



J. Dairy Sci. 96:4678–4687

<http://dx.doi.org/10.3168/jds.2012-6406>© American Dairy Science Association®, 2013. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Model comparison on genomic predictions using high-density markers for different groups of bulls in the Nordic Holstein population

H. Gao,*†¹ G. Su,*¹ L. Janss,* Y. Zhang,† and M. S. Lund*

*Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark

†College of Animal Science and Technology, China Agricultural University, 100193 Beijing, P. R. China

ABSTRACT

This study compared genomic predictions based on imputed high-density markers (~777,000) in the Nordic Holstein population using a genomic BLUP (GBLUP) model, 4 Bayesian exponential power models with different shape parameters (0.3, 0.5, 0.8, and 1.0) for the exponential power distribution, and a Bayesian mixture model (a mixture of 4 normal distributions). Direct genomic values (DGV) were estimated for milk yield, fat yield, protein yield, fertility, and mastitis, using deregressed proofs (DRP) as response variable. The validation animals were split into 4 groups according to their genetic relationship with the training population. Group_{sms} had both the sire and the maternal grandsire (MGS), Group_{sire} only had the sire, Group_{mgs} only had the MGS, and Group_{non} had neither the sire nor the MGS in the training population. Reliability of DGV was measured as the squared correlation between DGV and DRP divided by the reliability of DRP for the bulls in validation data set. Unbiasedness of DGV was measured as the regression of DRP on DGV. The results showed that DGV were more accurate and less biased for animals that were more related to the training population. In general, the Bayesian mixture model and the exponential power model with shape parameter of 0.30 led to higher reliability of DGV than did the other models. The differences between reliabilities of DGV from the Bayesian models and the GBLUP model were statistically significant for some traits. We observed a tendency that the superiority of the Bayesian models over the GBLUP model was more profound for the groups having weaker relationships with training population. Averaged over the 5 traits, the Bayesian mixture model improved the reliability of DGV by 2.0 percentage points for Group_{sms}, 2.7 percentage points for Group_{sire}, 3.3 percentage points for Group_{mgs}, and 4.3 percentage points for Group_{non} compared with GBLUP. The results indicated that a Bayesian model with

intense shrinkage of the explanatory variable, such as the Bayesian mixture model and the Bayesian exponential power model with shape parameter of 0.30, can improve genomic predictions using high-density markers.

Key words: genomic prediction, reliability, high-density marker, genetic relationship

INTRODUCTION

Many factors influence the accuracy of genomic prediction, one of the crucial factors being marker density (Solberg et al., 2008; Habier et al., 2009; Harris and Johnson, 2010). It is expected that the reliability of genomic predictions will be greatly improved using high-density (HD) SNP markers because of stronger linkage disequilibrium (LD) between the SNP markers and the QTL affecting the traits of interest (Solberg et al., 2008; Meuwissen and Goddard, 2010). However, a recent study on genomic predictions in Nordic Holstein and Red populations using BLUP methods only showed a small improvement when using ~777,000 (777K) SNP markers, compared with using ~54,000 (54K) SNP markers (Su et al., 2012a). The authors argued that more sophisticated variable selection methods and models were required to exploit the potential advantage of HD markers for genomic prediction.

When using medium-density SNP chips (e.g., 54K), many studies have shown that a linear model assuming that effects of all SNP are normally distributed with equal variance performs as well as variable selection models for most traits in dairy cattle (Hayes et al., 2009a; VanRaden et al., 2009). Therefore, such linear models (genomic BLUP, **GBLUP**) have been used by many countries as the routine genomic evaluation models because of their simplicity and low computational requirement. For high-density SNP chips, it is uncertain if such GBLUP models can take full advantage of the LD information (Meuwissen and Goddard, 2010). Therefore, it is important to compare different models for genomic prediction using HD markers.

Breeding values can be accurately predicted using genome-wide dense markers, in part due to LD between markers and all QTL affecting the trait, and in part because markers capture genetic relationships among

Received November 22, 2012.

Accepted March 22, 2013.

¹Corresponding authors: guosheng.su@agrsci.dk and hongdinggao@gmail.com

Table 1. Heritability of the traits and number of bulls in training and validation data sets

Trait	h ²	Training	Validation
Milk	0.39	2,987	1,395
Fat	0.39	2,987	1,395
Protein	0.39	2,987	1,395
Fertility	0.04	3,021	1,378
Mastitis	0.04	2,990	1,461

genotyped animals (Habier et al., 2007). In general, genomic predictions are more accurate for animals having closer relationships with the training population (Lund et al., 2009; Meuwissen, 2009; Habier et al., 2010). However, the contribution of LD and relationship information to accuracy of genomic predictions may not be the same when using different models. It can be hypothesized that predictions from models that better capture LD between markers and QTL also persist better when genetic relationships get weaker. The advantage of one model over another model would thereby depend on the relationship between the predicted animals and the training population. This would be more profound when using HD markers, because of stronger LD between markers and QTLs.

