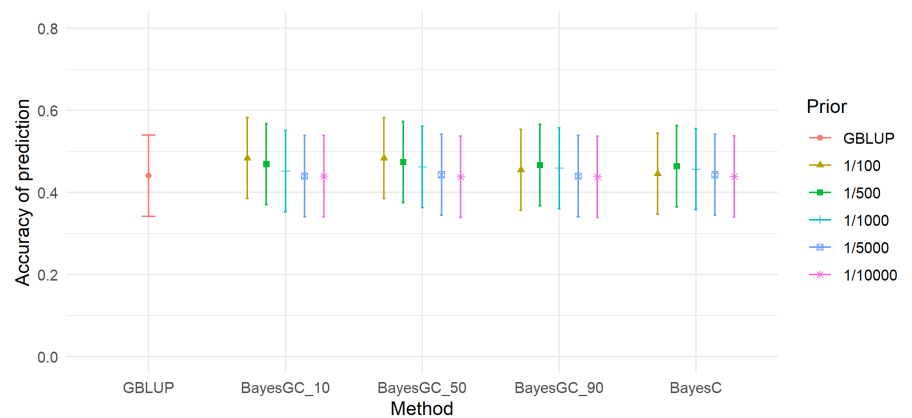
**FIGURE 1** The accuracy of prediction for the trait total number born (TNB), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



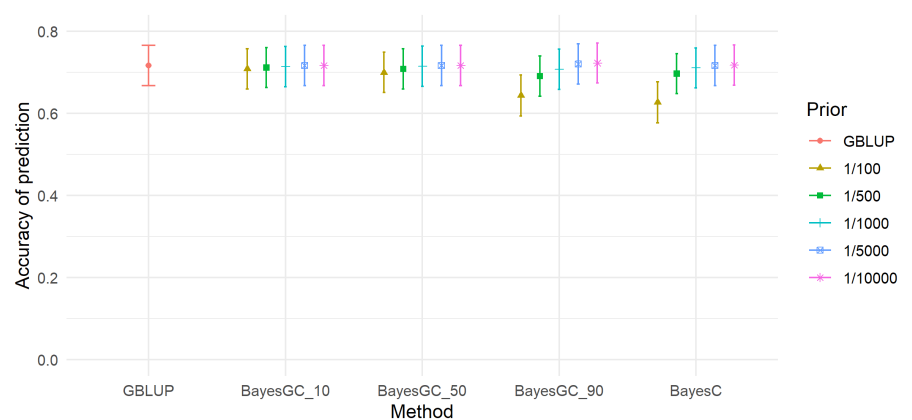
**FIGURE 2** The accuracy of prediction for the trait number of stillborn (STB), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



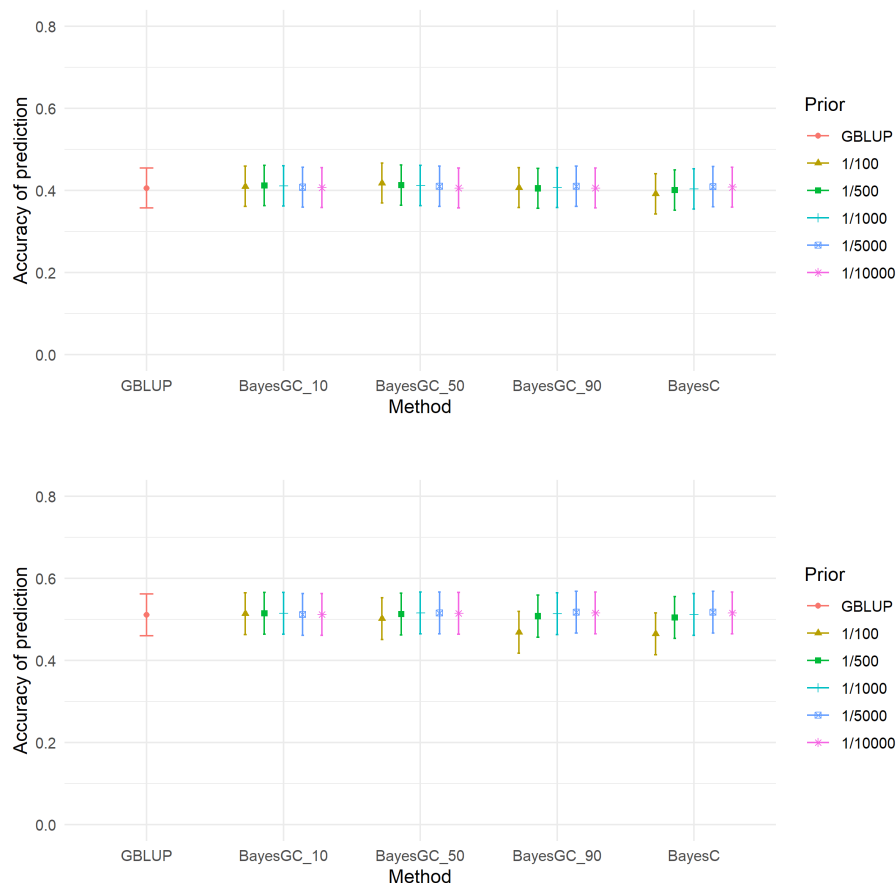
**FIGURE 3** The accuracy of prediction for the trait mortality within 3 weeks (M3W), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** The accuracy of prediction for the trait total litter weight at 3 weeks (LW3W), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]







