Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses

D. A. L. Lourenco,*[1] I. Misztal,* S. Tsuruta,* I. Aguilar,† E. Ezra,‡ M. Ron,§ A. Shirak,§ and J. I. Weller§
*Department of Animal and Dairy Science, University of Georgia, Athens 30602
†Instituto Nacional de Investigacion Agropecuaria, Las Brujas 90200, Uruguay
‡Israel Cattle Breeders Association, Caesaria Industrial Park, Caesaria 38900, Israel
§Institute of Animal Sciences, ARO, The Volcani Center, Bet Dagan 50250, Israel

ABSTRACT

Methods for genomic prediction were evaluated for an Israeli Holstein dairy population of 713,686 cows and 1,305 progeny-tested bulls with genotypes. Inclusion of genotypes of 343 elite cows in an evaluation method that considers pedigree, phenotypes, and genotypes simultaneously was also evaluated. Two data sets were available: a complete data set with production records from 1985 through 2011, and a reduced data set with records after 2006 deleted. For each production trait, a multitrait animal model was used to compute traditional genetic evaluations for parities 1 through 3 as separate traits. Evaluations were calculated for the reduced and complete data sets. The evaluations from the reduced data set were used to calculate parent average for validation bulls, which was the benchmark for comparing gain in predictive ability from genomics. Genomic predictions for bulls in 2006 were calculated using a Bayesian regression method (BayesC), genomic BLUP (GBLUP), single-step GBLUP (ssGBLUP), and weighted ssGBLUP (WssGBLUP). Predictions using BayesC and GBLUP were calculated either with or without an index that included parent average. Genomic predictions that included elite cow genotypes were calculated using ssGBLUP and WssGBLUP. Predictive ability was assessed by coefficients of determination (R^2) and regressions of predictions of 135 validation bulls with no daughters in 2006 on deregressed evaluations of those bulls in 2011. A reduction in R^2 and regression coefficients was observed from parities 1 through 3. Fat and protein yields had the lowest R^2 for all the methods. On average, R^2 was lowest for parent averages, followed by GBLUP, BayesC, ssGBLUP, and WssGBLUP. For some traits, R^2 for direct genomic values from BayesC and GBLUP were lower than those for parent averages. Genomic estimated breeding values using ssGBLUP were the least biased, and this method appears to be a suitable tool for genomic evaluation of a small genotyped population, as it automatically accounts for parental index, allows for inclusion of female genomic information without preadjustments in evaluations, and uses the same model as in traditional evaluations. Weighted ssGBLUP has the potential for higher evaluation accuracy.

Key words: genomic selection, few genotyped animals, single-step method, multitrait model

INTRODUCTION

Genomic selection has become a standard procedure in dairy cattle breeding because of its potential for increasing genetic gain through improved evaluation accuracy and for reducing generation interval (Hayes et al., 2009; VanRaden et al., 2009). Genomic selection was first introduced for US Holsteins in 2008 with evaluation of 5,285 progeny-tested bulls (VanRaden et al., 2008). The number of genotyped US Holstein bulls has increased by more than 10-fold, but numbers of genotyped animals in most other countries or of other breeds are substantially lower. The accuracy of genomic predictions is strongly dependent on the number of genotyped animals (VanRaden et al., 2009; Calus, 2010; Daetwyler et al., 2010). Therefore, accuracies of genomic EBV (**GEBV**) from small reference populations could be as low as in traditional BLUP evaluations that are based on pedigrees and performance of an animal or its progeny (VanRaden et al., 2009; Van Grevenhof et al., 2012).

Several methods that add genotypic information to pedigree and phenotypic information have been developed since genomic selection was first proposed in animal breeding (Meuwissen et al., 2001). The current method used in dairy cattle evaluations is based on a multistep procedure (VanRaden, 2008; Hayes et al., 2009; Lund et al., 2011; Wiggans et al., 2011b). In this procedure, the first step involves the estimation of traditional breeding values. In the second step, daughter yield deviations (**DYD**) or deregressed evaluations (**DD**) are calculated. In the third step, direct genomic

values (**DGV**) are computed for genotyped animals using methods based either on estimation of SNP effects or on genomic relationships using DYD or DD as phenotypic observations. A fourth step can be added to combine DGV with parent averages (**PA**) by selection index theory (VanRaden et al., 2009) to obtain GEBV (**GEBV$_i$**).