The objective of this study was to compare a GBLUP model and 5 Bayesian shrinkage and variable selection models on the accuracy of genomic predictions using HD markers. The comparison was carried out for different groups of animals with varying degrees of close relationship with the animals in the training data set in the Nordic Holstein population.

MATERIALS AND METHODS

Data

The data used in this study consisted of 4,539 genotyped Nordic Holstein bulls born between 1974 and 2008. The bulls were divided into a training population and a validation population by birth date of October 1, 2001. Five traits (sub-indices) in the Nordic Total Merit index were analyzed: milk yield, fat yield, protein yield, fertility, and mastitis. The numbers of bulls in the training and validation data sets varied over traits

and are shown in Table 1. For bulls in the validation data set, 4 groups were constructed: (1) bulls that had both sire and maternal grandsire (**MGS**) in the training data set (**Group_{smgs}**); (2) bulls that had sire but no MGS in training data set (**Group_{sire}**); (3) bulls that had MGS but no sire in training data set (**Group_{mgs}**); and (4) bulls that had neither sire nor MGS in the training data set (**Group_{non}**). To balance numbers among these 4 groups, 16 bulls were removed from the training data set; the numbers of bulls in each group before and after removing the 16 bulls are presented in Table 2. Although Group_{mgs} and Group_{non} did not have the sire in the training data set, 177 bulls in Group_{mgs} and 191 bulls in Group_{non} had the paternal grandsire in the training data set.

The bulls were genotyped using the Illumina Bovine SNP50 BeadChip (Illumina Inc., San Diego, CA). In total, 557 bulls in the EuroGenomics project (Lund et al., 2011) were re-genotyped using the Illumina BovineHD BeadChip (777K). Among the 557 bulls, 161 bulls appeared in the training data and 16 bulls were in the validation data. The marker data of the bulls genotyped using the 54K chip were imputed to the HD genotypes applying Beagle package (Browning and Browning, 2009) and using the 557 HD genotyped bulls as reference. Detailed description of the imputed HD markers can be found in Su et al. (2012a).

A total of 14,588 progeny-tested bulls and 42,144 individuals in the pedigree were used to derive the deregressed proofs (**DRP**), which were used as the pseudo phenotype data in this study. The deregression procedure was implemented by using the iterative method described in (Jairath et al., 1998; Schaeffer, 2001) using the MiX99 package (Strandén and Mäntysaari, 2010) and with the heritabilities presented in Table 1, which were supplied by Nordic cattle routine genetic evaluation (<http://www.nordicebv.info/Routine+evaluation/>).

Statistical Models

The statistical models used in this study were a GB-LUP model, 4 Bayesian exponential power (**EPOW**) models, and a Bayesian mixture model.

Table 2. Number of bulls for each group in the validation data set before and after removing 16 bulls from the training data set¹

Item	Group _{smgs}	Group _{sire}	Group _{mgs}	Group _{non}
Before removing	902	218	213	62
After removing	344	351	347	353

¹Group_{smgs}: bulls had both sire and maternal grandsire (MGS) in training data set; Group_{sire}: bulls had sire but no MGS in training data set; Group_{mgs}: bulls had MGS but no sire in training data set; Group_{non}: bulls had neither sire nor MGS in training data set.

GBLUP Model

The GBLUP model (VanRaden, 2008; Hayes et al., 2009b) used to predict direct genomic breeding value (DGV) was as follows:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{g} + \mathbf{e}, \quad [1]$$

where \mathbf{y} is the data vector of DRP of genotyped bulls, $\mathbf{1}$ is a vector of ones, μ is the overall mean, \mathbf{Z} is a design matrix allocating records to breeding values, \mathbf{g} is a vector of genomic breeding values to be estimated, and \mathbf{e} is the vector of residuals. Using the GBLUP model, the estimate of (\hat{g}_i) was taken as the DGV of animal i .

It is assumed that $\mathbf{g} \sim N(\mathbf{0}, \sigma_g^2)$, where σ_g^2 is the additive genetic variance, and \mathbf{G} is the marker-based genomic relationship matrix (VanRaden, 2008; Hayes et al., 2009b). Matrix \mathbf{G} is defined as $\mathbf{G} = \mathbf{M}\mathbf{M}' / \sum 2p_i(1-p_i)$, where elements in column i of \mathbf{M} are $0 - 2p_i$, $1 - 2p_i$, and $2 - 2p_i$ for genotypes A_1A_1 , A_1A_2 , and A_2A_2 , respectively, and p_i is the allele frequency of A_2 , which was calculated from observed markers in the present study. For random residuals, it is assumed that $\mathbf{e} \sim N(\mathbf{0}, \mathbf{D}\sigma_e^2)$, where σ_e^2 is the residual variance, and \mathbf{D} is a diagonal matrix containing the element $d_{ii} = 1/w_i$, which was used to account for heterogeneous residual variances due to differences in reliabilities of DRP. The weights w_i were defined as $w_i = r_i^2 / (1 - r_i^2)$, where r_i^2 is the reliability of DRP for animal i . This weight expresses the inverse residual variance (in a standardized scale) of DRP. In the current data, reliability of DRP for animals in the training data ranged from 0.618 to 0.990 with an average of 0.939 for the 3 milk yield traits, from 0.250 to 0.990 with an average of 0.681 for fertility, and from 0.161 to 0.983 with an average of 0.822 for mastitis. The variation between reliabilities of DRP for a given trait was caused by different numbers of daughter records. To avoid possible problems resulting from extremely high weight values caused by the residual variances of DRP approaching zero, reliabilities larger than 0.98 were set to 0.98.

