Improving genetic evaluation of litter size and piglet mortality for both genotyped and nongenotyped individuals using a single-step method

X. Guo,* O. F. Christensen,* T. Ostersen,† Y. Wang,‡ M. S. Lund,* and G. Su*2

*Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark; †Pig Research Centre, Danish Agricultural and Food Council, DK-1609 Copenhagen, Denmark; and ‡College of Animal Science and Technology, China Agricultural University, 100193 Beijing, P. R. China

ABSTRACT: A single-step method allows genetic evaluation using information of phenotypes, pedigree, and markers from genotyped and nongenotyped individuals simultaneously. This paper compared genomic predictions obtained from a single-step BLUP (SSBLUP) method, a genomic BLUP (GBLUP) method, a selection index blending (SELIND) method, and a traditional pedigree-based method (BLUP) for total number of piglets born (TNB), litter size at d 5 after birth (LS5), and mortality rate before d 5 (Mort; including stillbirth) in Danish Landrace and Yorkshire pigs. Data sets of 778,095 litters from 309,362 Landrace sows and 472,001 litters from 190,760 Yorkshire sows were used for the analysis. There were 332,795 Landrace and 207,255 Yorkshire animals in the pedigree data, among which 3,445 Landrace pigs (1,366 boars and 2,079 sows) and 3,372 Yorkshire pigs (1,241 boars and 2,131 sows) were genotyped with the Illumina PorcineSNP60 BeadChip. The results showed that the 3 methods with marker information (SSBLUP, GBLUP, and SELIND) produced more accurate predictions for genotyped animals than the pedigree-based method. For genotyped animals, the average of reliabilities for all traits in both breeds using traditional BLUP was 0.091, which increased to 0.171 when using GBLUP and to 0.179 when using SELIND and further increased to 0.209 when using SSBLUP. Furthermore, the average reliability of EBV for nongenotyped animals was increased from 0.091 for traditional BLUP to 0.105 for the SSBLUP. The results indicate that the SSBLUP is a good approach to practical genomic prediction of litter size and piglet mortality in Danish Landrace and Yorkshire populations.

Key words: genomic selection, litter size, piglet mortality, single-step method


INTRODUCTION

Selection for total number of piglets born (TNB) to improve litter size at weaning is generally associat-
ed with an increase of piglet mortality (Johnson et al., 1999; Lund et al., 2002; Su et al., 2007). Accordingly, the Danish breeding program changed breeding goal from selection for TNB to litter size at d 5 after birth (LS5) in 2004, which has led to an increase of litter size at weaning and a decrease of piglet mortality (Nielsen et al., 2013). With the development of high-throughput genotyping technologies, it is possible to predict genomic breeding values using genomewide markers. Many studies have shown that a genomic BLUP (GBLUP) model performs as well as Bayesian variable selection models for most traits (Hayes et al., 2009a; VanRaden et al., 2009), and therefore it is widely used for practical genomic prediction because it is simple and has low computational demands. Generally, genomic
predictions use information of genotyped animals. In practice, however, not all individuals can be genotyped. The accuracy of genomic prediction can be improved by combining information of nongenotyped animals. VanRaden et al. (2009) proposed a selection index blending (SELIND) approach that includes direct estimated genomic breeding value and traditional EBV. Furthermore, a single-step BLUP (SSBLUP) method to predict breeding values using information from genotyped and nongenotyped animals simultaneously was proposed (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010). Experiences with field data in cattle, pigs, and chickens indicate that SSBLUP is more accurate and much simpler than multistep methods (Aguilar et al., 2011; Chen et al., 2011; Forni et al., 2011; Christensen et al., 2012; Gao et al., 2012).

The objective of this study was to test the hypothesis that genomic models predict more accurately than traditional BLUP and that SSBLUP predicts more accurately than a GBLUP model for litter size and piglet mortality, based on data from Danish Landrace and Yorkshire populations.

MATERIALS AND METHODS

Data

The data from breeding herds and multiplier herds were supplied by the Danish Agriculture and Food Council, Pig Research Centre (Copenhagen, Denmark). The records in the years before measurements started on LS5, in 2004, were removed. After editing, 778,095 litters of 309,362 Landrace sows and 472,001 litters of 190,760 Yorkshire sows were used for the analysis. All sows with records were born from 1998 to 2012. The DMU Trace program (Madsen, 2010) was used to trace the pedigree back to 1994. Consequently, there were 332,795 Landrace and 207,255 Yorkshire animals in the pedigree data, among which 3,445 Landrace pigs (1,366 boars and 2,079 sows) and 3,372 Yorkshire pigs (1,241 boars and 2,131 sows) were genotyped using Illumina PorcineSNP60 BeadChip (Illumina, San Diego, CA). A total of 38,434 and 38,630 SNP markers for Landrace and Yorkshire, respectively, met the following requirements. Each animal had a call rate greater than 0.9. Each marker had a minor allele frequency greater than 0.05, a call frequency score greater than 0.9, and not strong deviation from Hardy-Weinberg equilibrium ($P > 10^{-7}$). For each animal and SNP–genotype combination, the GenCall score (Oliphant et al., 2002) was greater than 0.6. Genotypes less than 0.6 were defined as missing. Animals with missing genotypes were assigned to be the population mean for the missing genotypes. The traits in the analysis were TNB, LS5, and mortality rate before d 5 ($\text{Mort}$), which was calculated as the difference between TNB and LS5 and divided by TNB. All litters had records for the 3 traits as long as they had records of TNB.