Generally, the amount of phenotypic information used to calculate DGV is smaller than the amount used to obtain PA, because only a subset of animals in the pedigree are genotyped (VanRaden et al., 2009). In this case, accounting for the standard sire-dam PA and the PA computed from the subset of genotyped ancestors in an index with DGV can help to improve genomic predictions (VanRaden et al., 2012). According to Su et al. (2012), the benefits from using such an index are dependent on high accuracies of DGV. Simpler and possibly less effective indices that combine only PA and DGV were also proposed (Guo et al., 2010).

Aguilar et al. (2010) and Christensen and Lund (2010) introduced a single-step approach for genomic evaluations called single-step genomic BLUP (**ssGB-LUP**), in which phenotypes, pedigree, and genotypes are utilized in a single analysis with GEBV as a direct output. In ssGBLUP, the pedigree-based relationship matrix (**A**) is replaced by a matrix (**H**) that integrates **A** and a genomic-based relationship matrix (**G**). All genotyped animals can be considered, and the inflation and bias of genomic evaluation can be minimized with small modifications to **H** (Aguilar et al., 2010; Tsuruta et al., 2011; Vitezica et al., 2011). The primary advantages of ssGBLUP are simplicity and lack of approximations. For example, Gray et al. (2012) compared several methods for genomic prediction of milk flow traits. For all traits, ssGBLUP accuracy was the highest.

Genotypes can be obtained for both bulls and cows. The number of progeny-tested bulls is limited, and elite cows used as bull dams can provide additional information. Use of genotypes of females in a multistep method initially decreased the accuracy of prediction for US Holstein bulls, mainly because females were selected for genotyping based upon their high genetic merit, which may be biased due to preferential treatment (Wiggans et al., 2011a); however, this decrease in accuracy could be reversed heuristically. The use of female genotypes in ssGBLUP increased prediction accuracy for both males and females without heuristics and modifications to the model (Tsuruta et al., 2013).

Production traits can be analyzed using repeatability or multitrait models. Weller and Ezra (2004) showed that a multitrait model, which assumes that each parity is a different trait, more accurately ranks animals of different ages than a repeatability model. In a multistep genomic evaluation, the BLUP evaluation may use a multitrait model, but the step for genomic prediction is usually single-trait because of computing limitations (Calus and Veerkamp, 2011). The ssGBLUP method allows for any type of BLUP model, including a multitrait model.

In general, methods based on genomic relationships assume the same weight per SNP. Such an assumption could undermine accuracy when selected SNP explain a large fraction of the genetic variance. Wang et al. (2012) modified ssGBLUP for genome-wide association to obtain SNP weights and incorporate them in a weighted ssGBLUP (**WssGBLUP**).

The primary objective of this study was to evaluate several methods of genomic prediction using a multitrait animal model for a population with a relatively small number of genotyped animals. An additional objective was to evaluate the effect on evaluation accuracy of adding female genotypes to ssGBLUP and WssGBLUP analyses.

MATERIALS AND METHODS

Data

Two Israeli Holstein data sets were analyzed. The complete data set contained 305-d milk, fat, protein, fat percentage, and protein percentage records for 713,686 cows with records for parity 1, 503,827 cows for parity 2, and 326,317 cows for parity 3. The cows calved from 1985 through 2011. The reduced data set included only production records through 2006 for 563,870 cows with records for parity 1, 391,977 cows for parity 2, and 249,954 cows for parity 3. These records were used to calculate genomic predictions using single-step genomic methods, traditional predictions, and DD. To estimate DGV, DD were used as phenotypic records for genotyped sires in multistep genomic methods.

A total of 1,305 bulls and 343 elite cows were genotyped for the Illumina BovineSNP50 BeadChip (Illumina Inc., San Diego, CA), which includes approximately 54,000 markers. Quality control retained SNP with call rates >0.9 for both SNP and animals genotypes, minor allele frequency >0.05, and departure from Hardy-Weinberg equilibrium (difference between expected and observed frequency) <0.15. Parent–progeny pairs were tested for conflicts. Single nucleotide polymorphisms with unknown position or located on sex chromosomes were not considered in the analyses. After the quality control, 30,359 SNP remained in the genotype file. Numbers of genotyped bulls and elite cows by birth year are shown in Table 1. Elite cows were the top 2% for the Israeli breeding index at the time they were selected for genotyping. Effects of SNP were estimated from a set of training animals using multistep genomic methods,

