Multi-trait genomic prediction in pigs using single and multistep methods based on the absorption of ungenotyped animals

Tu Luan[1,*] tu.luan@nmbu.no;

Øyvind Nordbø[2] ;

Ina Andersen-Ranberg[2] ;

Theo H. E. Meuwissen[1]

[1]Faculty of Biosciences
Norwegian University of Life Sciences
Ås N-1432, Norway

[2]Norsvin SA
Storhamargata 44, Hamar 2317, Norway

[*]**Correspondence**
Tu Luan, Faculty of Biosciences, Norwegian University of Life Sciences, Ås N-1432, Norway.
~~ing author~~Email: tu.luan@nmbu.no

Abstract

Many quantitative traits measured in breeding programs are genetically correlated. The genetic correlations between the traits indicate that the measurement of one trait ~~carry~~ carries information on others. To benefit from this information, multi-trait genomic prediction (MTGP) is preferable to use. However, MTGP is more difficult to implement compared to single-trait genomic prediction (STGP), and even more challenging for the goal to exploit not only the information on other traits but also the information on ungenotyped animals. This could be accomplished ~~by~~ using both single and multistep methods. The single-step method was achieved by implementing a single-step genomic best linear unbiased prediction~~BLUP~~ (ssGBLUP) approach using a multi-trait model. Here, we examined a multistep analysis based on an approach called “Absorption” to achieve this goal. The Absorption approach absorbed

all available information including the phenotypic information on ungenotyped animals as well as the information on other traits if applicable, into mixed model equations of genotyped animals. The multistep analysis included (1) to apply the Absorption approach that exploits all available information~~,~~ and (2) to implement genomic BLUP (GBLUP) prediction on the absorbed dataset. In this study, the ssGBLUP and multistep analysis were applied to 5 traits in Duroc pigs, which were slaughter percentage ~~(SP)~~, feed consumption from 40 ~~kg~~ to 120 kg (FC40_120), days of growth from 40 ~~kg~~ to 120 kg (D40_120), age at 40 kg (A40) and lean meat percentage ~~(LMP)~~. The results showed that MTGP yielded ~~a~~ higher accuracy than STGP, which on average was 0.057 higher for the multistep method and 0.045 higher for ssGBLUP. The m~~M~~ultistep method achieved similar prediction accuracy as ssGBLUP. However, the prediction bias of the multistep method was in general lower than that of ssGBLUP.

Keywords

absorption of phenotype
genomic selection
m~~M~~ulti-trait genomic prediction
pig

1. INTRODUCTION

With the availability of high-density panels of DNA markers covering the whole genome, genomic selection (GS) (Meuwissen et al., 2001) has become feasible as an effective tool for animal and plant breeding. This method has been successfully implemented in livestock breeding programs, and most extensively in dairy cattle (Wiggans & Carrillo, 2022). Selection of elite bulls and cows based on the genomic estimated breeding value (GEBV) doubles genetic gains mainly due to a reduction of the generation intervals (Garcia-Ruiz et al., 2016). This also reduced the cost of proving bulls by more than 90% (Schaeffer, 2006). GS is also a promising procedure to increase genetic gain in ~~the~~ pig breeding, especially for the traits that are not easy to measure on selection candidates and/or have low heritability, such as meat quality (Lopez et al., 2020), and also on traits obtained late in pig's life (Mote et al., 2019).

Actual breeding often targets multiple traits that are genetically correlated, and the practical routine genetic evaluation of the breeding value is usually calculated using multi-trait models. Multi-trait models for GEBV prediction have been reported including Bayesian approaches (Villar-Hernandez et al., 2021) and the genomic best linear unbiased prediction (GBLUP) method (Karaman et al., 2020). Studies have shown that multi-trait genomic prediction (MTGP), which accounts for the relationships between the traits, may result in more accurate GEBV than single-trait genomic prediction (STGP) (Semagn et al., 2022; Song et al., 2020).

MTGP could be implemented by a multiple-~~-~~step procedure. This includes to run traditional multi-trait genetic evaluation for each individual; to create pseudo-records by adjusting the phenotypes; multi-trait estimation of allelic effects for each SNP~~;~~ and to combine genomic predictions and traditional evaluations in a selection index (VanRaden, 2008). Although genomic evaluations are more accurate than the parent average, using approximations for adjusting phenotype can inflate GEBV and hence cause bias (VanRaden et al., 2009).

The approach referred to as single-step GBLUP (ssGBLUP) combines the pedigree relationship matrix (**A** matrix) and genomic relationship matrix (**G** matrix) into a single relationship matrix called **H** matrix (Legarra et al., 2009; Misztal et al., 2009). The inverse of the **H** matrix has a simple form and can substitute for the inverse of the traditional relationship matrix. Compared to the multistep method, this approach makes use of all data (pedigree, genotypes and phenotypes) simultaneously to maximize~~maximise~~ the accuracy of the GEBV.

With ~~the~~ extensive research and the development of efficient computing algorithms to solve the challenges that limit the practical implementation of ssGBLUP, for instance metafounder approach (Kudinov et al., 2020; Legarra et al., 2015) and the use of the J factor (Belay et al., 2022; Stranden et al., 2022) to improve the compatibility between genomic and pedigree information, ssGBLUP has become the most popular methodology for genetic evaluations including genotyped and ungenotyped animals and has been successfully implemented in almost all livestock populations (Bermann et al., 2022). However, even though ssGBLUP took over the multistep method as the

most chosen genetic evaluation methodology, the multistep method may show merits for instance the more straightforward extension of the method to variable selection models.

Here, we examine a multistep method to achieve MTGP based on an approach called "Absorption" that approximately absorbs all available information on ungenotyped animals as well as the information on other traits, into mixed model equations of genotyped animals. This Absorption approach creates pseudo-records referred to as absorbed records. MTGP was achieved by performing genomic prediction with absorbed records. The procedure involves (1) regular genetic evaluation to predict breeding value (EBV) for each individual and the reliability of EBV prediction is calculated; (2) creation of pseudo-observations and weights by absorbing the phenotypic information of ungenotyped animals into mixed model equations for genotyped animals and~~;~~ (3) GEBV prediction using pseudo-records and variable or non-variable selection methods.

In this study, the accuracy and bias of MTGP ~~was~~ were investigated for this multistep approach using absorbed records and compared with ssGBLUP. MTGP were conducted on 5 traits of Duroc boars. STGP were also performed to compare with MTGP. The accuracy of GEBV prediction was assessed by 1~~-~~028 validation boars.