Statistical Models

The statistical methods used for predicting breeding values in this study were a traditional BLUP method with pedigree-based relationship matrix, a GBLUP method with marker-based relationship matrix, a SELIND approach, and a SSBLUP method with a combined relationship matrix constructed from marker and pedigree information.

Traditional BLUP

A traditional animal model with a pedigree-based relationship matrix was used to estimate the genetic parameters, predict breeding values, and derive the response variable for genomic prediction using GBLUP method. The analysis was performed for each trait and each breed separately.

The basic model was

$$y = Xb + Z_a a + Z_{pe} pe + Z_s s + e,$$

in which $y$ was the vector of phenotypic values of a trait in Landrace or Yorkshire; $b$ was the vector of fixed effects including herd–year–season, parity, month at farrowing, and regressions on hybrid indicator ($0 =$ pure litter and $1 =$ hybrid litter), age at first farrowing (age and age$^2$, only for first parity), parity correction ($0 =$ no correction and $1 =$ correction), farrowing interval (interval and interval$^2$, not for first parity), AI ($0 =$ natural mating and $1 =$ AI); $a$ was the vector of additive genetic effects; $pe$ was the vector of permanent effects of sows; $s$ was the vector of service sire effects; $e$ was the vector of random residuals; and $X$, $Z_a$, $Z_{pe}$, and $Z_s$ were incidence matrices associating $b$, $a$, $pe$, and $s$ with $y$. The random effects were assumed to be independent of each other and normally distributed, that is, $a \sim N(0, \mathbf{A} \sigma^2_a), pe \sim N(0, \mathbf{I} \sigma^2_{pe}), s \sim N(0, \mathbf{I} \sigma^2_s)$, and $e \sim N(0, \mathbf{I} \sigma^2_e)$, in which $\mathbf{A}$ was the matrix of additive genetic relationships constructed based on the pedigree; $\mathbf{I}$ was the identity matrix; and $\sigma^2_a$, $\sigma^2_{pe}$, $\sigma^2_s$, and $\sigma^2_e$ were the variances of additive genetic effects, permanent environment effects, service sire effects, and residuals, respectively.

Variance components and corrected phenotypic values ($\hat{y}$) were estimated using the full data. Phenotypic variance ($\sigma^2_p$) was defined as $\sigma^2_p = \sigma^2_a + \sigma^2_{pe} + \sigma^2_s + \sigma^2_e$. 
Heritability \( (h^2) \) was defined as the ratio of additive genetic variance to phenotypic variance \( (h^2 = \sigma_a^2 / \sigma_i^2) \).

The \( y_c \) of sows were calculated as EBV plus the average of estimated residuals over parities for a sow \( (y_c = \hat{a} + \sum \hat{e}_i / n_p, \) in which \( n_p \) was number of parities for a sow). Heritability of corrected phenotypic value \( (h^2_c) \) was defined as \( h^2_c = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2) \). The reliabilities of \( y_c \) for sows were:

\[
r_{y_i}^2 = \frac{n_p h_{y_i}^2}{(n_p - 1) h_{y_i}^2 + 1}.
\]

For boars, the \( y_c \) were calculated from \( y_c \) of nongenotyped daughters, \( y_{do} = \sum (w_d \times y_d) / \sum w_d \), in which \( w_d \) is relative weight of a daughter\’s \( y_d \) in terms of its contribution to the \( y_c \) of the boar and was calculated as \( w_d = 1/k \), with \( k = (4 - r_i^2) \). The reliabilities of \( y_c \) for boars were calculated as \( r_{y_i}^2 = n_d/(n_d + \lambda) \), in which

\[
\lambda = \frac{4 - \sum r_i^2}{\sum r_i^2/n_d}
\]

and \( n_d \) was the number of nongenotyped daughters of a boar. The \( y_c \) were used as the response variable for GBLUP method and in the procedure of validation of predictions.