Bayesian EPOW Models

We implemented a Bayesian sparse shrinkage model by using an exponential power distribution for marker effects, here referred to as EPOW model. The EPOW model can be seen as a variation on Bayesian LASSO (Tibshirani, 1996; Park and Casella, 2008; Yi and Xu, 2008) with a tunable sparsity parameter. With q_i (the effect of SNP i), Bayesian LASSO assumes an exponential distribution on $|q_i|$, whereas EPOW uses an

exponential distribution on $|q_i|^\beta$. Using values of $\beta < 1$, a relatively sharper and longer-tailed distribution is made, leading to more intense shrinkage and higher sparsity in the marker effects, compared with Bayesian LASSO.

The model to describe the data, based on marker effects, is as follows:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{M}\mathbf{q} + \mathbf{e}, \quad [2]$$

where \mathbf{y} is the data vector of phenotypes of genotyped bulls (DRP), $\mathbf{1}$ is a vector of ones, μ is the overall mean, \mathbf{M} is the design matrix of marker genotypes as defined above, \mathbf{q} is the vector of SNP effects, and \mathbf{e} is the vector of residuals. The distribution of SNP effects is

$$p(\mathbf{q}) = \prod_{i=1}^m \frac{1}{2} \lambda e^{-\lambda |q_i|^\beta}, \quad [3]$$

where λ is a rate parameter, m is the number of markers, and β is the shape parameter controlling the sparsity. In the current study, 4 Bayesian EPOW models were used for genomic predictions. The 4 models differed in the shape parameters, which were set to be 0.3, 0.5, 0.8, or 1.0 (the ordinary Bayesian LASSO). These models were denoted as EPOW0.3, EPOW0.5, EPOW0.8, and EPOW1.0. The residuals were distributed as defined in Model [1].

For the Markov chain Monte Carlo (MCMC) implementation of this model, the conditional posterior distribution of SNP effects is not in a standard form. Combining a part coming from the likelihood (which will be Gaussian) and the prior distribution as given in [3], the conditional distribution for a SNP effect is in the form

$$p(q_i | \mathbf{y}, \text{other parameters}) \propto \exp\left[-(q_i - \hat{q}_i)^2 \frac{m'_i m_i}{2\sigma_e^2}\right] \exp\left(-\lambda |q_i|^\beta\right), \quad [4]$$

where m_i is column i of \mathbf{M} , $\hat{q}_i = (m'_i m_i)^{-1} m'_i \tilde{y}$, and \tilde{y} is the data corrected for the mean and all other SNP effects. The technique described by Damien et al. (1999) was used to sample parameters in this nonconjugate case by replacing [4] with

$$\mathbf{I}\left[u_1 < \exp\left[-(q_i - \hat{q}_i)^2 \frac{m'_i m_i}{2\sigma_e^2}\right]\right] \mathbf{I}\left[u_2 < \exp\left(-\lambda |q_i|^\beta\right)\right], \quad [5]$$

where $\mathbf{I}[\]$ denotes indicator function. In this technique, u_1 and u_2 are auxiliary variables, and the marginal dis-

tribution of [5] with respect to u_1 and u_2 is the needed conditional distribution of q_i (Damien et al., 1999). From [5], the conditional distributions for u_1 , u_2 , and q_i are all uniform.

The Bayesian model also estimates residual variance and the hyperparameter λ , using flat prior distributions. All parameters other than the SNP effect have standard distributions; that is, normal for the model mean, scaled inverse χ^2 for the residual variance, and Gamma for the exponential rate parameter λ .

