J. Dairy Sci. 93:1175–1183
doi:10.3168/jds.2009-2192

Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population

G. Su,[1] B. Guldbrandtsen, V. R. Gregersen, and M. S. Lund
Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Aarhus University, DK-8830, Tjele, Denmark

ABSTRACT

This study investigated the reliability of genomic estimated breeding values (GEBV) in the Danish Holstein population. The data in the analysis included 3,330 bulls with both published conventional EBV and single nucleotide polymorphism (SNP) markers. After data editing, 38,134 SNP markers were available. In the analysis, all SNP were fitted simultaneously as random effects in a Bayesian variable selection model, which allows heterogeneous variances for different SNP markers. The response variables were the official EBV. Direct GEBV were calculated as the sum of individual SNP effects. Initial analyses of 4 index traits were carried out to compare models with different intensities of shrinkage for SNP effects; that is, mixture prior distributions of scaling factors (standard deviation of SNP effects) assuming 5, 10, 20, or 50% of SNP having large effects and the others having very small or no effects, and a single prior distribution common for all SNP. It was found that, in general, the model with a common prior distribution of scaling factors had better predictive ability than any mixture prior models. Therefore, a common prior model was used to estimate SNP effects and breeding values for all 18 index traits. Reliability of GEBV was assessed by squared correlation between GEBV and conventional EBV ($r^2_{GEBV, EBV}$), and expected reliability was obtained from prediction error variance using a 5-fold cross validation. Squared correlations between GEBV and published EBV (without any adjustment) ranged from 0.252 to 0.700, with an average of 0.418. Expected reliabilities ranged from 0.494 to 0.733, with an average of 0.546. Averaged over 18 traits, $r^2_{GEBV, EBV}$ was 0.13 higher and expected reliability was 0.26 higher than reliability of conventional parent average. The results indicate that genomic selection can greatly improve the accuracy of preselection for young bulls compared with traditional selection based on parent average information.

Key words: cross validation, genomic estimated breeding value, genomic selection, reliability

INTRODUCTION

The application of molecular genetic information has become an important issue in animal breeding. In cattle, an assay for simultaneous genotyping of more than 50,000 SNP markers is commercially available. This opens an opportunity for effective selection using dense markers through the whole genome (i.e., genomic selection). Genomic selection is based on breeding values that are directly estimated from genome-wide dense marker panels. Therefore, genetic evaluation can be performed as soon as DNA is obtained, which allows accurate selection in both genders early in life. Genomic selection is expected to lead to considerably higher genetic gains than conventional quantitative genetic selection (Meuwissen et al., 2001; Schaeffer, 2006). It is expected that by using genomic selection in dairy cattle breeding, the genetic progress would be doubled whereas the cost for proving bulls would be reduced by 92% (Schaeffer, 2006).

Several statistical models and algorithms have been proposed to predict breeding values based on dense markers (Meuwissen et al., 2001; Xu, 2003; Meuwissen and Goddard, 2004; Gianola et al., 2006). Among the proposed methods, BLUP, BayesA, and BayesB have been widely used to analyze simulated data and real data. A linear BLUP approach (Meuwissen et al., 2001; VanRaden, 2008) assumes that effects of all SNP are normally distributed with same variance. BayesA and BayesB (Meuwissen et al., 2001) allow each marker to have its own variance of allele effects, and each variance is a sample of a scaled inverse chi-squared distribution. BayesB also models most SNP having zero effect, but a few having moderate to large effects. To simplify the computing algorithm in BayesA and BayesB (especially the Metropolis-Hastings step in BayesB), alternative Bayesian approaches similar to BayesA and BayesB have been proposed for genomic prediction (Meuwissen and Goddard, 2004; Villumsen et al., 2009). These approaches model SNP effects as a product of a scaled effect and a scaling factor (which can be understood

Received March 6, 2009.
Accepted December 3, 2009.
[1]Corresponding author: guosheng.su@agrsci.dk

as standard deviation of allele effects in a marker). It assumes that the prior distribution of scaling factors is either a normal distribution or a mixture of 2 normal distributions. Some simulation studies showed that the prediction ability of BayesA and BayesB was greater than BLUP approach, based on the simulated scenarios assuming that few QTL had a large effect and most QTL had a small effect (Meuwissen et al., 2001; Lund et al., 2009). Based on real data from dairy cattle, the Bayesian approaches or the analogous approaches [e.g., nonlinear BLUP in VanRaden (2008)] gave higher reliabilities than linear BLUP approach for the traits having known major genes, but the differences between these approaches were very small for the traits without major genes (Cole et al., 2009; Hayes et al., 2009; VanRaden et al., 2009).

So far, most reports on genomic selection in the literature were based on simulated data. Recently, many results based on the data from real livestock populations have been published (e.g., Harris et al., 2008; González-Recio et al., 2009; Hayes et al., 2009; VanRaden et al., 2009). However, to apply this new technology in practical breeding programs, it is necessary to evaluate the accuracy of genomic prediction in the target population. Therefore, the objectives of this study were to assess the predictive ability of models with different prior densities of marker effects and to investigate the reliability of genomic estimated breeding values for 18 traits based on the data from the Danish Holstein population.