and then predictions based on genomic information were tested with a set of validation bulls. Validation bulls included 135 genotyped bulls born from 2002 to 2005 and with no daughter records in the reduced data set (2006) but ≥20 daughters in the complete data set (2011). For multistep evaluations, training bulls were required to have DD in the reduced data set (732 bulls). For single-step evaluation, the data set included all 1,305 genotyped bulls (the 732 multistep training bulls, 34 bulls without DD in the reduced data set, 82 bulls born from 2002 to 2005, 135 validation bulls, and 322 young bulls born since 2006), but only production records up to 2006. To evaluate the effect of including cow genotypes in single-step methods, genotypes of 343 cows were included, which was straightforward, because the single-step method uses observed phenotypes rather than pseudo-phenotypes.

Model and Analysis

The model adopted in Israel for evaluation of production traits in Holstein is a multitrait animal model (Weller and Ezra, 2004). This model was used to compute traditional genetic evaluation for parities 1 through 3 as separate traits. The EBV for individual parities were combined into a single total evaluation (total breeding value, **BVT**) with the index of Weller and Ezra (2004):

$$BVT = (EBV_1 + 0.73EBV_2 + 0.51EBV_3)/2.24,$$

where EBV_1, EBV_2, and EBV_3 are EBV for parities 1, 2, and 3, respectively. Traditional evaluations were calculated for the complete and reduced data sets. Using EBV and approximate accuracies for each data set, DD

were calculated as deregressed evaluations (VanRaden et al., 2009) as follows:

$$DD_i = \frac{EBV_i - PA_i}{R_i} + PA_i,$$

where the fraction is the deregressed Mendelian sampling; i is the ith parity; and R_i is the deregression factor obtained as in Wiggans et al. (2012):

$$R_i = \frac{DE_{animal}}{DE_{animal} + DE_{PA} + 1},$$

where DE_{animal} is equivalent daughter contributions from the animal and its progeny, and DE_{PA} is equivalent daughter contributions from parent averages.

Four different methods were used to calculate genomic predictions. The first was genomic BLUP (**GBLUP**), which used a **G** matrix as constructed by VanRaden (2008):

$$G = \frac{(M - P)(M - P)'}{2\sum_{j=1}^{n} p_j(1 - p_j)},$$

where **M** is a matrix of SNP genotypes for each animal, **P** is a matrix of 2 times the allele frequency of the second allele p at locus j (p_j). The denominator is a scaling factor for **G**.

The second method applied was a Bayesian regression model known as BayesC (Kizilkaya et al., 2010). This method assumes that a fixed amount π of SNP markers have no effect, whereas $1 - \pi$ have effects sampled from a normal distribution with an a priori

Table 1. Numbers of genotyped Israeli Holsteins by animal birth year

	Bulls				
		Training			
Birth year	All	With daughter records in 2006 or earlier	Validation	Young	Elite cows[1]
1975–1979	10	10	—	—	—
1980–1984	72	72	—	—	—
1985–1989	162	162	—	—	—
1990–1994	206	205	—	—	—
1995–1999	216	208	—	—	—
2000–1901	100	75	—	—	12
2002	41	—	11	—	7
2003	7	—	44	—	12
2004	18	—	48	—	32
2005	16	—	32	—	60
2006–2012	—	—	—	322	220
All	848	732	135	322	343

[1]Top 2% for the Israeli breeding index at the time of selection for genotyping.

variance assumed to be scaled in accordance with an inverse χ^2 distribution.

Both GBLUP and BayesC used DD in the reduced data set as phenotypic information. As the bulls in the training population had different number of daughters, DD was weighted by its reliability (**Rel**) expressed as equivalent daughter contributions; DE, the weighting factor for the DD, was obtained as in VanRaden and Wiggans (1991):

$$DE = \frac{\text{Rel of DD}}{1 - \text{Rel of DD}}.$$

Because the number of genotyped animals was small, the quality of predictions by both methods could be improved by blending DGV with information from BLUP analyses (VanRaden et al., 2009). The index used was based on selection index theory with modifications: $GEBV_i = b_{DGV}DGV + b_{PA}PA$; the weights ($b$) for DGV and PA were calculated as in Guo et al. (2010).