2. MATERIALS AND METHODS

2.1. Genotypic and phenotypic data

The phenotypic data of 9~~-~~641 Duroc pigs were provided by Norsvin SA (www.norsvin.no). There were 5 traits used in the study: slaughter percentage (SP), feed consumption from 40~~ kg~~ to 120 kg (FC40_120), days of growth from 40~~ kg~~ to 120 kg (D40_120), age at 40 kg (A40) and lean meat percentage (LMP). A description of the phenotype for these 5 traits ~~are~~ is shown in Table 1. All data were obtained through operational breeding procedures in Norsvin and all animals in the study were reared according to the laws and regulations for keeping pigs in Norway (Animal Welfare Act 2009-06-19-97, Regulation for the keeping of pigs in Norway 2003-02-18-175).

Within the Duroc pigs in the dataset, 5~~-~~045 boars born between 2010 and 2015 were

genotyped at Cigene (http://www.cigene.no/), using the iScan (Illumina, San Diego, CA, USA) platform with the PorcineSNP60 array according to manufacturer's instructions. Image intensity data processing, clustering and genotype calling ~~was~~ were performed using the genotyping module in the Genome Studio software (Illumina, San Diego, CA, USA). A total of 36,551 single nucleotide polymorphisms (SNPs) remained after removing SNP with minor allele frequency (MAF) below 0.01.

2.2. Genotyped animals, their ancestors and ungenotyped animals

The animals used in the study are defined as two types: Genotyped animals and their ancestors (GA-set) and ungenotyped animals (D-set) that are generally descendants of the GA-set. GA-set comprised of 9750 animals, of which 9 generations of ancestors preceded 5045 genotyped animals (G-set). Included in the GA-set are 195 founders and 9555 non-founders. For 4705 ungenotyped ancestors in GA-set, their genotype probabilities were calculated using the LDMIP program (Meuwissen & Goddard, 2010). Ungenotyped D-set animals are generally descendants of GA-set animals. There are in total 127,825 ungenotyped D-set animals whose information will be absorbed into the genotyped animals.

2.3. Absorption of phenotypic information of ungenotyped descendants to GA-set animals

To absorb information ~~of~~ on ungenotyped animals (D-set) into the genotyped animals (G-set) in GA-set, the EBV of the animals and their reliabilities were required to be known. These can be obtained from, for example, a large-scale (national) pedigree-based genetic evaluation. In the presented study, this pedigree-based genetic evaluation was implemented using the DMU package (Madsen & Jensen, 2013; Madsen et al., 2014). The (co)variance matrix was from Norsvin's routine genetic evaluation of EBV.

The mixed model equation (MME) for the GA-set animals absorbing D-set information, based on the numerator relationship matrix (**A**-matrix), may be expressed as:

$$(\boldsymbol{M} + \boldsymbol{A}^{-1}\lambda)\boldsymbol{EBV}_{GA} = \boldsymbol{d}$$

where $\boldsymbol{M}$ is an information matrix resulting from the absorption; $\boldsymbol{EBV}_{GA}$ is the vector of EBV of the GA-set animals; λ is the variance ratio as σ_e^2/σ_a^2, where σ_a^2 and σ_e^2 are additive genetic variance and error variance, respectively~~;~~ and $\boldsymbol{d}$ is the right-hand-side resulting from the absorption process. Here the exact form of $\boldsymbol{M}$ and $\boldsymbol{d}$ is not defined, since they depend on fixed and random effects in the model. However, we can assume the information matrix $\boldsymbol{M}$ to be approximated by a diagonal weight matrix $\boldsymbol{W}$ to achieve the same $\boldsymbol{EBV}_{GA}$ and reliabilities result as for the complete data set (GA + D set). In addition, the right-hand-side $\boldsymbol{d}$ of MME can be approximated as $\boldsymbol{d} = \boldsymbol{W}\boldsymbol{y}_a$, where $\boldsymbol{y}_a$ is the vector of absorbed records yielding the complete set $\boldsymbol{EBV}_{GA}$. Thus, the MME for the absorbed records $\boldsymbol{y}_a$ with weights $diag(\boldsymbol{W})$ may be written as:

$$(\boldsymbol{W} + \boldsymbol{A}^{-1}\lambda)\boldsymbol{EBV}_{GA} = \boldsymbol{W}\boldsymbol{y}_a$$

The weights $\boldsymbol{W}$ that approximately give the same reliabilities as for the complete data ($GA + D$ set), from (national) genetic evaluations, are calculated following the approach of ~~Ricard et al., (~~Ricard et al.~~,~~ (2012). The absorbed records $\boldsymbol{y}_a$ are calculated by multiplying ($\mathbf{I} + \boldsymbol{W}\boldsymbol{A}^{-1}\lambda$) with the known $\boldsymbol{EBV}_{GA}$, where $\mathbf{I}$ is the identity matrix and $\boldsymbol{EBV}_{GA}$ is the vector of EBVs of genotyped animals from large-scale genetic evaluation.

When considering genomic relationships for the GA-set animals, the absorbed genomic ~~mixed model equation (~~MME~~)~~ may be expressed as:

$$(\boldsymbol{W} + \boldsymbol{G}^{-1}\lambda)\boldsymbol{GEBV}_{GA} = \boldsymbol{W}\boldsymbol{y}_a$$

where $\boldsymbol{W}$ and $\boldsymbol{y}_a$ are the same as for the $\boldsymbol{A}$ matrix-based equations; $\boldsymbol{G}$ is the genomic relationship matrix~~;~~ and $\boldsymbol{GEBV}_{GA}$ is the vector of GEBV of the genotyped animals. The absorption of D-set animals is not affected by the known marker genotypes since the D-set animals have no marker information, nor have their descendants (Meuwissen & Goddard, 1999). The absorbed MME model was implemented ~~by~~ using the package ASReml (Gilmour et al., 2006).

The absorption relied on the EBV and reliability obtained by a large-scale genetic evaluation based on ***A*** matrix relationships. If there were multiple traits that were genetically correlated, the genetic evaluation could be implemented either one by one on each trait through a single-trait model, or simultaneously on all traits through a multi-trait model. For the absorption based on the EBV and reliability from single-trait genetic evaluation, it is referred to as single-trait absorption, and single-trait absorbed records are obtained. For the absorption based on multi-trait EBV and reliability, it is referred to as multi-trait absorption and multi-trait absorbed records are obtained.