MATERIALS AND METHODS

Data

Holstein bulls from 258 half-sib families (1–41 bulls each), born between 1986 and 2004, were genotyped using Illumina Bovine SNP50 BeadChip (Illumina, San Diego, CA). The marker data were edited using the following procedures: 1) deleted the locus with minor allele frequency less than 5%; 2) deleted the locus with average GenCall score less than 0.65; 3) deleted the individual with call rate score less than 0.85; and 4) for a marker with GenCall score less than 0.6 in an individual, set the marker as unknown in this individual. After the editing, there were 3,330 bulls and 38,134 SNP markers available. In the analysis of SNP effects and genomic prediction, any missing SNP at a particular marker in some animals was treated as an extra allele. This corresponded to replacing the effect of missing SNP at a marker with population mean of this marker.

Published conventional EBV were used as response variables to estimate SNP effects. The EBV and their reliability for the genotyped bulls were obtained from official evaluations in April 2009. In total, 18 index traits were analyzed in this study. Except for fat percentage and protein percentage, the traits are the subtraits in the new Nordic Total Merit index. Detailed descriptions of these index traits and their EBV are given in Danish Cattle Federation (2006).

Statistical Model

In this study, all individual SNP markers were used as predictors and conventional EBV were used as response variables weighted by a function of reliability of EBV (see detail later). A Bayesian method, which captures the features of BayesA and BayesB but simplifies the computing algorithm, was used to estimate marker effects for genomic prediction. The method applies the methodology of variable selection presented by George and McCulloch (1993). A detailed description of the method was presented by Villumsen et al. (2009) and Meuwissen and Goddard (2004). The following model was used to fit EBV data:

$$\mathbf{y} = 1\mu + \sum_{\mathrm{i=1}}^{\mathrm{m}} \mathbf{X}_i \mathbf{q}_i \nu_i + \mathbf{e},$$

where $\mathbf{y}$ is the vector of published conventional EBV, μ is the intercept, m is the number of SNP markers, $\mathbf{X}_i$ is the design matrix of allele types in marker i, $\mathbf{q}_i$ is the vector of scaled SNP effects (scaled by SD) of marker i with $\mathbf{q}_i \sim N(\mathbf{0}, \mathbf{I})$, v_i ($v_i > 0$) is a scaling factor (SD) for SNP effects of marker i, and $\mathbf{e}$ is the vector of residual with $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, where I is a identity matrix. The effects of SNP alleles of marker i are the products of v_i and $\mathbf{q}_i$.

Scaling factors v_i were assumed to have either a common prior distribution or a mixture prior distribution. A common prior distribution across the variances of chromosome segment effects, which leads to a slight or moderate differentiation between small and large effects of markers, was assumed to be a positive half-normal distribution (TN),

$$v_i \sim TN(0, \sigma_v^2),$$

where $v_i > 0$. Mixture prior distributions, which lead to strong differentiation between small and large effects of markers, assume that a proportion (π_0, typically large) of markers have very small effects, and another proportion ($\pi_1 = 1 - \pi_0$, typically small) of markers have moderate or large effects. This was achieved by assuming that the prior distribution of v_i was sampled from

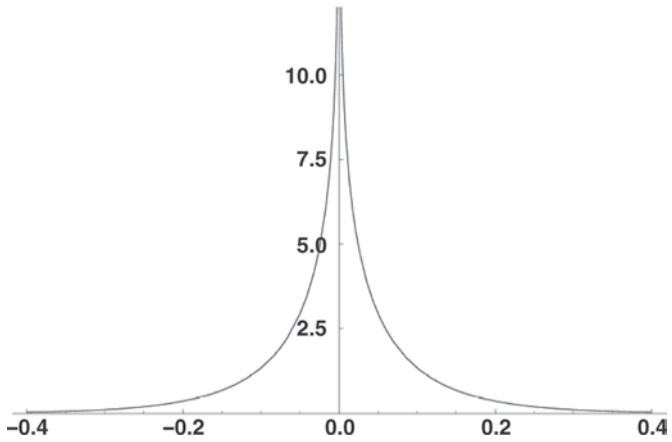

Figure 1. Prior distribution of q × v, where q ∼N(0, 1) and v ∼N(0, 0.01), and q is a variable of scaled SNP effect and v is a variable of scaling factor.

either a positive half-normal distribution with a small variance (σ^2_{v0}) or a positive half-normal distribution with large variance (σ^2_{v1}):

$$\nu_i \sim \pi_0 TN(0, \sigma^2_{v0}) + \pi_1 TN(\sigma^2_{v1}).$$

It is known that the probability density function for the product (z) of 2 independent normal random variables, $x \sim N(0, \sigma_x^2)$ and $y \sim N(0, \sigma_y^2)$, is

$$p(z) = \frac{1}{\pi\sigma_x\sigma_y} K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)$$

(Jagdish and Campbell, 1996), where K_0 is a modified Bessel function of the second kind. According to the features of normal distribution, the probability density function holds also for the product of a normal variable and a half normal variable. As shown in Figure 1, the prior distribution of q × v has a shape of a mass around zero and thicker tails, which allows for more extreme values of SNP effects.

The prior distributions of μ, σ^2_v, and σ^2_{v1} were assumed to be improper uniform distributions, whereas σ^2_{v0} was fixed at a small value. In this study, σ^2_{v0} was set to 0.0001 for all traits.