Bayesian Mixture Model

The Bayesian mixture model used in this study was extended from George and McCulloch (1993) and Meuwissen (2009). Notably, we applied here a version with a 4-mixture distribution and applied Bayesian learning by estimating all variances in the mixture distribution. However, because of LD between SNP, confounding existed between the number of SNP with large effects and the size of the large effects. Thus, it is not particularly feasible to estimate both mixture distribution proportions and mixture distribution variances. Here, we chose to constrain the proportions in the mixture distribution and learn the variances. Use of a multi-mixture distribution improves computational efficiency by improved mixing of mixture indicators and SNP effects. High-density SNP data in cattle can show blocks of dozens of SNP in very high LD. The model to describe data is the same as model [2] but assumed that the distribution of marker effects was a mixture of 4 normal distributions:

$$q_i \sim \pi_1 N(0, \sigma_{\pi_1}^2) + \pi_2 N(0, \sigma_{\pi_2}^2) + \pi_3 N(0, \sigma_{\pi_3}^2) + \pi_4 N(0, \sigma_{\pi_4}^2).$$

Mixing proportions in this distribution were taken as known and set to $\pi_1 = 0.889$, $\pi_2 = 0.1$, $\pi_3 = 0.01$, and $\pi_4 = 0.001$; the variances were taken as model parameters and were estimated with flat prior distributions under the constraint $\sigma_{\pi_1}^2 < \sigma_{\pi_2}^2 < \sigma_{\pi_3}^2 < \sigma_{\pi_4}^2$. Model residuals were distributed as defined in Models [1] and [2]. The MCMC implementation of this mixture model adds an indicator variable to indicate membership of each SNP to one of the mixtures (but which may vary during MCMC cycles). Further MCMC implementation is straightforward with recognizable conditional distributions for all model parameters as described elsewhere (George and McCulloch, 1993; Meuwissen, 2009). The constraint on the mixture variances was implemented using a rejection sampler.

For all models, variances were estimated from the reference data. The analysis of GBLUP model was performed using the DMU package (Madsen and Jensen, 2010). The analysis of the Bayesian models was per-

formed using BayZ package (<http://www.bayz.biz/>). Each of the Bayesian analysis was run as a single chain with a length of 50,000 samples, and the first 20,000 cycles were regarded as the burn-in period.

Validation

The primary criterion to evaluate differences between genomic models and between relationship groups was the reliability of genomic predictions, evaluated as squared correlations between the predicted breeding values and DRP for each group of bulls in the validation data set and then divided by reliability of DRP (Su et al., 2012b). A Hotelling-Williams *t*-test (Dunn and Clark, 1971; Steiger, 1980) was used to test the difference between the validation correlations among these prediction models. Unbiasedness of genomic predictions was measured as the regression of DRP on the genomic predictions. A necessary condition for unbiased prediction was that the regression coefficient should not deviate significantly from 1 (Su et al., 2012a).

RESULTS

The reliabilities of genomic predictions using different models for different groups of bulls are shown in Tables 3, 4, 5, and 6, respectively. Genetic relationship between validation and training populations had a large effect on reliability of DGV, especially for the sires being included in or excluded from the training data set (Group_{smsg} vs. Group_{mgs}, and Group_{sire} vs. Group_{non}). Averaged over the 5 traits and the 6 models, the difference in reliability of DGV was 11.5 percentage points between Group_{smsg} and Group_{mgs}, and 10.4 percentage points between Group_{sire} and Group_{non}. Moreover, the influence of sire status in training population on reliability of DGV was larger for the 3 production traits than for fertility and mastitis. Maternal grandsire status in the training population (Group_{smsg} vs. Group_{sire}, and Group_{mgs} vs. Group_{non}) increased reliability of DGV for the 3 production traits, but not for fertility or mastitis. Averaged over the traits and the models, the difference in reliability of DGV was 6.4 percentage points between Group_{smsg} and Group_{sire}, and 5.3 percentage points between Group_{mgs} and Group_{non}. On average, the difference between Group_{smsg} and Group_{non} was 16.8 percentage points. In fact, about half of the animals in Group_{mgs} and Group_{non} had the paternal grandsire in the training data. If there was no paternal grandsire in the training data, the reliability of genomic prediction in these 2 groups could further reduce.

In general, the Bayesian models led to higher reliability of DGV than the GBLUP model, and the mixture and EPOW0.3 models performed better than the other

Table 3. Reliabilities (%) of genomic predictions using different models for the animals having sire and maternal grandsire in reference population (Group_{smsg})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	51.8	57.6	56.6	55.0	52.6	57.5
Fat	52.5	54.5	54.3	54.3	53.2	55.1
Protein	56.6	58.1	58.0	56.8	56.5	57.9
Fertility	35.1	34.7	34.8	34.9	35.4	35.0
Mastitis	39.0	39.7	39.3	38.9	38.8	39.4
Mean	47.0	48.9	48.6	48.0	47.3	49.0