**Genomic BLUP**

The model used to predict genomic breeding value for the animals with genotypes was as follows:

\[
y_c = \mathbf{1}\mu + \mathbf{Z}_g\mathbf{g} + \mathbf{Z}_u\mathbf{u} + \mathbf{e}, \tag{2}
\]

in which \( y_c \) was the data vector of corrected phenotypic values of genotyped boars and sows, \( \mu \) was the overall mean, \( \mathbf{1} \) was a vector of 1s, \( \mathbf{g} \) was a vector of genomic breeding value to be estimated, \( \mathbf{u} \) was the vector of residual polygenic effects that were not captured by the SNP, \( \mathbf{Z}_g \) and \( \mathbf{Z}_u \) were design matrices allocating records to \( \mathbf{g} \) and \( \mathbf{u} \), and \( \mathbf{e} \) was the vector of residuals. It was assumed that \( \mathbf{g} \sim \text{N}(\mathbf{0}, \mathbf{G} \sigma_g^2) \) and \( \mathbf{u} \sim \text{N}(\mathbf{0}, \mathbf{A} \sigma_u^2) \), in which \( \sigma_g^2 \) and \( \sigma_u^2 \) were the variances of additive genetic effect and residual polygenic effect, \( \mathbf{A} \) was the pedigree-based relationship matrix, and \( \mathbf{G} \) was the marker-based genomic relationship matrix (VanRaden, 2008; Hayes et al., 2009b). The allele frequencies used in constructing the \( \mathbf{G} \) matrix were estimated from observed markers. For random residuals, it was assumed that \( \mathbf{e} \sim \text{N}(\mathbf{0}, \mathbf{D} \sigma_e^2) \), in which \( \sigma_e^2 \) was the residual variance and \( \mathbf{D} \) was a diagonal matrix containing the elements \( d_{ii} = 1/w_i^* \), in which \( w_i^* \) was standardized weight of \( y_i \). The weight of \( y_i \) was calculated as \( w = r_i^2/(1 - r_i^2) \) and then standardized by average weight \( (w_i^* = w_i / \bar{w}) \), and thus the mean of \( w_i^* \) was 1.

**Selection Index Blending**

According to VanRaden et al. (2009), genomic enhanced breeding value (GeBv) can be obtained using a selection index as follows:

\[
\text{GeBv} = b_1 \text{DGP} + b_2 \text{PI}_s + b_3 \text{PI}_t,
\]

in which \( \text{DGP} \) was direct genomic prediction, \( \text{PI} \) was the traditional pedigree index (i.e., EBV for the animals without their own records and the offspring\’s records), \( \text{PI}_s \) was estimated using the subset data set of genotyped animals used in GBLUP model, and \( \text{PI}_t \) was estimated using the full data set used in traditional BLUP model. In the present study, DGP was computed from GBLUP model without polygenic effect, and both \( \text{PI}_s \) and \( \text{PI}_t \) were computed using model [1]. The GBLUP model used to predict DGP was as follows:

\[
y_c = \mathbf{1}\mu + \mathbf{Z}_g\mathbf{g} + \mathbf{e}, \tag{3}
\]

in which \( y_c, \mu, \mathbf{1}, \mathbf{Z}_g, \mathbf{g}, \) and \( \mathbf{e} \) were same as in model [2].

The equation system for the selection index was as presented in the study by VanRaden et al. (2009) and Su et al. (2012):

\[
b_1 V_{11} + b_2 V_{12} + b_3 V_{13} = V_{11},
\]

\[
b_1 V_{12} + b_2 V_{22} + b_3 V_{23} = V_{22}, \quad \text{and}
\]

\[
b_1 V_{13} + b_2 V_{23} + b_3 V_{33} = V_{33},
\]

in which \( V_{11}, V_{22}, V_{33} \) were the model-based reliabilities \( (1 - \text{PEV}/\sigma_e^2) \) of DGP, \( \text{PI}_s \), and \( \text{PI}_t \), respectively, and PEV was prediction error variance. We defined \( V_{12} = V_{22}, V_{23} = V_{22}, \) and \( V_{13} = V_{22} + (V_{11} - V_{22}) / (V_{33} - V_{22}) \). To ensure that matrix \( \mathbf{V} \) was positive definite, \( V_{11} \) and \( V_{33} \) were constrained to be greater than \( V_{22} \). In the present study, \( V_{11}, V_{22}, \) and \( V_{33} \) were obtained by inverting the coefficient matrix of the corresponding model.

**Single-Step BLUP**

A SSBLUP method uses information from both genotyped and nongenotyped individuals simultaneously by combining the genomic relationship matrix (\( \mathbf{G} \)) with the pedigree-based numerator relationship matrix (\( \mathbf{A} \)). The model was the same as model [1] but it was assumed that \( \mathbf{a} \sim \text{N}(\mathbf{0}, \mathbf{H} \sigma_a^2) \). The matrix \( \mathbf{H} \) was
the modified genetic relationship matrix by combining marker and pedigree information (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010):

$$
H = \begin{bmatrix}
G_{\omega} & G_{\omega}A_{11}^{-1}A_{12} \\
A_{21}A_{11}^{-1}G_{\omega} & A_{21}A_{11}^{-1}G_{\omega}A_{11}^{-1}A_{12} + A_{22} - A_{21}A_{11}^{-1}A_{12}
\end{bmatrix},
$$