The genomic estimated breeding value (**GEBV**) for individual k was defined as the sum of predicted effects of SNP over all markers:

$$\text{GEBV}_\text{k} = \hat{\mu} + \sum_{\text{i=1}}^{\text{m}} \mathbf{x}_{\text{i(k)}} \hat{\mathbf{q}}_\text{i} \hat{\nu}_i.$$

The effect of prior density on accuracy of GEBV was investigated using 5 scenarios: 1) mixture prior to scaling factors with $\pi_1 = 5\%$, 2) $\pi_1 = 10\%$, 3) $\pi_1 = 20\%$, and 4) $\pi_1 = 50\%$, and 5) common prior to scaling factors for all markers (i.e., $\pi_1 = 100\%$). A low proportion (π_1) of markers with large effect indicated a high intensity of shrinkage for SNP effects.

A pilot study on fertility and udder health using a common prior model was carried out to investigate the effect of weighting factor of response variable (EBV) on accuracy of genomic prediction. A cross validation (see detail below) showed that squared correlations between GEBV and EBV were 0.412, 0.402, and 0.394 for fertility and 0.435, 0.423 and 0.418 for udder health using weighting factor of 1/(1 − reliability of EBV), reliability of EBV, and a constant weight of 1 for all response variables, respectively. Therefore, weighting factor of 1/(1 − reliability of EBV), which is inversely proportional to prediction error variance of EBV, was used in the current analysis. To avoid possible problems caused by extremely high weights, reliabilities of EBV larger than 0.98 were replaced with 0.98 in the calculation of weight.

Model Validation and Evaluation of Reliability of GEBV

The models with different priors for scaling factors and the accuracy of GEBV were evaluated using a 5-fold cross validation. In the cross validation, 134 half-sib families that have at least 1 bull born after 1993 were divided into 5 test datasets by the following procedures. First, the 134 half-sib families were assigned into 10 year classes (1994–2003) according to birth year for the most of half-sibs. Then, each 2 year classes formed a test dataset (i.e., 1994–1995 formed test set 1, 1996–1997 formed test set 2, and so on). The 5 test datasets comprised a total of 2,393 bulls. In each fold cross validation, the whole data excluded 1 test dataset to form a training dataset that was used to estimate marker effects and predict genomic breeding values of the left-out animals (the animals in the test data). Then, GEBV of the animals in test data were compared with their conventional EBV. The detailed information on the whole data and 5 test datasets is shown in Table 1.

Two criteria were used to assess accuracy of genomic prediction. The first criterion was the squared correlation between GEBV and published conventional EBV ($\mathbf{r^2_{GEBV,\ EBV}}$) in test datasets. In the present study, the total test dataset covered a period of more than 10 yr. Because of selection, mean breeding values changed among years, which would have an influence

Table 1. Structures of the whole dataset and 5 test datasets

Dataset	No. of bulls	No. of half-sib families	Interval of birth years
Whole	3,330	258	1986–2004
Test1	538	21	1989–1997
Test2	469	20	1994–2000
Test3	472	22	1997–2003
Test4	458	29	1999–2004
Test5	456	42	2001–2004

on $r^2_{GEBV,\ EBV}$. For the purpose of evaluating reliability of GEBV, both GEBV and EBV were adjusted for the year mean (subtracted by the mean of the year when the bull was born); that is, within-year squared correlation. The annual averages of GEBV and EBV were calculated from the animals in the total test dataset. The second criterion was the expected genomic reliability, obtained from prediction error variance (**PEV**). The PEV for a GEBV was measured as the variance of the posterior samples of this GEBV. To avoid strong influence of the dependency between test data and training data on the evaluation, the sires in test data having sons or grandsons in the training data were excluded from the validation.

Five scenarios of prior distribution for scaling factors (SD, v_i) of SNP effects were evaluated by analyzing 4 index traits (protein, fat percentage, udder health, and female fertility). Model predictive ability was assessed by $r^2_{GEBV,EBV}$ in the 5-fold cross validation. The best model (which was a common prior distribution in this study) was used to analyze all the 18 index traits.

The analyses were carried out using the IBAY package (Janss Luc, Faculty of Agricultural Sciences, Aarhus University, Tjele, Denmark). The Gibbs sampler was run as a single chain with a length of 50,000 samples. Convergence was monitored by graphical inspection of the variance of scaling factors and the correlation between GEBV from 2 separate rounds. The first 20,000 samples were discarded as burn-in. Every tenth sample of the remaining 30,000 was saved to estimate the features of the realized posterior distributions.