The third method was ssGBLUP, in which phenotypes, pedigrees, and genotypes are blended in a single analysis (Aguilar et al., 2010; Christensen and Lund, 2010). This method is similar to traditional BLUP in the way that phenotypes and pedigree-based relationships are combined in the mixed model equations, all recorded animals can be considered in the analysis, and pseudo-phenotypes (i.e., deregressed evaluations) are not needed because bulls are related to daughters' production through **A**. The inverse of the modified relationship matrix **H** was

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau(0.95\mathbf{G} + 0.05\mathbf{A}_{22} + \alpha)^{-1} - \omega\mathbf{A}_{22}^{-1} \end{bmatrix},$$

where τ is a scaling factor for $\mathbf{G}^{-1}$ and ω is a scaling factor for $\mathbf{A}_{22}^{-1}$, the inverse of the pedigree-based numerator relationship matrix for genotyped animals. Changes in both weights were investigated by Misztal et al. (2010) for final score for US Holsteins. Changes in τ had little effect on accuracy and prediction bias, but $\omega < 1$ helped to reduce the inflation of GEBV. A τ of 1.0 and a ω of 0.7 were chosen as indicated in Tsuruta et al. (2011). Weights for **G** (0.95) and $\mathbf{A}_{22}$ (0.05) can avoid singularity problems and can slightly improve predictions (VanRaden, 2008). The constant α is the difference between average values of $\mathbf{A}_{22}$ and **G**, and it accounts for the fact that genotyped animals are more related through **A** than **G** is able to reflect (Vitezica et al., 2011). Although the choice of α greatly influenced the biases of GEBV in broiler chicken (Chen et al., 2011), it seemed to have minimal effect in dairy cattle (results not shown).

The fourth method was WssGBLUP (Wang et al., 2012). In this procedure, GEBV was converted to SNP solutions and used to create estimates of each SNP variance. Subsequently, the variances were used to create weighted **G**, which is applied in ssGBLUP. The procedure can be iterative, although most improvements in accuracy are obtained after one round of iteration (Wang et al., 2012). A single round was applied for this study, except as indicated otherwise. In multitrait models using WssGBLUP, **G** needs to be identical for all the traits. For this study, the weight for **G** was derived separately for each parity and trait. Subsequent analyses used **G** with average weight across parities. The GEBV were also computed by (W)ssGBLUP with genotypes of elite cows included in the analysis.

GEBV (DGV) Validation

The GEBV validation test (Mäntysaari et al., 2010) was performed to determine the predictive ability of the 4 methods for each parity and BVT for all production traits. The deregressed evaluations from complete data set were used as the expected future progeny performance of validation bulls to assess predictive ability based on the reduced data set. The EBV from the reduced data set was used to calculate PA for validation bulls, which was the benchmark used to compare the gain in predictive ability due to genomics. The following weighted regression model was applied:

$$DD_{complete} = \mu + \delta X_{reduced} + e,$$

where $DD_{complete}$ is DD of the validation bulls computed based on the complete data set; μ is the mean; δ is a regression coefficient; $X_{reduced}$ is a bull's PA, DGV, or GEBV based on the reduced data set; and e is the residual. The linear regression was weighted by the reliability of $DD_{complete}$. Ideally, values of δ close to 1 indicate that early evaluations are successful in predicting the actual magnitude of differences among animals; therefore, for values of δ closer to 1, $X_{reduced}$ is an unbiased prediction of $DD_{complete}$. Deviations of $\pm 15\%$ from unity are generally acceptable (Tsuruta et al., 2011). The coefficient of determination (R^2) was used to assess the validation reliability of evaluation models.

Computation

The GS3 software of Legarra et al. (2013) was used to compute GBLUP DGV. The software estimates DGV for a subset of training animals and also the SNP effects, which are then used to obtain DGV for validation animals in a separate computation. The convergence criterion was set to 10^{-8}. Analyses involving

the BayesC method were performed using the GenSel software of Fernando and Garrick (2009); assumptions were that 95% of SNP had no effect and that priors for variance of SNP and residual effects followed a scaled inverse χ^2 distribution with 4 and 10 df, respectively. A total of 50,000 samples were generated with a burn-in of 5,000. Estimates of SNP effects and DGV were obtained from the posterior distribution of the 45,000 remaining samples. The convergence of Markov chains was analyzed through diagnostic tests available in the R coda package (Plummer et al., 2013). Traditional evaluations with ssGBLUP were computed using a modified version of BLUP90IOD (Tsuruta et al., 2001; Aguilar et al., 2011a). The convergence criterion was set to10^{-14} for all production traits.