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

Bayesian models, especially for production traits in Group_{non} and Group_{smsg}. Based on the data pooled over the 4 relationship groups, the Hotelling-Williams *t*-test showed that the differences between reliabilities of DGV from different models were statistically significant ($P < 0.05$) for production traits, except for those between the mixture, EPOW0.3, and EPOW0.5 for milk, between the mixture and EPOW0.3 for fat, and between the GBLUP, EPOW0.8, and EPOW1.0 and between the mixture and EPOW0.3 for protein. For fertility, a significant difference existed only between the mixture model and EPOW0.8. For mastitis, reliabilities of DGV obtained from the mixture, EPOW0.3, and EPOW0.5 models were significantly or near significantly ($P = 0.014$ to 0.062) higher than those from the GBLUP, EPOW0.8, and EPOW1.0, and the mixture model performed significantly better than EPOW0.5. Averaged over the 5 traits and the 4 relationship groups, the reliability of DGV was 40.9% using Bayesian mixture model; 40.6, 40.0, 39.4, and 38.3% using the EPOW models with shape parameters of 0.3, 0.5, 0.8, and 1.0, respectively; and 37.8% using the GBLUP model. The difference in reliability of DGV from the 6 models was large for the 3 production traits, but small for fertility

and mastitis. Moreover, the superiority of the Bayesian models over the GBLUP model was related to the genetic relationship between validation animals and training animals. Compared with the GBLUP model, on average over the 5 traits, the Bayesian mixture model increased reliability by 2.0, 2.7, 3.3, and 4.2 percentage points, and the Bayesian EPOW0.3 model increased reliability by 1.9, 2.8, 3.2, and 3.3 percentage points for Group_{smsg}, Group_{sire}, Group_{smsg}, and Group_{non}, respectively.

Pooled over the 4 relationship groups, the number of overlaps between the 200 top bulls based on DGV and 200 bulls based on DRP was calculated. Averaged over the 5 traits, the numbers of overlapped bulls were 83.6, 84.0, 85.6, 86.4, 86.8, and 87.6, according to DGV from GBLUP, EPOW1.0, EPOW0.8, EPOW0.5, EPOW0.3, and the mixture model, respectively. The rank was consistent with the one according to validation reliabilities.

Tables 7, 8, 9, and 10 present the regression coefficients of DRP on DGV from different models for each group of validation bulls, respectively. The patterns of regression coefficients in relation to models and groups differed among the traits. For milk yield, the Bayesian mixture model and the EPOW0.3 model led to more

Table 4. Reliabilities (%) of genomic predictions using different methods for the animals having sire but not maternal grandsire in reference population (Group_{sire})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	40.8	43.6	45.0	43.9	41.6	43.0
Fat	39.9	48.8	48.0	46.4	41.3	48.0
Protein	42.8	43.9	43.8	42.8	42.7	43.7
Fertility	37.2	38.3	37.3	37.2	36.8	38.4
Mastitis	39.5	39.7	39.6	39.2	39.9	40.3
Mean	40.0	42.8	42.7	41.9	40.5	42.7

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

Table 5. Reliabilities (%) of genomic predictions using different methods for the animals having maternal grandsire but not sire in reference population (Group_{mgs})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	46.9	53.8	52.2	51.7	48.1	52.2
Fat	33.9	42.9	38.3	37.4	35.2	42.0
Protein	40.0	40.8	41.2	40.1	39.9	41.5
Fertility	33.0	32.4	32.2	32.5	33.3	32.7
Mastitis	20.7	20.6	20.9	20.7	20.5	22.7
Mean	34.9	38.1	36.9	36.5	35.4	38.2

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

bias for all groups. For fat yield, these 2 models were worse than the other models for Group_{s_{mgs}}, but better than the other models for Group_{n_{on}}, with regard to bias of DGV. For protein yield, the same 2 models resulted in more bias than the other models for Group_{s_{ire}} and Group_{m_{gs}}. For fertility and mastitis, the differences in the regression coefficients among the models were small for all groups. With regard to genetic relationship groups, the largest bias of DGV for the 3 production traits arose in Group_{n_{on}} (weakest relationship), and for mastitis in Group_{m_{gs}} and Group_{n_{on}}. The differences in the regression coefficient between groups were relatively small for fertility. Averaged over the 5 traits, the differences in regression coefficient between the models were small, and we found a tendency that bias of genomic predictions increased with decreasing relationship between training and validation populations.

DISCUSSION

The present study investigated the influences of different models and genetic relationships between validation and training animals on the accuracy of genomic predictions based on HD markers in the Nordic Holsteins.

The Bayesian mixture model and Bayesian EPOW0.3 led to the highest reliabilities, followed by the EPOW0.5 and EPOW0.8 models. The EPOW model with shape parameter of 1.0 (Bayesian LASSO) and the GBLUP model resulted in the lowest reliabilities.

The advantage of the Bayesian mixture and EPOW0.3 models was more profound, with weak relationships between training and validation data sets, showing that these models indeed capture more LD between markers and QTL. Compared with the GBLUP model, the Bayesian mixture model increased the reliabilities of DGV by 2.0 percentage points for the validation animals with sire and MGS in training population (Group_{s_{mgs}}) to 4.2 percentage points for the validation animals without sire and MGS in training population (Group_{n_{on}}). For production traits, the difference was even higher (increasing from 3.2 to 6.2 percentage points). Su et al. (2012a) studied genomic predictions for protein yield, fertility, and mastitis based on HD markers in Nordic Holsteins, and reported that a Bayesian mixture model performed slightly better (0.5 percentage points higher) than a GBLUP model. However, they used a mixture model with 2 distributions. Those authors discussed that a mixture model with 2 distributions might not