RESULTS

Table 2 shows the mean, across-year standard deviation, and within-year standard deviation of the published EBV and their reliabilities for the genotyped bulls, and the range of heritabilities for the component traits of each complex trait. The heritabilities were provided by Nordic Genetic Evaluation (Danish Agricultural Advisory Service, Aarhus, Denmark). The published EBV were standardized to a mean of 100 for the cows born 3 to 5 yr (for production and conformation traits, animal model) or for the bulls born 7 to 9 yr (for the remaining traits, sire model) before publication, and standardized to a standard deviation of 10 for bulls born in 1997 and 1998. The across-year standard deviations for yield index, protein, milk, fertility, and other–disease were higher than 10, reflecting a genetic change over the years for these traits. Within-year standard deviations were close to 10 for all traits except for longevity (8.6) and growth (11.5), indicating that the genotyped bulls

Table 2. Mean, cross-year standard deviation $\left(\sigma_t^2\right)$, and within-year standard deviation $\left(\sigma_w^2\right)$ of EBV, reliability of EBV (REL_{EBV}) for the genotyped bulls, and the range of heritabilities (h^2) for the component traits of each complex trait

	EBV				
Trait	Mean	σ_t^2	σ_w^2	REL_{EBV}	h^2
Birth index	100.9	10.4	10.2	75.7	0.01–0.20
Body conformation	97.3	10.9	10.2	81.1	0.10–0.61
Calving index	99.5	10.1	9.9	70.5	0.01–0.07
Fat	98.8	11.2	9.4	93.4	0.29–0.36
Fat percentage	100.0	10.4	10.4	93.5	0.50
Fertility	104.6	10.9	9.9	69.1	0.01–0.04
Other–disease	102.0	10.6	9.9	61.2	0.01–0.04
Feet–legs	98.8	10.0	9.9	62.5	0.10–0.28
Longevity	100.4	8.5	8.4	61.7	0.12
Milk	98.9	12.6	10.6	93.4	0.27–0.43
Udder conformation	97.2	10.1	9.5	77.9	0.12–0.41
Milking ability	99.5	10.9	10.7	72.8	0.10–0.25
Protein	97.1	14.0	10.4	93.4	0.25–0.35
Protein percentage	98.2	10.3	10.2	93.5	0.50
Temperament	100.3	9.8	9.4	63.5	0.15
Udder health	101.4	10.3	10.1	75.9	0.04–0.05
Yield index	97.1	13.5	10.0	93.4	0.25–0.43
Growth	100.5	11.6	11.5	87.9	0.16–0.29
Average	99.6	10.9	10.0	78.9	

represented the genetic variation of bulls in the population. Reliabilities of EBV differed among 18 traits and were consistent with heritabilities of the traits.

Influence of Prior Distribution on Genomic Prediction

The effect of changing the prior distribution of scaling factors on the predictive ability was investigated on fertility, protein, udder health, and fat percentage. The predictive abilities of the models with different priors for scaling factors were evaluated by $r^2_{GEBV,\ EBV}$ based on a 5-fold cross validation. Table 3 shows that there was a clear trend that $r^2_{GEBV,EBV}$ increased with increasing prior proportion (π_1) of SNP, with large effects within each subsets. Pooled over 5 subsets, $r^2_{GEBV,EBV}$ increased from 0.347 ($\pi_1 = 0.05$) to 0.412 (common prior; $\pi_1 = 1.0$) for fertility, from 0.279 to 0.412 for protein, from 0.338 to 0.435 for udder health, and from 0.670 to 0.700 for fat percentage. The $r^2_{GEBV,EBV}$ for fertility were similar when using models with $\pi_1 = 0.50$ and common prior; for fat percentage, the values were similar when using models with $\pi_1 = 0.20$, $\pi_1 = 0.50$, and common prior. It was found that variation of $r^2_{GEBV,\ EBV}$ among the 5 subsets was larger in fertility and udder health (the traits having low heritability) than protein and fat percentage (high heritability). As shown in Table 3, average reliability of EBV for bulls in the test dataset 5 was lower than those in the other test datasets. However, no clear association between $r^2_{GEBV,\ EBV}$ and average reliability of EBV was observed, probably because of small variation in average reliability of EBV between these datasets.

It was found that the prior distribution of scaling factors had a considerable influence on the estimates of marker effects. Taking fertility as an example, the marker effects (expressed as absolute value of the difference between 2 allele effects) followed a Gamma distribution for all scenarios (Figure 2). With a prior that assumes a lower proportion (π_1) of markers with large effect, the distribution becomes more L-shaped. The posterior percentages of the markers with an estimated effect less than 0.005 were 91, 84, 70, 41, and 33%, and the maximal values of estimated effects were 1.646, 1.390, 0.581, 0.181, and 0.120 for $\pi_1 = 0.05$, 0.10, 0.20, and 0.50 and a common prior, respectively.

Table 3. Reliability of EBV (REL_{EBV}) and squared correlation between genomic EBV and EBV ($r^2_{GEBV,\ EBV}$) for bulls in each test dataset of a cross validation[1]