2.4. Single-trait multistep genomic prediction

Single-trait multistep (ST-multistep) genomic prediction with absorbed records could be implemented using the GBLUP model expressed as:

$$\mathbf{y}_a = \mathbf{1}\mu + \mathbf{Za} + \mathbf{e}$$

where $\mathbf{y}_a$ is a vector of absorbed pseudo-phenotypes for a trait; μ is the overall mean; **Z** is a design matrix linking the animals to the absorbed records; **a** is a vector of additive genetic effects of the animals and **e** is the vector of random residuals. It is assumed that $\mathbf{a} \sim N(\mathbf{0}, \mathbf{G}\sigma_g^2)$ where **G** is the genomic relationship matrix and σ_g^2 is the genetic variance associated with **G**, and $\mathbf{e} \sim N(\mathbf{0}, \mathbf{W}^{-1}\sigma_e^2)$ where **W** is the diagonal weight matrix obtained from the absorption.

There are various methods for calculating the **G** matrix. Here, we used the **G** matrix referred to as the $\mathbf{G}_{\mathbf{LDLA}}$ matrix constructed by the method of ~~Meuwissen et al.~~ Meuwissen et al. (2015). $\mathbf{G}_{\mathbf{LDLA}}$ matrix was a relationship matrix that combined linkage disequilibrium (LD) and linkage analysis (LA) relationship information as~~:~~ $\mathbf{G}_{\mathbf{LDLA}} = \boldsymbol{\Delta}\hat{\mathbf{G}}\boldsymbol{\Delta} + \mathbf{D}\hat{\mathbf{A}}\mathbf{D}$, where $\hat{\mathbf{G}} = \mathbf{XX}'/N_m$, as N_m is the number of markers and **X** is a matrix of the standardized marker genotypes, $X_{ij} = (g_{ij} - 2p_j)/\sqrt{2p_j(1-p_j)}$, where g_{ij} is the genotype of animal i for SNP j, with $g_{ij} = 0$, 1 or 2 for genotypes "0 0," ~~;~~ "1 0" or "1 1", respectively, and p_j is the frequency of allele 1 of SNP j. Standardization is such that the mean and the variance of X_{ij} are 0 and 1, respectively (Iversen et al., 2017); $\boldsymbol{\Delta}$ is a diagonal matrix as $\Delta_{ii} = 1/\sqrt{G_{ii}}$ if $G_{ii} \geq 1$ or $\Delta_{ii} = 1$ if $G_{ii} < 1$; $\hat{\mathbf{A}}$ is a pedigree-based gametic

relationship matrix and~~;~~ **D** is a diagonal matrix as $D_{ii} = \sqrt{1 - G_{ii}}$ if $G_{ii} \leq 1$ or $D_{ii} = 0$ if $G_{ii} > 1$.

2.5. Multi-trait multistep genomic prediction

Multi-trait multistep (MT-multistep) genomic prediction must be implemented using multi-trait absorbed records. To perform the absorption of information of ungenotyped animals into genotyped animals for multiple traits, the first multi-trait pedigree-based genetic evaluation on a complete data set was executed using the MiX99 package (Stranden & Lidauer, 1999; Vuori et al., 2006) to predict EBVs of the traits, $\boldsymbol{EBV}$.

For trait i in the multi-trait genetic model, the phenotype $\boldsymbol{y}_i$ can be expressed as $\boldsymbol{y}_i = \boldsymbol{a}_i + \boldsymbol{e}_i$, where $\boldsymbol{a}_i$ and $\boldsymbol{e}_i$ are additive effect and residual for trait i, and $\mathrm{Var}(\boldsymbol{a}_i) = \mathbf{G}$ and $\mathrm{Var}(\boldsymbol{e}_i) = \mathbf{R}$. Canonical transformation can be applied to trait i so that the transformed trait $\mathbf{X}'\boldsymbol{y}_i$ can be independently evaluated with a single trait model, and $\mathrm{Var}(\mathbf{X}'\boldsymbol{a}_i) = \mathbf{L}$ and $\mathrm{Var}(\mathbf{X}'\boldsymbol{e}_i) = \mathbf{I}$. Genetic variance matrix **L** is diagonal and the residual variance matrix is an identity matrix **I**, as $\mathbf{L} = \mathbf{X}'\mathbf{G}\mathbf{X}$ and $\mathbf{I} = \mathbf{X}'\mathbf{R}\mathbf{X}$.

Canonical transformed EBVs, $\boldsymbol{EBV}^*$, are given by $\boldsymbol{EBV}^* = \mathbf{X}'\boldsymbol{EBV}$. The reliabilities of predicting transformed EBVs were calculated using ApaX in the MiX99 package. With transformed EBVs and reliabilities, ~~the~~ single-trait absorption was implemented to obtain absorbed records $\boldsymbol{y}_a^*$ and weight $\boldsymbol{W}^*$. Then the absorbed MME model was executed ~~by~~ using the package ASReml to predict transformed GEBV, $\boldsymbol{GEBV}^*$, as $(\boldsymbol{W}^* + \boldsymbol{G}^{-1}\lambda)\boldsymbol{GEBV}^* = \boldsymbol{W}^*\boldsymbol{y}_a^*$. The GEBV predicted using multi-trait absorbed records is obtained as $\boldsymbol{GEBV} = (\mathbf{X}')^{-1}\boldsymbol{GEBV}^*$.

2.6. Single-trait and multi-trait ssGBLUP

The single-trait ssGBLUP (ST-ssGBLUP) is defined as:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{Z}\boldsymbol{a} + \boldsymbol{e}$$

where **y** is a vector of phenotypes for the traits; $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the design matrices; $\boldsymbol{b}$ and $\boldsymbol{a}$ denote the fixed effects and the additive genetic effects, respectively~~,~~ and **e** is the random residual. It is assumed that $\boldsymbol{a} \sim N(\mathbf{0}, \mathbf{H}\sigma_a^2)$ where σ_a^2 is additive genetic

variance, **H** is the pedigree-genomic relationship matrix which combines SNP marker and pedigree information. A detailed description of how **H** is computed can be found in Aguilar et al. (2010).

The mixed model equations are:

$$\begin{pmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z + H^{-1}\lambda} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{a}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{pmatrix}$$

where $\lambda = \sigma_e^2/\sigma_a^2$. The pedigree and genomic relationship matrices (VanRaden, 2008) were used to build the combined relationship matrices (Aguilar et al., 2010; Christensen & Lund, 2010; Legarra et al., 2009). The ST-ssGBLUP was implemented using the DMU package (Madsen & Jensen, 2013) with the G-ADJUST option to adjust elements in the genomic relationship so that the average of diagonal elements and the average of off-diagonal elements equal the same average in the additive relationship for the genotyped animals (Christensen et al., 2012).