**FIGURE 5** The accuracy of prediction for the trait shoulder lesions (SHL), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**FIGURE 6** The accuracy of prediction for the trait body condition score (BCS), from the different prediction methods at the different priors for fraction of variance explained by a single SNP (bars denote standard errors) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

BayesGC\_90, 90% of the total genetic variance is fitted with markers and 10% with the polygenic effect. For the Bayesian methods, the priors on the fraction of variance explained by a single SNP ( $Fr$ ) seem more important than how much genetic variance is explained by either polygenic effect ( $\sigma_{pol}^2$ ) or the marker effects ( $\sigma_m^2$ ), i.e. there are more differences within the methods BayesGC\_10, BayesGC\_50 or BayesGC\_90 than between them.

M3W (Figure 3) showed the largest differences in accuracy between GBLUP and BayesGC and it seems the accuracy increases gradually as  $Fr$  becomes larger (fewer SNPs fitted) but only when the ratio between  $\sigma_m^2$  and  $\sigma_{pol}^2$  is favouring  $\sigma_{pol}^2$  in such a way that the model fits 50%–90% of the genetic variance as  $\sigma_{pol}^2$  and the remaining variance is fitted with very few SNPs that in turn get fitted with a relatively high emphasis through  $Fr$ . The traits SHL and BCS (Figures 5 and 6) show a similar pattern, i.e. fitting a few SNPs is not improving the overall prediction accuracy unless it is also accompanied by a high emphasis on  $\sigma_{pol}^2$ . This could indicate that finding QTL and fitting them on their own is not sufficient to obtain high prediction accuracy. One also needs the support of a polygenic effect through, e.g. a genomic relationship matrix. However, when fitting many SNPs through the BayesC term, the Bayesian variable selection method would also fit many SNPs with a small effect—similar

to GBLUP. The benefit of a Bayesian variable selection method compared with GBLUP is thus expected to be lower for the methods with a higher  $\pi$ -value, like the  $Fr$  1/5000 and 1/10,000.

#### 4.4 | Further developments

Further development of the BayesGC would be to expand the model to include nongenotyped animals in the estimation of breeding values through, e.g. single-step genomic prediction (Christensen & Lund, 2010; Fernando et al., 2014; Legarra et al., 2009). The challenge of including nongenotyped animals with genomic prediction is the need to impute genotypes. With linear methods, there are methods such as ssGBLUP (Aguilar et al., 2010; Christensen & Lund, 2010; Legarra et al., 2009; Misztal et al., 2009) where an additive relationship matrix  $H$  is combining information from both pedigree and SNP data. Bayesian methods for combining genotyped and nongenotyped animals have been developed that could be adapted to this model by imputing genotypes for nongenotyped animals using MCMC methods that could be used with whole-genome data (Fernando et al., 2014). BayesGC includes a polygenic effect in the form of a  $G$ -matrix, which could also be exchanged with an  $H$ -matrix

to include nongenotyped animals, and the marker-model-based single-step approach of Fernando et al. (2014) could be used for the additional SNPs fitted by the BayesGC model. Other options for using BayesGC results in routine genomic evaluations would be to use the analysis of genotyped animals to find SNPs that need extra weight. In a regular GBLUP/ssGBLUP analysis, these SNPs would thus attain extra weights when constructing **G** and implicitly the **H**-matrix. The information on SNP variance from a Bayesian analysis could thus be used to improve the genomic relationship matrix for GBLUP or ssGBLUP analyses.