			$r^2_{GEBV,EBV}$				
Trait	Dataset	REL_{EBV}	Mixture[2] $\pi_1 = 5\%$	Mixture $\pi_1 = 10\%$	Mixture $\pi_1 = 20\%$	Mixture $\pi_1 = 50\%$	Common[3]
Fertility	Test1	70.0	0.275	0.304	0.314	0.342	0.362
	Test2	68.2	0.348	0.378	0.389	0.388	0.399
	Test3	68.9	0.300	0.340	0.359	0.374	0.376
	Test4	67.4	0.416	0.405	0.434	0.441	0.444
	Test5	63.0	0.419	0.444	0.438	0.495	0.493
	Pooled	67.6	0.347	0.370	0.384	0.407	0.412
Protein	Test1	93.8	0.284	0.315	0.357	0.393	0.401
	Test2	93.2	0.304	0.363	0.405	0.371	0.413
	Test3	93.6	0.283	0.331	0.354	0.374	0.375
	Test4	93.1	0.283	0.352	0.392	0.407	0.438
	Test5	92.0	0.233	0.309	0.368	0.410	0.420
	Pooled	93.1	0.279	0.337	0.378	0.394	0.412
Udder health	Test1	76.1	0.279	0.301	0.330	0.332	0.351
	Test2	75.7	0.275	0.317	0.369	0.377	0.410
	Test3	76.5	0.415	0.448	0.481	0.498	0.505
	Test4	75.3	0.372	0.395	0.395	0.421	0.431
	Test5	71.4	0.322	0.381	0.433	0.464	0.466
	Pooled	75.0	0.338	0.373	0.404	0.417	0.435
Fat percentage	Test1	93.9	0.681	0.709	0.725	0.711	0.716
	Test2	93.2	0.662	0.678	0.694	0.694	0.685
	Test3	93.5	0.709	0.729	0.748	0.741	0.751
	Test4	92.8	0.695	0.705	0.714	0.708	0.703
	Test5	92.0	0.591	0.611	0.622	0.632	0.640
	Pooled	93.1	0.670	0.688	0.702	0.698	0.700

[1]GEBV were predicted using models with different prior distributions of scaling factors.

[2]Mixture: prior distribution of scaling factors was a mixture of two positive half-normal distributions with probability π_1 and $(1 - \pi_1)$, respectively.

[3]Common: prior distribution of scaling factors was a positive half-normal distribution common to all markers.

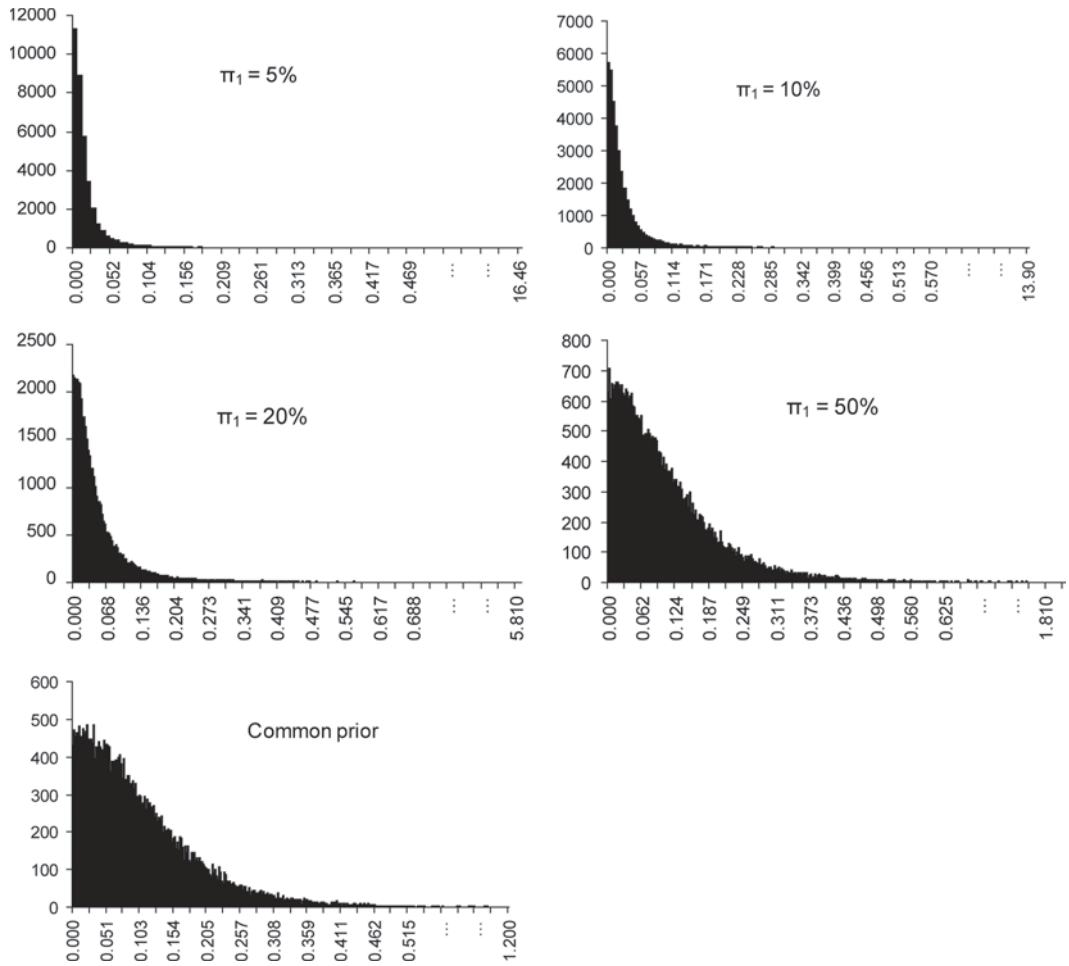

Figure 2. Distributions of SNP effects for fertility estimated from models with different priors for SNP variance. Y-axis: frequency; x-axis: absolute SNP effect (×10).