in which $A_{11}$ was the submatrix of pedigree-based relationship matrix ($A$) for genotyped animals, $A_{22}$ was the submatrix of $A$ for nongenotyped animals, $A_{12}$ (or $A_{21}$) was the submatrix of $A$ describing the relationships between genotyped and nongenotyped animals, and $G_{\omega} = (1 - \omega)G^* + \omega A_{11}$, in which $\omega$ was a relative weight, which could explain the fraction of the genetic variance not captured by markers and $G^*$ was an adjusted $G$ (see below). In the present study, a set of different $\omega$ were used to construct $G_{\omega}$, and the detailed results to be reported were those from the $H$ matrix with $\omega = 0.5$.

The inverse of $H$ matrix was

$$
H^{-1} = \begin{bmatrix}
G_{\omega}^{-1} - A_{11}^{-1} & 0 \\
0 & 0
\end{bmatrix} + A_{11}^{-1}.
$$

The $G$ matrix was adjusted for the differences in location and scale from pedigree relationship matrix ($A_{11}$) using 2 parameters, $\alpha$ and $\beta$ (Christensen et al., 2012), that is,

$$
G^* = G\beta + \alpha.
$$

The parameters $\alpha$ and $\beta$ were derived from the following equations:

$$
\text{Avg.diag}(G)\beta + \alpha = \text{Avg.diag}(A_{11}) \quad \text{and} \quad \text{Avg.offdiag}(G)\beta + \alpha = \text{Avg.offdiag}(A_{11}),
$$

in which Avg.diag is the average of diagonal elements, Avg.offdiag is the average of off-diagonal elements, and then $G^*$ was used to replace $G$ to construct the combined relationship matrix in the SSBLUP method. The adjustment using parameters $\alpha$ and $\beta$ ensured the genomic relationship matrix and the pedigree-based relationship matrix to be compatible.

The variance components and predictions were computed using the DMUAI procedure in the DMU software (Madsen and Jensen, 2012).

Validation

The birthdate of April 1, 2012, was chosen as the cut-off date to divide the population into training and validation data sets. For GBLUP, the training data sets consisted of 1,366 boars and 1,529 sows in Landrace and 1,241 boars and 1,610 sows in Yorkshire. For traditional BLUP and SSBLUP methods, the training data sets consisted of 764,442 records from 296,195 sows for Landrace and 461,501 records from 180,579 sows for Yorkshire. For all methods, 550 Landrace and 521 Yorkshire genotyped sows were used as validation data. Furthermore, the validation data included 12,617 Landrace and 9,660 Yorkshire nongenotyped sows when using traditional BLUP and SSBLUP methods.

The methods were compared with regard to reliability of predictions ($r^2$), evaluated as squared correlations between the predicted breeding values (EBV) and $y_c$ for individuals in the validation data sets, divided by $h^2_{y_c}$, that is, $r^2 = \text{cor}^2(EBV, y_c)/h^2_{y_c}$.

The differences were assessed using the Hotelling–Williams $t$ test (Dunn and Clark, 1971; Revelle, 2014) at a significance level of 5%. Furthermore, to assess possible inflation of predictions, a regression of corrected phenotypic values on EBV was calculated. A regression coefficient significantly less than 1 would reflect an inflation (Su et al., 2012). These comparisons were made for all animals in the validation data set and for genotyped and nongenotyped animals separately.

Results

The descriptive statistics for analyzed traits are shown in Table 1. In the Landrace population, the average TNB was 15.04 and the average LS5 was 12.25, which incurred a mortality rate of 0.18. In the Yorkshire population, even though the sows produced 0.5 more piglets, LS5 was only 0.3 higher than in Landrace with a bit higher mortality rate.

<table>
<thead>
<tr>
<th>Breed</th>
<th>Trait</th>
<th>N-obs</th>
<th>Average</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landrace</td>
<td>TNB</td>
<td>778,095</td>
<td>15.04</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>778,095</td>
<td>12.25</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>778,095</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>Yorkshire</td>
<td>TNB</td>
<td>472,001</td>
<td>15.54</td>
<td>4.21</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>472,001</td>
<td>12.54</td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>472,001</td>
<td>0.18</td>
<td>0.17</td>
</tr>
</tbody>
</table>

$^1$TNB = total number of piglets born; LS5 = litter size at Day 5 after birth; Mort = mortality rate before Day 5 (including stillbirth).

$^2$N-obs = number of observations.

The variance components and heritabilities are shown in Table 2. The range of heritabilities for the 3 traits was between 0.08 and 0.11 in both Landrace and Yorkshire. Across the 3 analyzed traits, the estimated heritability of TNB in Landrace was the highest, and
the other traits in both breeds had similar estimated heritabilities (0.08–0.09).