Table 6. Reliabilities (%) of genomic predictions using different methods for the animals having neither sire nor maternal grandsire in reference population (Group_{n_{on}})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	27.6	34.7	33.0	31.7	28.6	35.8
Fat	35.1	42.3	39.9	39.3	36.5	43.7
Protein	27.9	29.4	28.9	28.0	27.8	29.7
Fertility	34.8	34.9	34.9	34.7	34.9	36.6
Mastitis	21.5	22.4	21.7	21.7	21.3	22.5
Mean	29.4	32.7	31.7	31.1	29.8	33.6

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

Table 7. Regression coefficient of deregressed proofs on genomic predictions from different models for the animals having sire and maternal grandsire in reference population (Group_{smsg})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	0.964	0.945	0.951	0.961	0.964	0.950
Fat	0.945	0.873	0.882	0.908	0.940	0.872
Protein	0.939	0.933	0.940	0.941	0.934	0.935
Fertility	0.949	0.934	0.948	0.944	0.951	0.930
Mastitis	0.893	0.896	0.893	0.886	0.891	0.886
Mean	0.938	0.916	0.923	0.928	0.936	0.915

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

be adequate to describe the distribution of true SNP effects. The current study suggests that a model with a mixture of 4 normal distributions as the prior distribution of SNP effects could be more reasonable, because a mixture of 4 normal distributions could describe the distribution of true SNP effects better than a mixture of 2 normal distributions. Ostersen et al. (2011) compared a GBLUP model, Bayesian LASSO, and Bayesian mixture model based on pig 60K data and found no difference among these models. The authors suggested that the advantage of the Bayesian models over the GBLUP model being able to efficiently capture the LD information could not be realized because the pig data was highly related.

Small improvements in reliability of predictions can have important effects on genetic progress in breeding programs. Genetic progress linearly depends on accuracy of genetic evaluation. For a trait such as milk yield in this study, the accuracy (square root of the provided reliability) of genomic prediction increases from 0.719 (using GBLUP) to 0.759 (using EPOW03) for strong relationships (Table 3), and from 0.525 to 0.589 for weak relationships (Table 6). This can be translated to increase in genetic gain of 5.4 and 12%,

respectively. Considering a large dairy cattle population, the improvements are less for other traits, but a small improvement in reliability as low as 1 or 2% is relevant for breeding. The disadvantage of the Bayesian models is the long computing time. For analysis of the current data in our computing system (Intel Xeon 2.93 GHz processor), the Bayesian models with 50,000 samples for one trait took about 120 h using 1 CPU. In practical implementations, it could be a good strategy to save the estimated SNP effects for prediction of new candidates and update SNP effects periodically (e.g., once or twice per year). Compared with the potential increases in genetic gain, the computing costs for using the Bayesian models are negligible.

Among the 4 Bayesian EPOW models, the EPOW0.3 model performed best in terms of DGV reliability, followed closely by EPOW0.5. Genomic predictions using the EPOW0.3 model were as accurate as those using the Bayesian mixture model. Less intense shrinkage models, using EPOW0.8 and EPOW1.0 (Bayesian LASSO), did not show clear advantages over the GBLUP model. The results indicate that the shape parameter has a considerable influence on the accuracy of genomic predictions, and an intense shrinkage of

Table 8. Regression coefficient of deregressed proofs on genomic predictions from different methods for the animals having sire but not maternal grandsire in reference population (Group_{sire})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	0.920	0.881	0.916	0.924	0.924	0.886
Fat	0.889	0.900	0.894	0.915	0.894	0.891
Protein	0.897	0.881	0.887	0.890	0.899	0.875
Fertility	0.912	0.922	0.917	0.912	0.905	0.924
Mastitis	1.023	1.020	1.022	1.019	1.031	1.005
Mean	0.928	0.921	0.927	0.932	0.931	0.916

¹EPOW_x = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

Table 9. Regression coefficient of DRP on genomic predictions from different methods for the animals having maternal grandsire but not sire in reference population (Group_{mgs})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	0.981	0.966	0.954	0.996	0.991	0.943
Fat	0.849	0.864	0.819	0.835	0.851	0.849
Protein	0.924	0.894	0.911	0.919	0.922	0.896
Fertility	0.941	0.935	0.932	0.934	0.941	0.929
Mastitis	0.782	0.782	0.787	0.784	0.784	0.820
Mean	0.895	0.888	0.881	0.894	0.898	0.887

¹EPOWx = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

explanatory variables is necessary for genomic prediction using HD data. A common concern exists with setup of the number of distributions and the mixing proportions in mixture models (e.g., BayesB, BayesC, BayesR, and the mixture model in this study) and sparsity parameter in EPOW models. An argument is that the parameters in the Bayesian models used in this study are not optimal. It will be interesting to further optimize the sparsity parameter in the EPOW model or the number of mixtures and mixture proportions in the multi-mixture model. This could be done by including additional hierarchies in the Bayesian models or in a machine learner’s fashion by cross validation.