Reliabilities of GEBV

Because models with a common prior distribution of scaling factors generally provided better predictive abilities than mixture prior models, this model was chosen to estimate SNP effects and predict breeding values for 18 traits in the Danish Holstein population. Table 4 presents $r^2_{GEBV,\ EBV}$ and expected reliability of GEBV calculated from PEV for bulls in the test data as well as reliability of conventional parent average (**PA**) at the time of birth of the bull calves in this population. The reliabilities of parent average were provided by Nordic Genetic Evaluation (Danish Agricultural Advisory Service, Aarhus, Denmark). The $r^2_{GEBV,\ EBV}$ ranged from 0.25 to 0.70, which were 0.01 to 0.25 greater than reliabilities of PA, except for body conformation, for which $r^2_{GEBV,\ EBV}$ was 0.05 lower than reliability of PA. Expected reliabilities ranged from 0.49 to 0.73, corresponding 0.13 to 0.41 higher than reliabilities of PA. Averaged over 18 traits, $r^2_{GEBV,\ EBV}$ was 0.13 higher (0.42 vs. 0.29) and expected reliability of GEBV was 0.26 higher (0.55 vs. 0.29) than reliability of PA.

It was observed that the variation in $r^2_{GEBV,\ EBV}$ among the 18 traits was larger than the variation in expected reliabilities, but the patterns of ranks were similar. Product moment correlation and rank correlation between the 2 parameters were 0.883 and 0.813, respectively. Although the heritabilities for these traits differed considerably, the difference in $r^2_{GEBV,\ EBV}$ or expected reliability between low-heritability traits and high-heritability traits were relatively small, indicating that the reliability of GEBV was not very strongly influenced by heritability. For example, fertility, feet–legs, udder health, and other–diseases had an expected reliability of GEBV and an $r^2_{GEBV,EBV}$ as high as or close to those for production traits.

Table 4. Reliability of parent average (REL_{PA}) at the time of birth of the bull calves in Danish Holstein population, squared correlation between EBV and genomic EBV (GEBV; $r^2_{GEBV,EBV}$), and expected reliability of GEBV (calculated from prediction error variance) for bulls in the test data[1]

Trait	REL_{PA}	$r^2_{GEBV,EBV}$	Expected reliability
Birth index	0.26	0.395	0.502
Body conformation	0.30	0.252	0.499
Calving index	0.22	0.369	0.525
Fat	0.38	0.487	0.569
Fat percentage	0.38	0.700	0.733
Fertility	0.22	0.412	0.566
Other–disease	0.18	0.426	0.593
Feet–legs	0.26	0.404	0.571
Longevity	0.21	0.317	0.494
Milk	0.29	0.481	0.562
Udder conformation	0.33	0.395	0.511
Milking ability	0.25	0.383	0.506
Protein	0.38	0.412	0.528
Protein percentage	0.38	0.518	0.559
Temperament	0.25	0.340	0.514
Udder health	0.25	0.435	0.557
Yield index	0.38	0.390	0.514
Growth	0.28	0.415	0.533
Average	0.29	0.418	0.546

[1]GEBV were predicted using a common prior model.

DISCUSSION

Reliability of GEBV is a critical criterion in deciding whether GEBV should be used for routine genetic evaluation. In this study, 5 scenarios of prior distribution for variance of SNP effects were assessed. The common prior model performed generally better than the mixture prior models; therefore, this model was used to investigate reliability of GEBV for 18 traits in the Danish Holstein population. In general, accuracy of GEBV was considerably higher than conventional PA.

In this study, reliability of GEBV was evaluated by $r^2_{GEBV,\ EBV}$ using a cross validation and expected reliability (calculated from PEV). It was found that $r^2_{GEBV,EBV}$ were lower than the expected reliabilities. The lower $r^2_{GEBV,EBV}$ could be caused by the facts that EBV contained error and the animals in the validation were selected from elite parents instead of random samples (VanRaden et al., 2009). On the other hand, it is also possible that the expected reliability may overestimate the reliability of GEBV. An alternative is to measure reliability of GEBV as $r^2_{GEBV,EBV}$ divided by reliability of EBV. This is to assume the correlation between GEBV and EBV was through their correlation with true breeding value (i.e., no correlation between prediction errors of GEBV and EBV). Thus, $r_{GEBV,EBV} = r_{GEBV,TBV} \times r_{EBV,TBV}$ and $r^2_{GEBV,TBV} = r^2_{GEBV,EBV}/r^2_{EBV,TBV}$, where TBV is true breeding value. However, based on the present data, reliability estimated using this approach seemed too high to be acceptable for some low-heritability traits (results not shown), implying prediction errors of GEBV and EBV were not completely independent. We suggest that the true reliability of GEBV ($r^2_{GEBV,TBV}$) in the present data could be between $r^2_{GEBV,EBV}$ and the expected reliability. Thus, averaged over 18 traits, reliability of GEBV could be in the interval between 42 and 55%. The figures are considerably greater than the reliability of the conventional PA. It indicates that genomic prediction can effectively improve the accuracy of preselection for young bulls compared with traditional selection based on PA.