The validation reliabilities of predictions using traditional BLUP and 3 genomic prediction methods are shown in Table 3. For the genotyped animals, the 3 methods with marker information (GBLUP, SELIND, and SSBLUP methods) generally provided higher reliabilities of predictions than traditional BLUP. The superiority of genomic prediction over traditional prediction was more obvious in Yorkshire than in Landrace for litter size traits. In contrary to litter size, the advantage of using marker information for Mort was larger in Landrace than in Yorkshire. Although the reliabilities of genomic predictions for genotyped animals were higher than traditional predictions in varying degrees, only the improvement of Mort in Landrace was statistically significant, which could be due to small number of genotyped validation animals.

As shown in Table 3, the superiority of using the SSBLUP method compared to the traditional BLUP method was also observed in nongenotyped animals. In Landrace, the validation reliabilities of TNB, LS5, and Mort were 0.126, 0.072, and 0.068, respectively, when using a traditional BLUP method, and these were increased by 1 to 2 percentage points when applying SSBLUP method. Similarly, the average of reliabilities for the 3 traits was increased by 1 to 2 percentage points in the Yorkshire population. The increases of reliabilities when using SSBLUP method were statistically significant for TNB in both Landrace and Yorkshire populations and for LS5 in Yorkshire.

When reliabilities were calculated based on the validation data including both genotyped and nongenotyped animals, superiority of SSBLUP over traditional BLUP became clearer in the significance test (Table 3). Across 3 analyzed traits, the reliability of predicted breeding values using traditional BLUP and SSBLUP methods were 0.088 and 0.107, respectively, in Landrace, and 0.094 and 0.115, respectively, in Yorkshire. The differences between reliabilities

### Table 2. Variance components1 with SE and heritabilities

<table>
<thead>
<tr>
<th>Breed</th>
<th>Trait</th>
<th>( \sigma^2_a )</th>
<th>( \sigma^2_p )</th>
<th>( \sigma^2_s )</th>
<th>( \sigma^2_e )</th>
<th>( \sigma^2_p )</th>
<th>( h^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TNB</td>
<td>1.76 (0.04)</td>
<td>1.28 (0.03)</td>
<td>1.28 (0.03)</td>
<td>12.00 (0.02)</td>
<td>16.32</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>Landrace</td>
<td>LS5</td>
<td>1.25 (0.03)</td>
<td>0.97 (0.02)</td>
<td>0.69 (0.01)</td>
<td>10.48</td>
<td>13.39</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.0025 (0.0017)</td>
<td>0.0003 (0.002)</td>
<td>0.0219 (0.02)</td>
<td>(6.5 × 10^{-5})</td>
<td>(4.73 × 10^{-5})</td>
<td></td>
</tr>
<tr>
<td>Yorkshire</td>
<td>TNB</td>
<td>1.32 (0.04)</td>
<td>1.13 (0.03)</td>
<td>0.98 (0.03)</td>
<td>12.08</td>
<td>15.51</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>1.07 (0.04)</td>
<td>1.01 (0.03)</td>
<td>0.49 (0.014)</td>
<td>10.63</td>
<td>13.20</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.0026 (0.0023)</td>
<td>0.0002 (0.002)</td>
<td>0.0233 (0.02)</td>
<td>(8.91 × 10^{-5})</td>
<td>(6.89 × 10^{-5})</td>
<td></td>
</tr>
</tbody>
</table>

1 \( \sigma^2_a \) = additive genetic variance; \( \sigma^2_p \) = variance of permanent effect; \( \sigma^2_s \) = variance of service-sire effect; \( \sigma^2_e \) = residual variance; \( \sigma^2_p \) = phenotypic variance; \( h^2 \) = heritability of corrected phenotypic value.

2 TNB = total number of piglets born; LS5 = litter size at Day 5 after birth; Mort = mortality rate before Day 5 (including stillbirth).

### Table 3. Validation reliability of EBV

<table>
<thead>
<tr>
<th>Breed</th>
<th>Trait</th>
<th>All</th>
<th>Genotyped</th>
<th>Nongenotyped</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BLUP</td>
<td>GBLUP3</td>
<td>SELIND4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landrace</td>
<td>TNB</td>
<td>0.128</td>
<td>0.155*</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>0.071</td>
<td>0.081</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.066</td>
<td>0.086*</td>
<td>0.030</td>
</tr>
<tr>
<td>Yorkshire</td>
<td>TNB</td>
<td>0.148</td>
<td>0.178*</td>
<td>0.251</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>0.061</td>
<td>0.083*</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.074</td>
<td>0.085</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.091</td>
<td>0.111</td>
<td>0.091</td>
</tr>
</tbody>
</table>

1 TNB = total number of piglets born; LS5 = litter size at Day 5 after birth; Mort = mortality rate before Day 5 (including stillbirth).