For the multi-trait ssGBLUP (MT-ssGBLUP), the solution to mix model equations can be expressed as:

$$\begin{pmatrix} \boldsymbol{X'R^{-1}X} & \boldsymbol{X'R^{-1}Z} \\ \boldsymbol{Z'R^{-1}X} & \boldsymbol{Z'R^{-1}Z + H^{-1} \otimes G_0} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{a}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X'R^{-1}y} \\ \boldsymbol{Z'R^{-1}y} \end{pmatrix}$$

where $\boldsymbol{R = I \otimes R_0}$, $\boldsymbol{R_0}$ is the residual covariance matrix across traits and $\boldsymbol{G_0}$ is the genetic covariance matrix across traits (Legarra et al., 2014). The implementation of MT-ssGBLUP was also achieved using the DMU package with the G-ADJUST option.

2.7. Validation procedure

A validation dataset was constructed comprising 1028 boars born after 1 February ~~1st,~~ 2014. For ssGBLUP analysis, the reference data set consisted of available records in the period from January ~~of~~ 2008 to January of 2014. All of the 4017 genotyped animals in the reference data had their own records. For the multistep method using absorbed GA-set, the reference data set consisted of 8722 absorbed records. For the multistep method using absorbed G-set, the reference data set consisted of 4017 absorbed records.

The accuracy of GEBV prediction was calculated as the correlation between the predicted GEBVs and the adjusted phenotypes, divided by the square root of the heritability of the trait (Estaghvirou et al., 2013). The adjusted phenotypes were calculated as the sum of EBVs and residuals from the traditional genetic evaluation (Wang et al., 2022).

The bias was measured as the coefficients of regression of the adjusted phenotypes on GEBV. For an unbiased result, the regression coefficient equals to 1. A regression coefficient ~~<~~ 1 implies that extremely high (low) values of the GEBV over-(under)predict the adjusted phenotypes, and vice versa for a regression coefficient ~~>~~ 1. The degree of bias is hence judged by comparing the regression coefficients of the adjusted phenotypes on GEBV with the value 1.

3. RESULTS

3.1. Accuracy of MTGP and STGP by the multistep method using absorbed GA-set records and by ssGBLUP

Table 2 presents the accuracies of ST-multistep and MT-multistep using single-trait and multi-trait absorbed records, and the accuracies of ST-ssGBLUP and MT-ssGBLUP analyses for comparison. The absorbed records used by multistep method were obtained by absorbing information into GA-set animals. The GEBVs were obtained for the 1028 validation animals when 8722 animals were in the training set.

The results in Table 2 show that the multistep method performed similarly to ssGBLUP. Over the 5 traits, neither method yielded a more accurate prediction than the other. MTGP achieved higher accuracy than STGP except for trait SP where accuracies were similar. There is on average a larger difference in accuracy between MTGP and STGP for the multistep method (0.057) than for ssGBLUP (0.045). It is observed a larger difference in accuracy between MTGP and STGP for the GEBV prediction with lower accuracy. The largest difference between MTGP and STGP was observed for trait A40, which ~~were~~ was 0.111 for the multistep method and 0.077 for ssGBLUP analysis. The trait was also observed with the lowest GEBV accuracy. For traits SP and FC40_120 with relatively low heritability, the genomic predictions were

similarly accurate to trait LMP whose heritability was the highest among the 5 traits studied (Table 1).

3.2. Bias of MTGP and STGP by the multistep method using absorbed GA-set records and by ssGBLUP

Table 3 summarizes the bias as the coefficients of regression of the adjusted phenotypes on GEBV for 5 traits by ST-multistep and MT-multistep using single-trait and multi-trait absorbed GA-set records, and by ST-ssGBLUP and MT-ssGBLUP analyses. The table showed the regression coefficients were mostly lower than 1, which suggests that the variance of the GEBV was slightly too high, relative to the variance of the adjusted phenotype. However, a regression coefficient >~~>~~1 was observed for MT-multistep for traits SP and FC40_120, indicating the variance of the GEBV was slightly too low relative to the variance of the adjusted phenotype.

Table 3 shows that MT-multistep prediction for trait LMP achieved the lowest bias. The trait A40 with the lowest GEBV accuracy in Table 2 is the most biased. Results demonstrate the multistep method is less biased than ssGBLUP. Furthermore, there is on average a bigger difference in regression coefficient between MT-multistep and ST-multistep (0.12) than between MT-ssGBLUP and ST-ssGBLUP (0.04), indicating a bigger variance in bias for multistep than for ssGBLUP.

3.3. Accuracy and bias of MT-multistep and ST-multistep using absorbed G-set records

Accuracies and biases of MT-multistep and ST-multistep prediction using 5-045 absorbed G-set records are in Table 4. The accuracies were calculated as the correlation between the predicted GEBV for the 1-028 validation animals, when the 4 017 animals were in the training set, and the adjusted phenotypes, divided by the square root of the heritability of the trait. As previously observed in the accuracy of multistep using absorbed GA-set records (Table 2), MT-multistep using absorbed G-set records in general achieved higher accuracy than ST-multistep. For ST-multistep, the prediction using absorbed G-set records achieved very similar accuracy to using absorbed GA-set records. However, for MT-multistep, the accuracy for the prediction

using absorbed G-set records over the 5 traits decreased ~~by~~ from 2.2% to 9.9% compared to using absorbed GA-set records.

For the bias results in Table 4, it is observed that generally there is less bias for the prediction that is more accurate. Among the 5 traits, GEBV prediction for trait A40 has the lowest accuracy~~,~~ and is most biased.

3.4. Accuracy and bias of ST-multistep method using multi-trait absorbed records

The absorption relied on the EBV and reliability obtained by a conventional pedigree-based genetic evaluation. For the situation of implementing traditional genetic evaluation on more than one trait, the evaluation can be implemented either one by one on each trait through a single-trait model, or simultaneously on all traits through a multi-trait model, resulting in either single-trait or multi-trait absorbed records. In the study, we have implemented ST-multistep based on both single-trait and multi-trait absorbed records.