Few reports exist on the reliability of GEBV based on data from real livestock. Hayes et al. (2009) reported that reliabilities of GEBV ranged from 18 to 53% for 5 traits in Australian Holstein population, which were 2 to 18% higher than reliabilities of PA. VanRaden et al. (2009) investigated reliability of GEBV for 27 traits in North American Holstein population and reported that reliabilities of the index combining GEBV and PA ranged from 35 to 78%, which, averaged over the traits, was 23% higher than reliability of PA (50 vs. 27%). Harris et al. (2008) reported that reliabilities of GEBV combining PA for 12 traits ranged from 45 to 60%, with an average of 53%, in New Zealand Holstein population. In the Netherlands, de Roos et al. (2009) reported that reliabilities of the index combining GEBV and national EBV for 12 traits of Holstein bulls were in the range of 27 to 68%, with an average of 51%. In Ireland, Berry et al. (2009) reported that reliability of the genetic evaluation blending GEBV and national EBV for 21 traits of Holstein bulls increased by 1% to 18% compared with reliabilities of PA. It is difficult to compare the reliabilities from different reports because sizes of reference data and methods to calculate reliability are different in different studies.

The difference in reliability of GEBV between low-heritability traits and high-heritability traits was relatively small. In the present study, marker effects were estimated from published EBV. The influence of heritability on GEBV was through its influence on reliability of EBV. However, the published EBV were predicted from a very large dataset, resulting in a relatively high accuracy even for the traits with low heritability. Moreover, in genomic prediction, each individual in reference data has a contribution to marker effects. In other word, the GEBV of a candidate is actually obtained from the information of all individuals in the reference data. The benefit from information of other animals for the traits with low heritability is relatively greater than that for the traits with high heritability. The weak dependency on heritability indicates that genetic evaluation based on GEBV would be relatively more beneficial for the traits with low heritability. Previous studies on marker-assisted selection have shown that gain in

response rate is larger for traits with lower heritability (Lande and Thompson, 1990; Meuwissen and Goddard, 1996). However, these calculations were conditional on the fact that QTL had been identified, which is much more difficult for low-heritability traits because of low statistical power of detection. Using genomic selection, the step of testing for QTL is circumvented. This is a reason that accurate GEBV can be obtained even for low-heritability traits. As a consequence of a relatively weak dependency of GEBV on heritability, it becomes relatively easier to improve functional traits and to obtain a balanced genetic progress between functional traits and production traits, compared with selection on conventional EBV.

Five scenarios of prior distributions of the variance of SNP effects were investigated in this study. It was found that, using single markers as explanatory variables, the model with a common distribution of scaling factors (SD) generally had better predictive ability than models assuming a mixture distribution. Similarly, Cole et al. (2009) reported that a heavy-tailed prior model (analogous to the common prior model in the present study) gave slightly higher reliability of GEBV than a finite locus model with heavy tails (analogous to the mixture prior model). VanRaden et al. (2009) reported that predictive ability of a nonlinear BLUP model (a heavy-tailed prior model) was considerably better than a linear BLUP model for fat percentage and protein percentage, whereas the predictive abilities were similar for 25 other traits. In simulation studies, Meuwissen et al. (2001) reported that the accuracy of GEBV using BayesB (similar to the mixture prior model in the present study) was higher than that using BayesA (common prior distribution), and Lund et al. (2009) found that mixture models predicted breeding value better than the models with a common prior distribution of variances or the models with equal variance for all SNP. Both studies were based on the data in which QTL effects were simulated from a Gamma distribution with shape parameter 0.4 (L shape).

There are many possible reasons why the models with a mixture prior distribution of scaling factors did not perform better than the model with a common prior distribution in the present data. First, the mixture prior distribution of scaling factors is based on the hypothesis that few genes have a large effect and a large number of genes have a small effect, and the distribution of QTL effects follows a Gamma distribution of L shape. The hypothesis is supported by the derived distribution of QTL effects reported by Hayes and Goddard (2001). However, the distribution of SNP effects is not necessary to be consistent with the distribution of QTL effects. Many SNP could be located in a chromosome segment with large effect; thus, the effect of the chromosome segment could be divided over many SNP (Hayes et al., 2009). On the other hand, effect of a QTL might not be fully accounted for by a single marker because of incomplete linkage between marker and QTL. A mixture prior model may lead to too many SNP with an effect regressed to zero.

Second, the data available for this study (about 2,850 individuals in the training data) may not be sufficient for satisfactory prediction using the mixture model. In the present study, the models with a mixture of 2 prior distributions distinguish the markers with small effect and the markers with large effect intensively by setting different variances for scaling factors. The variance of scaling factors for the marker with large effects was sampled from the conditional posterior distribution, whereas the variance of scaling factors for the markers with small effects was given a value of 0.0001. Both the given variance and the proportions may not be optimal. With insufficient data, the priors have strong influence on the estimates. This was observed in this study, where the distributions of SNP effects were greatly dependent on the given proportions of markers with large effects (Figure 2). Inappropriate priors could reduce the accuracy of the estimates of marker effects. However, it is difficult to find an optimal proportion and variance of scaling factors for the markers with small effect.

The accuracy of GEBV was evaluated using a 5-fold cross validation. The advantage of multiple-fold cross validation is that it can retain training data as large as possible while keeping the test data as large as required. In the cross validation, each set of training data left many half-sib families out instead of leaving a random sample out. This strategy greatly reduces the dependency between the training data and the test data because the individuals in the test data did not have their sibs in the training data.