2 SSBLUP = single-step BLUP.

3 GBLUP = genomic BLUP.

4 SELIND = selection index blending.

*Significantly differs from BLUP at \( P < 0.05 \).
obtained from SSBLUP and traditional BLUP were statistically significant for TNB in both Landrace and Yorkshire, Mort in Landrace, and LS5 in Yorkshire. The regression coefficients of predictions are shown in Table 4. For genotyped animals, the regression coefficients of different methods were not consistent among 3 traits in 2 breeds. However, the regression coefficients were not statistically significantly far from 1 in most cases due to large SE (the SE were in the range from 0.230 to 0.580). There was no clear trend that a model was better than the others in unbiasedness of predictions. For nongenotyped animals, the regression coefficients from BLUP and SSBLUP were similar to each other, except for LS5 in Landrace and Mort in Yorkshire, for which SSBLUP led to lower regression coefficients (larger inflation) than BLUP. For all animals in the validation data sets, which comprised mainly nongenotyped animals, the regression coefficients were similar to those for nongenotyped animals.

To test the effect of different weighting factors \( \omega \) in constructing \( G_\omega \) and \( H \) on genomic predictions, 9 values of \( \omega \) between 0.1 and 0.9 were used for the SSBLUP model. On average, reliabilities of EBV varied from 0.105 to 0.111 over the 9 scenarios for 3 traits in both populations. As shown in Fig. 1, the average reliability increased with the weight increasing from 0.1 to 0.4 and then decreased rapidly when the weight became larger than 0.6. The highest mean reliability was 0.111, obtained when using weight 0.5. The regression coefficients of genomic predictions ranged from 0.712 to 0.828. When the weight of \( A \) matrix increased from 0.1 to 0.9, the regression coefficient increased linearly to be closer to 1 (Fig. 2). However, when the weight was near 1, the regression coefficient dropped to a level similar to conventional BLUP (results not shown). As a trade-off between reliability and regression, the relative weight of 0.5 was appropriate for the present data.

### DISCUSSION

This study applied 4 methods, traditional BLUP, GBLUP, SELIND, and SSBLUP, for genetic evaluation of litter size and piglet mortality in Danish Landrace and Yorkshire populations. Predictive abilities of these 4 methods were compared in terms of reliability and regression coefficient. Results indicated that the methods with marker information were more accurate than the method based on only pedigree. In addition, the SSBLUP method generally provided more accurate predictions than the other genomic models for both genotyped and nongenotyped animals in Yorkshire and for nongenotyped animals in Landrace.

Comparison between the current study and the previous study in 2007 (Su et al., 2007) indicates that during the recent years, TNB has increased by 0.74 and LS5 has increased by 2.05 in Landrace and both litter size traits increased by 2.44 in Yorkshire. Moreover, in the current study, litter size is larger in Yorkshire, which is contrary to the results reported by Su et al. (2007). This is because the number of records is more concentrated in Yorkshire than in Landrace in recent years and litter size is increasing over the years. The

<table>
<thead>
<tr>
<th>Breed</th>
<th>Trait</th>
<th>All Genotyped</th>
<th>Genotyped</th>
<th>Nongenotyped</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BLUP</td>
<td>SSBLUP</td>
<td>BLUP</td>
</tr>
<tr>
<td>Landrace</td>
<td>TNB</td>
<td>0.856</td>
<td>0.762</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>0.877</td>
<td>0.756</td>
<td>0.589</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.749</td>
<td>0.756</td>
<td>1.266</td>
</tr>
<tr>
<td>Yorkshire</td>
<td>TNB</td>
<td>0.834</td>
<td>0.637</td>
<td>1.156</td>
</tr>
<tr>
<td></td>
<td>LS5</td>
<td>0.640</td>
<td>0.786</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td>Mort</td>
<td>0.874</td>
<td>0.757</td>
<td>0.850</td>
</tr>
</tbody>
</table>

1. TNB = total number of piglets born; LS5 = litter size at Day 5 after birth; Mort = mortality rate before Day 5 (including stillbirth).
2. SSBLUP = single-step BLUP.
3. GBLUP = genomic BLUP.
4. SELIND = selection index blending.

**Figure 1.** Reliability of genomic predictions obtained from the single-step method with different weights on pedigree-based relationship matrix, averaged over 3 traits and 2 breeds.
other reason for the difference could be that the results in Su et al. (2007) were based on the data of purebred litters in breeding farms, whereas the current study is based on litters from both purebred and crossbred litters in both breeding and multiplier farms. Comparing the results from the current study with the study by Su et al. (2007), the mortality rate decreased in both populations with an increase of both TNB and LS5. The trend of increasing litter size and decreasing piglet mortality during these years indicate that LS5 is an effective selection criterion to improve piglets alive.