Table 5 presents the accuracy and bias of ST-multistep using multi-trait absorbed GA-set and G-set records that were obtained from the absorption based on multi-trait EBV and reliability. Compared to the accuracy results for ST-multistep using single-trait absorbed GA-set records (Table 2) and G-set records (Table 4), it is observed that ST-multistep using multi-trait absorbed records in general achieved higher accuracy of the prediction. Over the 5 traits studied, when using multi-trait absorbed records, accuracy for ST-multistep increased by 2.2%~~ to ~~–17% for using absorbed GA-set records and by 1.3%~~ to ~~–13.1% for using absorbed G-set records. Trait D40_120 achieved the highest increase in ~~the ~~accuracy. However, the bias was not found ~~improved ~~to improve for the ST-multistep using multi-trait absorbed records.

3.5. Correlations of GEBV by the multistep method using absorbed GA-set records and by ssGBLUP

The GEBV correlations of 1-028 validation animals were compared between MTGP and STGP, and between the multistep method and ssGBLUP. Figure 1 shows the GEBV correlations between MTGP and STGP by the multistep method using multi-

trait and single-trait absorbed GA-set records (MTGP-STGP_multistep), and the GEBV correlations between MT-ssGBLUP and ST-ssGBLUP (MTGP-STGP_ssGBLUP). In the figure, traits SP, FC40_120 and LMP show similar correlations between MTGP and STGP, and the difference between the multistep method and ssGBLUP are small. GEBV of MTGP and STGP are less correlated for trait D40_120. For trait A40 both the multistep method and ssGBLUP results in the lowest GEBV correlations. The difference in GEBV correlations between the multistep method and ssGBLUP are larger for trait D40_120 and A40.

Figure 2 shows the GEBV correlations between MT-multistep using multi-trait GA-set records and MT-ssGBLUP (multistep-ssGBLUP_MTGP), and the GEBV correlations between ST-multistep using single-trait GA-set records and ST-ssGBLUP (multistep-ssGBLUP_STGP). One can see in the figure the similarly high GEBV correlations between the multistep method and ssGBLUP, varying from 0.807 for trait D40_120 to 0.874 for trait SP with an average of 0.848 in MTGP, and from 0.790 for trait D40_120 to 0.858 for trait SP with an average of 0.829 in STGP, which shows that the multistep method performed similarly to ssGBLUP.

3.6. Genetic trends in genotyped animals

Figure 3 shows the genetic trends in 5 traits as the average GEBV in genetic standard deviations for 5-045 genotyped animals born between 2010 and 2015. There were only 20 genotyped animals born in 2015 in Norsvin data. We plotted the genetic trends from 2014 to 2015 in dashed lines to indicate ~~the~~ that genetic trends may be strongly affected by the too–small data set. Figure 3 illustrates that for trait SP, FC40_120 and LMP, MTGP and STGP achieved similar genetic trends from 2010 to 2014. ~~Multistep~~ The multistep method may yield a slightly larger improvement in genetic trends for trait SP and FC40_120 than ssGBLUP. For traits D40_120 and A40, there is a difference of approximately $0.6\sigma_g$ in the average GEBV between MTGP and STGP.

4. DISCUSSION

4.1. Calculation of accuracy of GEBV

In this study, we examined a multistep method based on the Absorption approach and evaluated the accuracy and bias of the multistep method using both single-trait and multi-trait absorbed GA-set and G-set records for 5 traits in pig breeding, by comparison with ssGBLUP analyses. The accuracy of EBV is commonly defined as the correlation between animal's EBV and its true BV (TBV). In practice, usually the correlation between GEBV and (adjusted) phenotypes as an indicator of prediction accuracy since the TBV are unknown. Here, we used the correlation between GEBV and the adjusted phenotypes, divided by the square root of the heritability of the trait. The latter accounts for the imperfection of phenotypes as measures for TBV. ~~Heritability~~ The heritability of a trait measures the squared correlation between the TBV and the phenotypes. If the actual TBV of an animal can be completely predicted, ~~i.e.~~that is, EBV ~~=~~ = ~~-~~TBV, the correlation between these perfect EBV and the phenotypes equals the square root of the heritability. The square root of the heritability hence imposes an upper limit on how accurate the TBV of an animal can be predicted. In the presented study, the heritabilities ranged from 0.27 to 0.68 across the 5 traits used (Table 1), indicating the different levels of the predictability of TBV for the evaluation methods.

4.2. ssGBLUP and multistep method based on Absorption approach

In the study, the accuracy and bias results of ssGBLUP were used to compare with those of multistep methods using absorbed records. Both methods were able to exploit all available information. For ssGBLUP, an H-matrix was used to combine pedigree and genomic information, which enabled ssGBLUP to accommodate ungenotyped animals so that all available phenotypic information was used in the prediction. The multistep method relied on the absorption of phenotypic information of ungenotyped animals into the mixed model equations of genotyped ones to achieve the same goal of utilizing all available information. We have applied ssGBLUP and multistep methods both in STGP to explore the efficacy of the methods using only the phenotypic information of ungenotyped animals, and in MTGP to examine the methods utilizing not only the information from ungenotyped animals but also the

information from other traits. Table 2 hardly showed that one method achieved more accurate prediction than the other, suggesting that both methods may possess similar efficacy of exploiting all available information. However, the bias results in Table 3 showed that the multistep method achieved less bias compared to ssGBLUP. This agrees with the findings of Iheshuilor (Iheshiulor, 2016) that genomic predictions based on the absorbed dataset were generally less biased.

4.3. Single-t~~T~~rait and m~~M~~ulti-t~~T~~rait g~~G~~enomic p~~P~~rediction

When comparing the accuracy and bias between STGP and MTGP, it was observed in Table 2 that MTGP could generally lead to more accurate and less biased predictions. For the trait SP, MTGP achieved similar accuracy as STGP. This may indicate that records on other traits carry little information for the prediction of SP. Generally, the accuracy was a little improved ~~by~~ using a multi-trait instead of a single trait model~~s~~. However, for the traits D40_120 and A40, we found that the use of multi-trait models yielded more accurate predictions compared to using single-trait models. D40_120 and A40 had generally lower prediction accuracies and the multi-trait predictions helped to bring their prediction accuracies more in line with those of the other traits, and a multi-trait model is more recommended to use.

Table 2 showed on average a bigger difference in accuracy between MTGP and STGP for multistep methods than for ssGBLUP, indicating a greater improvement in the accuracy of MTGP using absorbed records. This suggests that the Absorption approach may benefit more from accounting for the information of other traits. Furthermore, the Absorption approach may possess the following merits: (1) the absorbed dataset may also be analysed by variable selection methods such as BayesA, B, C or R (Iheshiulor, 2016), whereas the extension of ssGBLUP to variable selection models is not straightforward, although the single-step Bayesian Regression approach (Fernando et al., 2014) could achieve this and~~;~~ (2) genomic prediction with absorbed data may avoid inversion of the G-matrix, for example by implementing SNP-BLUP, which would be computationally advantageous if the number of genotyped animals is high and thus the **G** matrix becomes very large.