Another way to improve BayesGC could be to expand the software towards multi-trait analyses as many routine breeding evaluations today are based on multi-trait models. Expanding the BayesGC model towards multi-trait analyses is relatively straightforward if one assumes that an SNP with a large effect, is affecting all the (related) traits (Karaman et al., 2019; Kemper et al., 2018). In situations where one cannot assume this, multi-trait variable selection modelling requires to sample which combination of traits is affected by each of the SNPs. If there are many traits, there are many such combinations. Applying the BayesGC results to multi-trait routine evaluations may be by giving extra weight to some SNP genotypes, resulting in a different **G**-matrix for each of the traits, and consequently also for different pairs of traits (since the **G**-matrix modelling covariances between traits *i* and *j* is constructed as the cross-product of the SNP genotypes weighted for trait *i* and those weighted for trait *j*). Modifications of routine software packages may be needed to accommodate these per trait alternative **G**-matrices.

Bayesian variable selection methods have a lot of potential for further development to be used in routine breeding value estimations. One of the biggest drawbacks today is the high computational costs of running the MCMC chains. However, computational power has historically increased and will most likely continue to increase, in addition to further research developing into more efficient algorithms using parallel computations.

## 5 | CONCLUSIONS

The accuracy of genomic prediction on six maternal traits in landrace pigs varied greatly ranging from 0.31 to 0.61. The prediction accuracies did not vary much between the different genomic prediction methods. The two traits M3W and BCS could benefit from using a BayesGC approach with a 9.8 and 3.0% increase in accuracy, respectively, while TNB, STB, LW3W and SHL showed only minor improvements. Although GBLUP,

BayesC and BayesGC all yielded similar genomic prediction accuracies, the accuracy of BayesGC was always as high as or higher than that of GBLUP. Within the BayesGC method, the accuracies could vary depending on the prior distributions. The models were more sensitive to how many markers were fitted in the model by varying the fraction of the total genetic variance explained by a single marker (*Fr*) compared with the amount of total genetic variance explained by marker effects as a whole (BayesGC\_10, BayesGC\_50 or BayesGC\_90), but overall, most traits were robust against varying the prior distributions.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the funding from the Norwegian Research Council (project 255297). Topigs Norsvin SA is acknowledged for providing data and genotype information. The helpful comments of two anonymous reviewers are also gratefully acknowledged.

## CONFLICT OF INTEREST

The authors confirm that there is no known conflict of interest associated with this publication.

## DATA AVAILABILITY STATEMENT

Restrictions apply to the availability of these data, which were used under licensed for this study. Data might be available upon reasonable request from the authors with the permission of Topigs Norsvin.

## ORCID

Maria V. Kjetså  <https://orcid.org/0000-0002-8018-9872>

Arne B. Gjuvsland  <https://orcid.org/0000-0002-4391-3411>

Øyvind Nordbø  <https://orcid.org/0000-0002-7000-1055>

## REFERENCES

- Aguilar, I., Misztal, I., Johnson, D., Legarra, A., Tsuruta, S., & Lawlor, T. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2), 743–752. <https://doi.org/10.3168/JDS.2009-2730>
- Brøndum, R. F., Su, G., Janss, L., Sahana, G., Guldbrandtsen, B., Boichard, D., & Lund, M. S. (2015). Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *Journal of Dairy Science*, 98(6), 4107–4116. <https://doi.org/10.3168/jds.2014-9005>
- Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, 42(1), 1–8. <https://doi.org/10.1186/1297-9686-42-2>
- Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de los Campos, G., & Hickey, J. M. (2013). Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics*, 193(2), 347–365. <https://doi.org/10.1534/genetics.112.147983>