The estimated heritabilities of 3 traits in the current study were around 0.1. In general, the estimates of heritability of TNB in both Landrace and Yorkshire were consistent with the average estimate (0.11) reported in the review by Rothschild and Bidanel (1998). The estimated heritabilities were also generally consistent with those in the study on the same populations by Nielsen et al. (2013) but higher than those reported by Su et al. (2007), where the estimated heritabilities of TNB and LS5 were 0.066 and 0.088 in Landrace and 0.053 and 0.070 in Yorkshire. In another study based on the data from breeding herds, Strathe et al. (2013) reported that heritabilities were 0.09 for TNB and 0.06 for LS5 in the Danish Landrace. Similarly, the current study showed that the estimated heritabilities of LS5 were smaller than those of TNB. Selection for LS5 during the last 10 yr may have caused a reduction of the genetic variation to some extent, which could partly explain that the genetic variation of LS5 is smaller than that of TNB. Mortality rate before d 5 is a ratio trait including TNB and LS5 and the estimates of heritability of Mort were similar to litter size traits. The estimated heritabilities of Mort in a previous study were 0.09 and 0.10 for the Danish Landrace and Yorkshire, respectively (Nielsen et al., 2013), which were similar to the results in current study.

In general, the ability of predicting litter size and piglet mortality was low. This could be because of the low heritability of these traits. Nevertheless, it has been reviewed that genomic prediction could be of special interest for increasing genetic gain for traits where the accuracy of selection is low, such as low heritability traits, if a sufficiently large number of genotypes and phenotypes are available (Ibañez-Escriche and Gonzalez-Recio, 2011). Lillehammer et al. (2011) reported that compared with conventional selection without progeny testing, the application of genomic information in pig breeding could improve genetic gain by increasing selection accuracy and reducing inbreeding rate. Based on reliabilities for the validation animals in current study, compared with traditional BLUP, SSBLUP method could improve the genetic gain by 6 to 18%.

In 1 study performed by Tusell et al. (2013), the predictive ability of several models for TNB was investigated. The average correlation between observed and predicted phenotypes in a 10-fold cross-validation was used to assess predictive ability. Similar to the current study, in most cases, the predictive ability of a genome-based model was higher than that of pedigree-based model. This was probably because of the fact that realized relationships among individuals are captured by marker information. However, not all the reliabilities of prediction using a genome-based model were higher than a pedigree-based model in the current study, depending on the phenotypic information used. For TNB in Yorkshire pigs, the predictive ability of the traditional BLUP model was somewhat higher than GBLUP for genotyped validation animals. This might be because GBLUP used the phenotypic information only from genotyped animals, whereas the traditional BLUP used phenotypic information of the whole population. In such a case, the superiority of a genome-based method to a pedigree-based method would be compromised by use of limited phenotypic information. Therefore, it is necessary to make the decision of selection based on the genomic information and all phenotypic information available. This can be realized using appropriate blending approaches or SSBLUP method.

For the genotyped animals, GBLUP used information of only genotyped animals, whereas the 2 blending approaches (SELIND and SSBLUP) used information of both genotyped and nongenotyped animals. In general, SELIND and SSBLUP provided more accurate predictions than GBLUP. However, for LS5 and Mort in Landrace, the blending approaches did not show superiority compared to GBLUP. The lower reliabilities of SELIND and SSBLUP than GBLUP could be partly explained by the low reliabilities of traditional BLUP for these 2 traits. However, differences between GBLUP, SELIND, and SSBLUP methods were not significant in prediction reliability for genotyped animals.
The SELIND and SSBLUP approaches used the same information sources but different algorithms. The reliabilities of predictions from the SSBLUP were generally higher than those from the SELIND. The SELIND method involved 2 steps. In the first step, \( P_{IT} \), \( P_{IT}^* \), and DGP, as well as their reliabilities, were estimated from different data sets and models. In the second step, these estimates were used to calculate GEBV. Therefore, the uncertainty from the first step was not taken into account in the following step (Christensen and Lund, 2010) whereas the core idea of a SSBLUP method is that information of genotyped and nongenotyped animals can be used simultaneously through the integration of the marker-based relationship matrix into the pedigree-based relationship matrix (Legarra et al., 2009; Misztal et al., 2009; Christensen and Lund, 2010). Clearly, the single-step method can avoid several assumptions and uncertainty of parameters required in multiple-step methods (Aguilar et al., 2010; Forni et al., 2011). This could be the reason why the SSBLUP, in general, produced more accurate predictions than the SELIND. In addition, the SSBLUP is easier to implement in practical genetic evaluations, and the nongenotyped candidates can also benefit from information of genotypes of genotyped animals.