4.4. Pseudo-phenotypic data set from Absorption approach

The multistep method we examined here was able to exploit all available information ~~by~~ using absorbed records of the GA or G animals. These absorbed records were pseudo-phenotypes of the GA or G animals corrected for information of all related ungenotyped animals in the complete data (GA ~~-~~+ ~~-~~D set) as well as the information on other traits if applicable, and weight-adjusted to achieve the same EBVs and reliabilities of the GA or G animals. Pseudo-phenotypes have been used as response variables in genetic evaluation. A typical example is deregressed proofs (DRP) in dairy cattle breeding. Calus et al. (2016) have compared the performance of different methods to compute DRP and weights for simultaneous deregression of cow and bull EBV.

In this study, two absorbed records were produced, ~~i.e.~~that is, absorbed G-set records and absorbed GA-set records, by applying the Absorption approach to G-set and GA-set. The G-set was a subset of GA-set. Comparison of the accuracy of multistep prediction using absorbed GA-set records (Table 2) and using absorbed G-set records (Table 4) showed that for MT-multistep, the prediction using multi-trait absorbed GA-set records achieved a higher accuracy than using multi-trait absorbed G-set. For ST-multistep the employment of single-trait absorbed GA-set records did not improve accuracy using a larger reference data set. A possible explanation is that for ST-multistep, the single-trait absorbed records used by the method only absorbed information on ungenotyped animals into genotyped animals. Although the size of the reference data set was different in predictions using absorbed GA-set and G-set records, the amount of available information on ungenotyped animals to absorb was about the same, ~~i.e.~~that is, the 5~~-~~045 genotyped animals in G-set have already obtained all available information from all ungenotyped animals, hence including 4~~-~~705 ancestral animals in GA-set did not improve accuracies. It was about the same reference information used to predict validation animals between ST-multistep using absorbed G-set records and that using absorbed GA-set records, which resulted in very similar accuracy of the prediction. However, for MT-multistep, the ancestors in

GA-set may get better imputed by the absorption of information from other traits. This may lead to the reference dataset in MT-multistep using absorbed GA-set records more informative than using absorbed G-set records, and hence an increase in the accuracy of the prediction. Furthermore, less bias was observed in MT-multistep using absorbed GA-set records (Table 3) than using absorbed G-set records (Table 4) which may also support a higher information content from the reference dataset in MT-multistep prediction using absorbed GA-set records.

A strategy to enable ST-multistep to accommodate the information from not only ungenotyped animals but also ~~the~~ other traits is to implement ST-multistep using multi-trait absorbed records. The accuracy results in Table 5 demonstrated the advantage of this method by an increase in accuracy on average by 7.2% (GA-set) and by 5% (G-set) compared to using single-trait absorbed records. Furthermore, ST-multistep using multi-trait absorbed records, which is more flexible than MT-multistep, allows to focus on only the traits of interest rather than predicting all the involved traits, and would effectively reduce the computational cost.

4.5. Practical implementation of the multistep method based on Absorption approach

For conventional multistep genetic evaluations, the drawbacks for instance of biased or inaccurate predictions for genotyped animals, the absence of gain in accuracy for ungenotyped animals, and incompatibility between EBVs for genotyped and ungenotyped animals (Bermann et al., 2022), undermine the prediction performance of the multiple-step method and may yield lower accuracy compared with ssGBLUP that includes both genotyped and ungenotyped animals simultaneously in a single genetic evaluation. In this study, we improved the prediction performance of the multistep method with Absorption approach, which achieved similar accuracy as ssGBLUP and in general lower bias.

Compared to raw phenotypes, EBVs may form a response variable data set of a higher quality to the prediction. This is because, for example in an animal model, all records that are available on an animal and its relatives are optimally used, while simultaneously adjusting for systematic environmental effects. For the breeders

involved in a long-term breeding program, they may have collected many conventional EBV and reliability data through traditional genetic evaluation. The implementation of multistep methods using absorbed records may be especially beneficial for these breeders since the absorption of information ~~of~~ on ancestral animals predicted in historical breeding practice may enable the breeders to rediscover the value of their previous traditional EBV assets in the genomic era.

With more and more genotyped animals, datasets used for genomic prediction might become huge. For example, the US dairy industry has now genotyped more than 3 million animals and the American Angus Association has more than 750,000 animals genotyped (Garcia et al., 2020). This requires a genomic prediction method that can handle huge datasets, with G-matrices that are computationally impossible to invert. The ssSNP-BLUP was developed to avoid the inversion of the G matrix. For the Absorption approach, one can simply calculate marker effects with absorbed records and predict GEBV with marker effects. In this way, the Absorption approach is able to handle huge datasets with millions of animals.

The multistep methods using absorbed records also have drawbacks. The absorbed records are pseudo records that may be complicated, and the weighting of the records requires approximations in complex models. Errors in EBVs due to poor conventional genetic evaluation may affect absorbed records and cause biased and inaccurate predictions. Furthermore, variance components cannot be estimated with the multistep approach.

5. CONCLUSION

The study shows that the multistep method using an absorbed dataset could achieve similarly accurate multi-trait prediction to the ssGBLUP method. But the multistep prediction showed in general less bias. For the genomic prediction where many traits are genetically correlated and may have different heritabilities, multi-trait models could yield higher accuracy than single-trait models, and hence are preferred. The implementation of the Absorption approach in multistep methods may be promising for the breeders to rediscover the value of previous traditional EBV estimation in historical breeding practices.

ACKNOWLEDGEMENTS~~ACKNOWLEDGEMENT~~

This study was financed by the Norwegian Research Council project: Genomics for improved sustainability and competitiveness of Norwegian breeding programs on livestock, aquaculture species and plants, project no. ES498330.

The authors acknowledge the Orion High Performance Computing Center (OHPCC) at the Norwegian University of Life Sciences (NMBU) for providing computational resources that have contributed to the research results reported within this paper. URL (internal): https://orion.nmbu.no.

CONFLICT OF INTEREST STATEMENT~~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 license~~d~~ for this study. Data might be available upon reasonable request from the authors with the permission of Norsvin SA.