Use of a 4-mixture distribution was also considered in “BayesR” by Erbe et al. (2012). However, in BayesR, one of the variances in the mixture distribution is set to zero, which does not allow sampling from full conditional distributions. From the equation given in Erbe et al. (2012) to sample SNP effects, it was unclear whether BayesR correctly overcomes this. We therefore used the parameterization of George and McCulloch (1993), where all 4 distributions have nonzero variances, which allows straightforward sampling of all model parameters from full conditional distributions.

The present study showed that genetic relationship between validation and training animals had a large influence on accuracy of genomic predictions for validation animals, especially sire-offspring relatedness. Similar results have been reported in several previous studies (Habier et al., 2007; Lund et al., 2009; Meuwissen, 2009; Habier et al., 2010; Clark et al., 2012; Pszczola et al., 2012). In this study, the genetic relationship between validation and training animals increased from Group_{non} to Group_{mgs}, and the accuracy of genomic predictions increased accordingly for all 6 models. This can be explained by the fact that with weaker relationship, less information from relatives was used to predict DGV (Habier et al., 2010). Habier et al. (2007) found that the accuracy of genomic predictions using only LD information was considerably lower than those using both LD and family information. Lund et al. (2009) reported that large differences in the accuracy of DGV between the group that has sires in the training data set and the group without sires in the training data set based on 54K SNP markers. Improving genomic predictions for the animals having a weak relationship with the training data set is very important when genomic predictions lead to the use of young bulls for breed-

Table 10. Regression coefficient of deregressed proofs on genomic predictions from different methods for the animals having neither sire nor maternal grandsire in reference population (Group_{non})

Trait	Genomic BLUP	Exponential power model ¹				Mixture ²
		EPOW0.3	EPOW0.5	EPOW0.8	EPOW1.0	
Milk	0.776	0.770	0.778	0.802	0.784	0.767
Fat	0.802	0.834	0.813	0.826	0.813	0.853
Protein	0.730	0.727	0.725	0.734	0.729	0.735
Fertility	1.013	1.021	1.016	1.009	1.017	1.039
Mastitis	0.825	0.832	0.829	0.836	0.828	0.830
Mean	0.829	0.837	0.832	0.841	0.834	0.845

¹EPOWx = exponential power model with shape parameter 0.30, 0.50, 0.80, and 1.0, respectively (the latter being Bayesian LASSO).

²Bayesian mixture model with 4 normal distributions.

ing. Many countries, such as the Nordic countries, have used a reasonable number of juvenile bulls selected on genomic EBV for breeding. In the near future, it will be a predominant situation that sires of young candidates will not be in the training data set because they do not have daughters' phenotypic information at the time of the candidates being selected. This means that the issue of evaluating young bulls without their fathers' progeny data is imminent. In this situation, as shown in this study, it is important that the Bayesian models perform better than the GBLUP model.

Although reliability of DGV reduced with decreasing relationship between validation and training animals, the amounts of reduction were different among the 6 models. The models with more intense shrinkage of SNP variables led to less reduction. The reductions of reliability from Group_{smsg} to Group_{non} were largest for the GBLUP model and Bayesian EPOW1.0 model, and smallest for the Bayesian mixture model. Correspondingly, the superiority of the Bayesian models over the GBLUP model was greater for the animals that had weaker genetic relationships with the training population. The results indicate that the contribution of population LD information and family information to genomic predictions may not be the same when using different models. Habier et al. (2010) did an analysis based on German Holsteins by controlling the genetic relationship between training data set and validation data set using BayesB and GBLUP models, and reported that the accuracy of genomic predictions decreased when genetic relationship decreased. In addition, they found that the Bayesian model exploits LD information much better than the GBLUP model.

Prediction bias was assessed by the regression coefficients of DRP on DGV (Tables 7, 8, 9, and 10). The patterns of regression coefficients in relation to the models and the relationship groups differed among the traits. Averaged over the 5 traits, the difference in regression coefficient between the models was very small. The GBLUP model led to least bias in Group_{smsg} and Group_{mgs}, EPOW0.8 resulted in the least bias in Group_{sire}, and the mixture model resulted in least bias in Group_{non}. The small difference in bias is in line with Su et al. (2012a), who found that the Bayesian mixture model did not reduce the bias of genomic prediction. On the whole, as the relationship between validation animals and training animals was weaker, the bias of genomic predictions became larger.

CONCLUSIONS

The results from this study indicate that a Bayesian model with intense shrinkage of the explanatory variable, such as the Bayesian mixture model and the

Bayesian EPOW0.3 in the current study, can improve genomic predictions using HD markers, especially for milk production traits. The improvement is more profound for the animals that have a weak relationship with the training population. This is important because the sires of candidates would not be in a future training data when the selection decision is made completely based on genomic predictions.