Forni et al. (2011) compared different genomic relationship matrices for SSBLUP analysis of TNB. Estimates of accuracy obtained using PEV with different genomic matrices showed that the increase of accuracy from genomic information was only for genotyped animals (Forni et al., 2011). However, the current study showed that the accuracies of predictions for both genotyped and nongenotyped animals were enhanced when genomic information was added, although only a small part (about 3,000) of animals were genotyped. In a study by Christensen et al. (2012), average daily gain (ADG) and feed conversion ratio (FCR) in Danish Duroc pigs were analyzed using SSBLUP method. For both ADG and FCR, the accuracies of prediction were increased by using the SSBLUP method compared with the conventional BLUP for the whole validation population including genotyped and nongenotyped animals. The accuracies were higher than in the current study, which might be due to the difference between traits. Average daily gain and FCR have much higher heritability than litter size and piglet mortality traits. In addition, the study on ADG and FCR was based on data from breeding herds where animals are more homozygous and management are more consistent than those in the current study, which was based on data from both breeding and multiplier herds and included both purebred and crossbred litters. The use of data from crossbred litters could include a heterosis effect and create genotype × hybrid interaction. In current study, although the basic model included a fixed effect of litters being crossbred or not, it did not account for genotype × hybrid interaction. Alternative approaches could be to use a more sophisticated model that accounts for genotype × hybrid interaction or use a multitrait model that takes purebred litter and crossbred litter as 2 different traits. Regardless of these differences among traits and breeds, the superiority of SSBLUP has been observed in both the previous study mentioned (Christensen et al., 2012) as well as in the current study.

In the current study, validation reliabilities of genotyped animals were lower than those of nongenotyped animals in Landrace, although the better performance of SSBLUP can be observed for both genotyped and nongenotyped populations. The reason for lower reliabilities of genotyped animals could be due to the preselection of these individuals for LS5. The directional selection would reduce the correlation between observations and genomic predictions and, consequently, underestimate the reliabilities of genomic predictions (Su et al., 2012). In addition, the number of genotyped animals in the test data was relatively small, resulting in a large sampling error of the validation reliability.

The results from the present study showed that increasing the weighting factor from 0.1 to 0.9 reduces bias and that weighting factors around 0.5 gave the highest reliability in average over the traits and breeds. However, the optimal weighting factors differed between traits (results not shown). A possible reason for different optimal weights on different traits within a population could be that for different traits, there may be different proportions of variance that cannot be accounted for by markers. Among all the analyzed traits, there was 1 in each breed that did not obtain the most reliable prediction when using weight as 0.5. However, the differences between the most suitable weight and 0.5 were only at the third decimal in reliability and not significant in regression. Christensen et al. (2012) compared weights from 0.15 to 0.3 together with adjustment of G matrix for the same scale of A matrix in construction of the joint relationship matrix for the prediction using SSBLUP method in pig and reported that 0.25 was the best one. Another study in dairy cattle (Gao et al., 2012) compared weights ranging from 0.05 to 0.40 and found that weighting factors around 0.15 to 0.20 give the highest reliability. Similar to the current study, it also showed that the increasing of weighting factor could reduce bias.

In general, the regression coefficients of traditional BLUP and SSBLUP were similar for nongenotyped validation animals. However, for the genotyped validation animals, the differences between 4 methods were large in some cases, especially in Landrace. Two possi-
Genetic evaluation using a single-step method

A multitrait model might increase the accuracy for genomic evaluation by using information from the correlated traits. The gain from a multitrait model depends on the traits and data in the analysis. Su et al. (2014) reported a bivariate GBLUP model for claw health and longevity increased the reliability of genomic prediction for claw health, which had small number of phenotypic records, compared with a univariate GBLUP model. Furthermore, the use of the multiple-trait single-step model leads to a large increase in computational costs (Tsutsu et al., 2011).

Genetic parameters and breeding values for mortality or survival rates are often estimated by treating them as Gaussian traits in practical breeding programs. However, in general, the distribution of mortality rate is typically far from normal distribution and is skewed with a large proportion of litters showing 0. More appealing models for the analysis of mortality could be treating mortality in piglet individual level as a binary trait. Thus, genetic parameters and breeding values for piglet mortality can be estimated using a liability threshold model. Some previous studies applied the threshold model to estimate genetic parameters for piglet mortality, based on the record at individual piglet level (Arango et al., 2005; Su et al., 2008). To our knowledge, so far no study reported comparison between linear model and threshold model in terms of prediction reliability for piglet mortality. The disadvantage of a threshold model based on individual piglet record is a high computational demand in the cases of large data sets and might not easy to be implemented in routine genomic prediction.

Conclusions

The results from the current study indicate that the reliabilities of EBV for litter size and piglet mortality traits can be increased when using genomic information. Furthermore, a SSBLUP model that uses all information available simultaneously can improve prediction of breeding value not only for genotyped animals but also for nongenotyped animals. In general, the SSBLUP method could be a suitable approach for practical genomic prediction.

LITERATURE CITED


Madsen, P. 2010. DMU trace, A program to trace the pedigree for a sub set of animals from a larger pedigree file. Faculty Agricultural Sciences, University of Aarhus, Dept. of Genetics and Biotechnology, Research Centre Foulum, Denmark.