- Daetwyler, H. D., Pong-Wong, R., Villanueva, B., & Woolliams, J. A. (2010). The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, *185*(3), 1021–1031. <https://doi.org/10.1534/genetics.110.116855>
- Erbe, M., Hayes, B. J., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., ... Goddard, M. E. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, *95*(7), 4114–4129. <https://doi.org/10.3168/jds.2011-5019>
- Eriksen, E. M. (2018). *Effects of changes in the breeding goal on genetic improvement for maternal traits in Landrace pigs*. Norwegian University of Life Science [NMBU]. Retrieved from <http://hdl.handle.net/11250/2572879>
- Fernando, R. L., Dekkers, J. C., & Garrick, D. J. (2014). A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics, Selection, Evolution: GSE*, *46*(1), 50. <https://doi.org/10.1186/1297-9686-46-50>
- Gianola, D., De Los Campos, G., Hill, W. G., Manfredi, E., & Fernando, R. (2009). Additive genetic variability and the Bayesian alphabet. *Genetics*, *183*(1), 347–363. <https://doi.org/10.1534/genetics.109.103952>
- Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *177*(4), 2389–2397. <https://doi.org/10.1534/genetics.107.081190>
- Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*, *12*(1), 186. <https://doi.org/10.1186/1471-2105-12-186>
- Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P., & Thaller, G. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genetics Selection Evolution*, *42*(1), 5. <https://doi.org/10.1186/1297-9686-42-5>
- He, Y., Zhou, X., Zheng, R., Jiang, Y., Yao, Z., Wang, X., ... Yuan, X. (2021). The association of an SNP in the EXOC4 gene and reproductive traits suggests its use as a breeding marker in pigs. *Animals: An Open Access Journal from MDPI*, *11*(2), 521. <https://doi.org/10.3390/ani11020521>
- Hu, Z.-L., Park, C. A., & Reecy, J. M. (2022). Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Research*, *50*(D1), D956–D961. <https://doi.org/10.1093/nar/gkab1116>
- Karaman, E., Lund, M. S., & Su, G. (2019). Multi-trait single-step genomic prediction accounting for heterogeneous (co)variances over the genome. *Heredity*, *124*(2), 274–287. <https://doi.org/10.1038/s41437-019-0273-4>
- Kemper, K. E., Bowman, P. J., Hayes, B. J., Visscher, P. M., & Goddard, M. E. (2018). A multi-trait Bayesian method for mapping QTL and genomic prediction. *Genetics Selection Evolution*, *50*(1), 1–13. <https://doi.org/10.1186/s12711-018-0377-y>
- Kjetså, M. H., Ødegård, J., & Meuwissen, T. H. E. (2020). Accuracy of genomic prediction of host resistance to salmon lice in Atlantic salmon (*Salmo salar*) using imputed high-density genotypes. *Aquaculture*, *526*, 735415. <https://doi.org/10.1016/j.aquaculture.2020.735415>
- Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, *92*(9), 4656–4663. <https://doi.org/10.3168/JDS.2009-2061>
- Lillehammer, M., Meuwissen, T. H. E., & Sonesson, A. K. (2011). Genomic selection for maternal traits in pigs. *Journal of Animal Science*, *89*(12), 3908–3916. <https://doi.org/10.2527/jas.2011-4044>
- Lillehammer, M., Meuwissen, T. H. E., & Sonesson, A. K. (2013). Genomic selection for two traits in a maternal pig breeding scheme. *Journal of Animal Science*, *91*(7), 3079–3087. <https://doi.org/10.2527/jas.2012-5113>
- Madsen, P., & Jensen, J. (2013). A user's guide to DMU a package for analysing multivariate mixed models. Retrieved from <http://dmu.agrsci.dk>
- Meuwissen, T., & Goddard, M. (2010). Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics*, *185*(2), 623–631. <https://doi.org/10.1534/genetics.110.116590>
- Meuwissen, T., van den Berg, I., & Goddard, M. (2021). On the use of whole-genome sequence data for across-breed genomic prediction and fine-scale mapping of QTL. *Genetics Selection Evolution*, *53*(1), 19. <https://doi.org/10.1186/s12711-021-00607-4>
- Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. Retrieved from <http://www.genetics.org/content/157/4/1819.short>
- Misztal, I., Legarra, A., & Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*, *92*(9), 4648–4655. <https://doi.org/10.3168/JDS.2009-2064>
- Ocepek, M., & Andersen, I. L. (2017). What makes a good mother? Maternal behavioural traits important for piglet survival. *Applied Animal Behaviour Science*, *193*, 29–36. <https://doi.org/10.1016/j.applanim.2017.03.010>
- Salek Ardestani, S., Jafarikia, M., Sargolzaei, M., Sullivan, B., & Miar, Y. (2021). Genomic prediction of average daily gain, back-fat thickness, and loin muscle depth using different genomic tools in canadian swine populations. *Frontiers in Genetics*, *12*, 735. Retrieved from <https://www.frontiersin.org/article/10.3389/fgene.2021.665344>
- Samorè, A. B., & Fontanesi, L. (2016). Genomic selection in pigs: State of the art and perspectives. *Italian Journal of Animal Science*, *15*(2), 211–232. <https://doi.org/10.1080/1828051X.2016.1172034>
- Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, *15*, 478. <https://doi.org/10.1186/1471-2164-15-478>
- Solberg, T. R., Sonesson, A. K., Woolliams, J. A., Ødegård, J., & Meuwissen, T. H. (2009). Persistence of accuracy of genome-wide breeding values over generations when including a polygenic effect. *Genetics Selection Evolution*, *41*(1). <https://doi.org/10.1186/1297-9686-41-53>
- Song, H., Zhang, J., Jiang, Y., Gao, H., Tang, S., Mi, S., Yu, F., Meng, Q., Xiao, W., Zhang, Q., & Ding, X. (2017). Genomic prediction for growth and reproduction traits in pig using an admixed reference population1. *Journal of Animal Science*, *95*(8), 3415–3424. <https://doi.org/10.2527/jas.2017.1656>
- Strandén, I., & Garrick, D. J. (2009). Derivation of equivalent computing algorithms for genomic predictions and reliabilities