ACKNOWLEDGMENTS

The authors thank the Danish Cattle Federation (Aarhus, Denmark), Faba Co-op (Hollola, Finland), Swedish Dairy Association (Stockholm, Sweden), and Nordic Cattle Genetic Evaluation (Aarhus, Denmark) for providing data. This work was performed in the project "Genomic Selection—from function to efficient utilization in cattle breeding (grant no. 3405-10-0137)," funded under GUDP by the Danish Directorate for Food, Fisheries and Agri Business (Copenhagen, Denmark), the Milk Levy Fund (Aarhus, Denmark), VikingGenetics (Randers, Denmark), Nordic Cattle Genetic Evaluation, and Aarhus University (Aarhus, Denmark).

REFERENCES

- Browning, B. L., and S. R. Browning. 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84:210–223.
- Clark, S. A., J. M. Hickey, H. D. Daetwyler, and J. H. J. van der Werf. 2012. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet. Sel. Evol.* 44:4.
- Damien, P., J. Wakefield, and S. Walker. 1999. Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables. *J. R. Stat. Soc. B Stat. Methodol.* 61:331–344.
- Dunn, O. J., and V. Clark. 1971. Comparison of tests of the equality of dependent correlation coefficients. *J. Am. Stat. Assoc.* 66:904–908.
- Erbe, M., B. J. Hayes, L. K. Matukumalli, S. Goswami, P. J. Bowman, C. M. Reich, B. A. Mason, and M. E. Goddard. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci.* 95:4114–4129.
- George, E. I., and R. E. McCulloch. 1993. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* 88:881–889.
- Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177:2389–2397.
- Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2009. Genomic selection using low-density marker panels. *Genetics* 182:343–353.
- Habier, D., J. Tetens, F. R. Seefried, P. Lichtner, and G. Thaller. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* 42:5.
- Harris, B., and D. Johnson. 2010. The impact of high density SNP chips on genomic evaluation in dairy cattle. Pages 40–43 in Proc. Interbull Mtg. Interbull, Uppsala, Sweden.
- Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009a. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92:433–443.

- Hayes, B. J., P. M. Visscher, and M. E. Goddard. 2009b. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res. (Camb.)* 91:47–60.
- Jairath, L., J. C. Dekkers, L. R. Schaeffer, Z. Liu, E. B. Burnside, and B. Kolstad. 1998. Genetic evaluation for herd life in Canada. *J. Dairy Sci.* 81:550–562.
- Lund, M. S., S. P. W. de Ross, A. G. de Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbbrandtsen, Z. Liu, and R. Reents. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. *Genet. Sel. Evol.* 43:43.
- Lund, M. S., G. Su, U. S. Nielsen, and G. P. Aamand. 2009. Relation between accuracies of genomic predictions and ancestral links to the training data. Pages 162–166 in *Proc. Interbull Mtg., Barcelona, Spain*. Interbull, Uppsala, Sweden.
- Madsen, P., and J. Jensen. 2010. A User's Guide to DMU. Version 6, Release 5.0. University of Aarhus, Faculty Agricultural Sciences (DJF), Department of Genetics and Biotechnology, Research Centre Foulum, Tjele, Denmark.
- Meuwissen, T., and M. Goddard. 2010. Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics* 185:623–631.
- Meuwissen, T. H. E. 2009. Accuracy of breeding values of “unrelated” individuals predicted by dense SNP genotyping. *Genet. Sel. Evol.* 41:35.
- Ostensen, T., O. F. Christensen, M. Henryon, B. Nielsen, G. Su, and P. Madsen. 2011. Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in purebred pigs. *Genet. Sel. Evol.* 43:38.
- Park, T., and G. Casella. 2008. The Bayesian lasso. *J. Am. Stat. Assoc.* 103:681–686.
- Pszczola, M., T. Strabel, H. A. Mulder, and M. P. L. Calus. 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389–400.
- Schaeffer, L. R. 2001. Multiple trait international bull comparisons. *Livest. Prod. Sci.* 69:145–153.
- Solberg, T. R., A. K. Sonesson, J. A. Woolliams, and T. H. Meuwissen. 2008. Genomic selection using different marker types and densities. *J. Anim. Sci.* 86:2447–2454.
- Steiger, J. H. 1980. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* 87:245.
- Strandén, I., and E. A. Mäntysaari. 2010. A recipe for multiple trait deregression. Pages 21–24 in *Proc. Interbull Mtg., Riga, Latvia*. Interbull, Uppsala, Sweden.
- Su, G., R. F. Brondum, P. Ma, B. Guldbbrandtsen, G. R. Aamand, and M. S. Lund. 2012a. Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *J. Dairy Sci.* 95:4657–4665.
- Su, G., O. F. Christensen, T. Ostensen, M. Henryon, and M. S. Lund. 2012b. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS ONE* 7:e45293.
- Tibshirani, R. 1996. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc., B* 58:267–288.
- 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.
- Yi, N., and S. Xu. 2008. Bayesian LASSO for quantitative trait loci mapping. *Genetics* 179:1045–1055.