REFERENCES

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

Belay, T. K., Eikje, L. S., Gjuvsland, A. B., Nordbo, O., Tribout, T., & Meuwissen, T. (2022). Correcting for base-population differences and unknown parent groups in single-step genomic predictions of Norwegian r~~R~~ed cattle. Journal of Animal Science, 100(9). skac227. doi:10.1093/jas/skac227

Bermann, M., Cesarani, A., Misztal, I., & Lourenco, D. (2022). Past, present, and future developments in single-step genomic models. Italian

Journal of Animal Science, 21(1), 673–-685. doi:10.1080/1828051x.2022.2053366

Calus, M. P. L., Vandenplas, J., ten Napel, J., & Veerkamp, R. F. (2016). Validation of simultaneous deregression of cow and bull breeding values and derivation of appropriate weights. Journal of Dairy Science, 99(8), 6403–-6419. doi:10.3168/jds.2016-11028

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

Christensen, O. F., Madsen, P., Nielsen, B., Ostersen, T., & Su, G. (2012). Single-step methods for genomic evaluation in pigs. Animal, 6(10), 1565–-1571. doi:10.1017/s1751731112000742

Estaghvirou, S. B. O., Ogutu, J. O., Schulz-Streeck, T., Knaak, C., Ouzunova, M., Gordillo, A., & Piepho, H. P. (2013). Evaluation of approaches for estimating the accuracy of genomic prediction in plant breeding. BMC Genomics~~Bmc Genomics~~, 14. doi:10.1186/1471-2164-14-860 860

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

Garcia, A. L. S., Masuda, Y., Tsuruta, S., Miller, S., Misztal, I., & Lourenco, D. (2020). Indirect predictions with a large number of genotyped animals using the algorithm for proven and young. Journal of Animal Science, 98(6). skaa154 doi:10.1093/jas/skaa154

Garcia-Ruiz, A., Cole, J. B., VanRaden, P. M., Wiggans, G. R., Ruiz-Lopez, F. J., & Van Tassell, C. P. (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proceedings of the National Academy of Sciences of the United States of America, 113(28), E3995-E4004. doi:10.1073/pnas.1519061113

Gilmour, A. R., Gogel, B. J., Cullis, B. R., & Thompson, R. (2006). ASReml User Guide Release 2.0.

Iheshiulor, O. O. M. (2016). Improving the accuracy of genomic predictions in numerically small populations. PhD, Norwegian University of Life Sciences, Ås.

Iversen, M. W., Nordbo, O., Gjerlaug-Enger, E., Grindflek, E., Lopes, M. S., & Meuwissen, T. H. E. (2017). Including crossbred pigs in the genomic relationship matrix through utilization of both linkage disequilibrium and linkage analysis. Journal of Animal Science, 95(12), 5197–-5207. doi:10.2527/jas2017.1705

Karaman, E., Lund, M. S., & Su, G. S. (2020). Multi-trait single-step genomic prediction accounting for heterogeneous (co)variances over the genome. Heredity, 124(2), 274–-287. doi:10.1038/s41437-019-0273-4

Kudinov, A. A., Mantysaari, E. A., Aamand, G. P., Uimari, P., & Stranden, I. (2020). Metafounder approach for single-step genomic evaluations of rRed dDairy cattle. Journal of Dairy Science, 103(7), 6299–6310. doi:10.3168/jds.2019-17483

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

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

Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral rRelationships uUsing Metafounders: Finite aAncestral pPopulations and aAcross pPopulation rRelationships. Genetics, 200(2), 455–468. doi:10.1534/genetics.115.177014

Lopez, B. I., Santiago, K. G., Lee, D., Cho, Y., Lim, D., & Seo, K. (2020). Single-sStep gGenomic eEvaluation for mMeat qQuality tTraits, sSensory cCharacteristics, and fFatty-aAcid cComposition in Duroc pPigs. Genes, 11(9). 1062 doi:10.3390/genes11091062

Madsen, P., & Jensen, J. (2013). A user's gGuide to DMU Ver. 6, rel. 5.2. Retrieved from http://dmu.agrsci.dk/DMU/Doc/Current/dmuv6_guide.5.2.pdf

Madsen, P., Jensen, J., Labouriau, R., Christensen, O. F., & Sahana, G. (2014). DMU—A pPackage for aAnalyzing mMultivariate mMixed mModels in quantitative gGenetics and gGenomics. Paper presented at the 10th wWorld cCongress of gGenetics aApplied to lLivestock pProduction; August 17–22, 2014; Vancouver, BC, Canada.

Meuwissen, T. H. E., & Goddard, M. E. (1999). Marker assisted estimation of breeding values when marker information is missing on many animals. Genetics Selection Evolution, 31(4), 375–394. doi:10.1051/gse:19990405

Meuwissen, T. H. E., & Goddard, M. E. (2010). The use of family relationships and linkage disequilibrium to impute phase and missing genotypes in up to whole-genome sequence density genotypic data. Genetics, 185(4), 1441–1449.

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.

Meuwissen, T. H. E., Svendsen, M., Solberg, T., & Ødegård, J. (2015). Genomic predictions based on animal models using genotype imputation on a national scale in Norwegian rRed cattle. Genetics Selection Evolution, 47(1), 79. doi:10.1186/s12711-015-0159-8

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. doi:10.3168/jds.2009-2064

Mote, B. E., Serenius, T. V., Supakorn, C., & Stalder, K. J. (2019). Genetic improvement of sow longevity and its economic impact on commercial pork production. South African Journal of Animal Science, 49(6), 1036–1046. doi:10.4314/sajas.v49i6.8

Ricard, A., Danvy, S., & Legarra, A. (2012). Computation of deregressed proofs for genomic selection when own phenotypes exist with an application in French show-jumping horses. Journal of Animal Science, 91(3), 1076–1085. doi:10.2527/jas.2012-5256

Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. Journal of Animal Breeding and Genetics, 123(4), 218–223. doi:10.1111/j.1439-0388.2006.00595.x

Semagn, K., Crossa, J., Cuevas, J., Iqbal, M., Ciechanowska, I., Henriquez, M. A., ...Randhawa H., Beres B.L., Aboukhaddour R., McCallum B., Brûlé-Babel A.L., N'Diaye A., Pozniak C., Spaner D. (2022). Comparison of single-trait and multi-trait genomic predictions on agronomic and disease resistance traits in spring wheat. Theoretical and Applied Genetics, 135(8), 2747–2767. doi:10.1007/s00122-022-04147-3

Song, H. L., Zhang, Q., & Ding, X. D. (2020). The superiority of multi-trait models with genotype-by-environment interactions in a limited number of environments for genomic prediction in pigs. Journal of Animal Science and Biotechnology, 11(1). doi:10.1186/s40104-020-00493-8, 88

Stranden, I., Aamand, G. P., & Mantysaari, E. A. (2022). Single-step genomic BLUP with genetic groups and automatic adjustment for allele coding. Genetics Selection Evolution, 54(1). doi:10.1186/s12711-022-00721-x, 38

Stranden, I., & Lidauer, M. (1999). Solving large mixed linear models using preconditioned conjugate gradient iteration. Journal of Dairy Science, 82(12), 2779–2787.