In this study, marker effects were estimated by fitting a model to published EBV. The advantage of using EBV is that they can be obtained directly from routine genetic evaluations. In addition, they contain little random error, which greatly reduces the prediction error variance. This could be important in situations where the number of genotyped animals in the reference data is small. An alternative type of response variable is daughter yield deviation. Further studies should be carried out to investigate the effect of different response variables on reliability of GEBV in different situations (e.g., different heritabilities and sizes of reference data).

The reliabilities of GEBV in this study indicate that genomic selection is promising. Moreover, genomic prediction can be further improved by several approaches. First, reliability of GEBV can increase with increasing data size (the number of individuals with

both genotypes and phenotypes) to estimate marker effects. Second, the reliability could be improved by using more sophisticated models. Third, the reliability of GEBV for an index trait is expected to be improved by predicting genomic breeding value for each single trait and then calculating the GEBV of the index trait instead of predicting the index trait directly. Finally, higher accuracy of genomic selection can be obtained by a genomic selection index that combines GEBV and other sources of information, such as parent EBV from conventional national genetic evaluation (VanRaden et al., 2009).

CONCLUSIONS

Averaged over all 18 index traits, $r^2_{GEBV, EBV}$ from a cross validation was 42% and expected reliability of GEBV was 55%, considerably greater than the reliability of conventional PA (29%). It indicates that genomic selection can greatly improve the accuracy of preselection for young bulls compared with traditional selection based on PA. Therefore, genomic prediction has been used in Nordic genetic evaluation of young candidates. Based on the data in this study, it was found that the model with a common prior distribution of scaling factors had generally better predictive ability than those models with a mixture prior distribution.

ACKNOWLEDGMENTS

We thank Danish Cattle Federation (Aarhus, Denmark), Swedish Dairy Association (Stockholm, Sweden), and Nordic Cattle Genetic Evaluation (Aarhus, Denmark) for providing phenotypic data. This work was performed in the project "Genomic selection—From function to efficient utilization in cattle breeding (grant no. 3412-08-02253)," funded by the Danish Directorate for Food, Fisheries and Agri Business (Copenhagen, Denmark), VikingGenetics (Aarhus, Denmark), Nordic Genetic Evaluation (Aarhus, Denmark), and Aarhus University (Aarhus, Denmark).

REFERENCES

Berry, D. P., F. Kearney, and B. F. Harris. 2009. Genomic selection in Ireland. Pages 29–34 in Proc. Interbull Workshop of Genomic Information in Genetic Evaluations, Uppsala, Sweden. Interbull Bulletin 39. Interbull, Uppsala, Sweden.

Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J. Dairy Sci. 92:2931–2946.

Danish Cattle Federation. 2006. Principles of Danish Cattle Breeding. 8th ed. The Danish Agricultural Advisory Centre, Aarhus, Denmark. http://www.landbrugsinfo.dk/Kvaeg/Avl/Sider/principles.pdf Accessed May 15, 2009.

de Roos, A. P. W., C. Schrooten, E. Mullaart, S. Van der Beek, G. De Jong, and W. Voskamp. 2009. Genomic selection at CRV. Pages 47–50 in Proc. Interbull Workshop of Genomic Information in Genetic Evaluations, Uppsala, Sweden. Interbull Bulletin 39. Interbull, Uppsala, Sweden.

George, E. I., and R. E. McCulloch. 1993. Variable selection via Gibbs sampling. J. Am. Stat. Assoc. 88:881–889.

Gianola, D., R. Fernando, and A. Stella. 2006. Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173:1761–1776.

González-Recio, O., D. Gianola, G. J. Rosa, K. A. Weigel, and A. Kranis. 2009. Genome-assisted prediction of a quantitative trait measured in parents and progeny: Application to food conversion rate in chickens. Genet. Sel. Evol. 41:3. doi:10.1186/1297-9686-41-3

Harris, B. L., D. L. Johnson, and R. J. Spelman. 2008. Genomic selection in New Zealand and the implications for national genetic evaluation. Proc. Interbull Meeting, Niagara Falls, NY. Interbull, Uppsala, Sweden.

Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433–443.

Hayes, B., and M. E. Goddard. 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33:209–229.

Jagdish, K. P., and B. R. Campbell. 1996. Handbook of the Normal Distribution. 2nd ed. Marcel Dekker Inc., New York, NY.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756.

Lund, M. S., G. Sahana, D.-J. de Koning, G. Su, and Ö. Carlborg. 2009. Comparison of analysis of the QTLMAS XII common dataset. I. Genomic selection. BMC Proc. 2009, 3(Suppl. 1):S1.

Meuwissen, T. H. E., and M. E. Goddard. 1996. The use of marker haplotypes in animal breeding schemes. Genet. Sel. Evol. 28:161–176.

Meuwissen, T. H. E., and M. E. Goddard. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genet. Sel. Evol. 36:261–279.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123:218–223.

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

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. *Invited review:* Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

Villumsen, T. M., L. Janss, and M. S. Lund. 2009. The importance of haplotype length and heritability using genomic selection in dairy cattle. J. Anim. Breed. Genet. 126:3–13.

Xu, S. 2003. Estimating polygenic effects using markers of the entire genome. Genetics 163:789–801.