- of animal merit. *Journal of Dairy Science*, 92(6), 2971–2975. <https://doi.org/10.3168/jds.2008-1929>
- Takeda, M., Inoue, K., Oyama, H., Uchiyama, K., Yoshinari, K., Sasago, N., ... Uemoto, Y. (2021). Exploring the size of reference population for expected accuracy of genomic prediction using simulated and real data in Japanese Black cattle. *BMC Genomics*, 22(1), 1–11. <https://doi.org/10.1186/S12864-021-08121-Z/FIGURES/3>
- Tan, C., Wu, Z., Ren, J., Huang, Z., Liu, D., He, X., & Prakapenka, D. (2017). Genome - wide association study and accuracy of genomic prediction for teat number in Duroc pigs using genotyping - by - sequencing. *Genetics Selection Evolution*, 49(1), 35. <https://doi.org/10.1186/s12711-017-0311-8>
- van den Berg, I., Fritz, S., & Boichard, D. (2013). QTL fine mapping with Bayes C( $\pi$ ): a simulation study. *Genetics, Selection, Evolution: GSE*, 45(1), 19. <https://doi.org/10.1186/1297-9686-45-19>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- VanRaden, P. M., & Wiggans, G. R. (1991). Derivation, calculation, and use of national animal model information. *Journal of Dairy Science*, 74(8), 2737–2746. [https://doi.org/10.3168/jds.S0022-0302\(91\)78453-1](https://doi.org/10.3168/jds.S0022-0302(91)78453-1)
- VanRaden, P. M. (2007). Genomic Measures of Relationship and Inbreeding. *Interbull Bulletin*, 25(37), 111–114. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Verbyla, K. L., Bowman, P. J., Hayes, B. J., & Goddard, M. E. (2010). Sensitivity of genomic selection to using different prior distributions. *BMC Proceedings*, 4(S1), 2–5. <https://doi.org/10.1186/1753-6561-4-s1-s5>
- Wientjes, Y. C. J., Veerkamp, R. F., & Calus, M. P. L. (2013). The effect of linkage disequilibrium and family relationships on the reliability of genomic prediction. *Genetics*, 193(2), 621–631. <https://doi.org/10.1534/genetics.112.146290>

**How to cite this article:** Kjetså, M. V., Gjuvsland, A. B., Nordbø, Ø., Grindflek, E., & Meuwissen, T. (2022). Accuracy of genomic prediction of maternal traits in pigs using Bayesian variable selection methods. *Journal of Animal Breeding and Genetics*, 139, 654–665. <https://doi.org/10.1111/jbg.12729>