VanRaden, P. M. (2008). Efficient mMethods to cCompute gGenomic pPredictions. Journal of Dairy Science, 91(11), 4414–4423.

VanRaden, P. M., Tooker, M. E., & Cole, J. B. (2009). Can you believe those genomic evaluations for young bulls? Journal of Dairy Science, 92(E-Suppl. 1), 314 (abstr. 279).

Villar-Hernández B.J.Villar-Hernandez, B. D., Pérez-Elizalde S.Perez-Elizalde, S., Martini, J. W. R., Toledo, F., Perez-Rodriguez, P., Krause, M., ...García-Calvillo I.D.Crossa, J., Covarrubias-Pazaran G., Crossa J. (2021). Application of multi-trait Bayesian decision theory for parental genomic selection G3-Genes Genomes Genetics, 11(2). jkab012 doi:10.1093/g3journal/jkab012

Vuori, K., Strandén, I., Lidauer, M., & Mäntysaari, E. A. (2006). MiX99--- Eeffective solver for large and complex linear mixed models. Minas Gerais.

Wang, X. Q., Wang, L. G., Shi, L. Y., Zhang, P. F., Li, Y., Li, M. Y., ... Tian J.Zhao, F. P., Wang L., Zhao F. (2022). GWAS of rReproductive

t~~T~~raits in l~~L~~arge w~~W~~hite p~~P~~igs on Chip and Imputed w~~W~~hole-g~~G~~enome s~~S~~equencing d~~D~~ata. International Journal of Molecular Sciences, 23(21). 13338 doi:10.3390/ijms232113338

Wiggans, G. R., & Carrillo, J. A. (2022). Genomic selection in United States~~United States~~ dairy cattle. Frontiers in Genetics, 13. 994466. doi:10.3389/fgene.2022.994466

TABLE 1. Heritabilities (h^2), genetic standard deviations (σ_g) and number of phenotypic records for the traits.

Trait	h^2	σ_g	Total number of records	Number of reference records for ssGBLUP	Comment
SP	0.27	1.00	8~~-~~661	7~~-~~633	Slaughter percentage
FC40_120	0.32	6.43	9~~-~~086	8~~-~~058	Feed consumption from 40~~-kg~~ to 120 kg
D40_120	0.48	4.85	9~~-~~248	8~~-~~220	Days from 40~~-kg~~ to 120 kg
A40	0.50	4.40	9~~-~~641	8~~-~~613	Age at 40 kg (days)
LMP	0.68	2.26	8~~-~~661	7~~-~~633	Lean meat percentage

TABLE 2. Accuracy of GEBV prediction using single-trait (ST-) and multi-trait (MT-) ssGBLUP and multistep method using absorbed GA-set records.

Trait	ST-ssGBLUP	MT-ssGBLUP	ST-multistep	MT-multistep
SP	0.657	0.667	0.669	0.665
FC40_120	0.637	0.673	0.649	0.698
D40_120	0.503	0.571	0.496	0.559
A40	0.414	0.491	0.391	0.502
LMP	0.627	0.661	0.616	0.675

TABLE 3. Regression coefficient of GEBV prediction using single-trait (ST-) and multi-trait (MT-) ssGBLUP and multistep method using absorbed GA-set records.

Trait	ST-ssGBLUP	MT-ssGBLUP	ST-multistep	MT-multistep
SP	0.89	0.91	0.96	1.09
FC40_120	0.87	0.88	0.97	1.09
D40_120	0.71	0.79	0.79	0.91
A40	0.59	0.64	0.62	0.80
LMP	0.85	0.90	0.94	1.01

TABLE 4. Accuracy and regression coefficient of GEBV prediction for single-trait (ST-) and multi-trait (MT-) multistep method using absorbed G-set records.

Trait	Accuracy		Bias	
	ST-multistep	MT-multistep	ST-multistep	MT-multistep
SP	0.667	0.650	0.95	0.98
FC40_120	0.641	0.657	0.97	0.95
D40_120	0.496	0.530	0.77	0.81
A40	0.394	0.452	0.60	0.67
LMP	0.621	0.640	0.89	0.78

TABLE 5. Accuracy and regression coefficient of GEBV prediction for single-trait multistep method using multi-trait absorbed GA-set and G-set records.

Trait	Accuracy		Bias	
	GA-set	G-set	GA-set	G-set
SP	0.684	0.676	1.21	0.91
FC40_120	0.668	0.657	1.12	0.89

D40_120	0.581	0.561	0.94	0.93
A40	0.413	0.403	0.64	0.72
LMP	0.666	0.655	1.09	0.96

FIGURE 1. GEBV correlations between MTGP and STGP by the multistep method using multi-trait and single-trait absorbed GA-set records (MTGP-STGP_multistep), and the GEBV correlations between MT-ssGBLUP and ST-ssGBLUP (MTGP-STGP_ssGBLUP).

FIGURE 2. GEBV correlations between MT-multistep using multi-trait GA-set records and MT-ssGBLUP (multistep-ssGBLUP_MTGP), and the GEBV correlations between ST-multistep using single-trait GA-set records and ST-ssGBLUP (multistep-ssGBLUP_STGP).

FIGURE 3. Genetic trends of trait SP, FC40_120, D40_120, A40 and LMP in genotyped animals; the *y*-axis shows average GEBV in genetic standard deviations. ssGBLUP_STGP and ssGBLUP_MTGP represents the average GEBV of ST-ssGBLUP and MT-ssGBLUP; multistep_STGP and multistep_MTGP represents the average GEBV of MT-multistep using multi-trait GA-set records and ST-multistep using single-trait GA-set records. The genetic trends from 2014 to 2015 were plotted in dashed lines since there were only 20 genotyped animals born in 2015 in Norsvin data and the genetic trends from 2014 to 2015 may be strongly affected by the too small data set for 2015.