RESULTS AND DISCUSSION

Predictive Ability with Different Genomic Methods

Initially, only genotypes from bulls were considered. Average reliabilities (SE) of PA in 2006 among traits for the 135 validation bulls were 0.36 (0.006), 0.34 (0.006), and 0.33 (0.006), for first, second, and third parities, respectively. For DD in 2011, average reliabilities were 0.90 (0.004), 0.87 (0.005), and 0.85 (0.005), respectively. The decrease in reliability of DD in 2011 with increase in parity was due to a smaller number of daughters with records in later parities. Correlations among PA as of 2006, $GEBV_i$ for BayesC and GBLUP, and GEBV for ssGBLUP based on the reduced data set for the 135 validation bulls are given in Table 2. As genomic predictions included PA, all the correlations were inflated by an autocorrelation due to the PA. Cor-

relations between PA and BayesC and GBLUP $GEBV_i$ were all high for all parities and production traits. Corresponding correlations of PA with ssGBLUP predictions were generally the same or lower for all parities compared with correlations for BayesC predictions, except for fat percentage. Compared with values for GBLUP predictions, correlations for ssGBLUP were higher in only a few cases. In general, BayesC was the method that had the greatest correlation with PA in 2006. For all traits and parities, correlations between the methods based on pseudo-phenotypes (BayesC and GBLUP) were higher than correlations between those methods and ssGBLUP. Lower correlations between single-step and other methods were expected, because the single-step method also incorporates information on nongenotyped animals that are related to genotyped animals and contribute phenotypes. Koivula et al. (2012) obtained low correlations between GBLUP and ssGBLUP predictions for a larger population of Nordic Red dairy cattle.

Table 3 shows heritabilities and R^2 for PA and genomic predictions from BayesC and GBLUP (DGV and $GEBV_i$) and from (W)ssGBLUP based on 2006 data for validation bulls. For all traits, R^2 generally decreased as parity increased, probably because of smaller heritabilities and fewer animals with records in later parities. Aguilar et al. (2011b) reported a reduction in R^2 across parities for conception rate in Holsteins. Luan et al. (2009) showed greater accuracy and lower bias for traits with greater heritabilities in Norwegian Red Cattle. For all parities, R^2 were much smaller for fat and protein yields than for the other traits. Animals have been intensely selected for those 2 traits, which could reduce R^2. According to VanRaden

Table 2. Correlations between genomic predictions[1] in 2006 obtained from different evaluation methods[2] and parent average (PA) in 2006 by parity for 135 Israeli Holstein validation bulls

Trait		Parity 1			Parity 2			Parity 3		
		BayesC	GBLUP	ssGBLUP	BayesC	GBLUP	ssGBLUP	BayesC	GBLUP	ssGBLUP
Milk yield	PA_{2006}	0.89	0.88	0.84	0.87	0.85	0.83	0.88	0.87	0.80
	BayesC		0.99	0.88		0.99	0.88		0.99	0.87
	GBLUP			0.87			0.87			0.86
Fat yield	PA_{2006}	0.93	0.89	0.85	0.90	0.87	0.84	0.91	0.88	0.89
	BayesC		0.98	0.84		0.98	0.82		0.98	0.87
	GBLUP			0.82			0.81			0.86
Protein yield	PA_{2006}	0.84	0.81	0.77	0.81	0.76	0.73	0.80	0.76	0.69
	BayesC		0.98	0.78		0.98	0.74		0.98	0.70
	GBLUP			0.77			0.73			0.69
Fat percentage	PA_{2006}	0.92	0.91	0.93	0.91	0.90	0.92	0.91	0.91	0.92
	BayesC		0.98	0.95		0.98	0.94		0.98	0.95
	GBLUP			0.94			0.94			0.95
Protein percentage	PA_{2006}	0.89	0.86	0.89	0.88	0.86	0.86	0.89	0.85	0.86
	BayesC		0.97	0.93		0.97	0.93		0.98	0.93
	GBLUP			0.90			0.91			0.92

[1]Genomic predictions for GBLUP and BayesC are genomic EBV obtained by an index that combines PA and direct genomic value.

[2]BayesC = Bayesian regression model; GBLUP = genomic BLUP; and ssGBLUP = single-step GBLUP.

et al. (2009) and Bijma (2012), the realized reliability could be much smaller than the published reliability if the trait is directly selected. The R^2 for component percentages were higher, even though those traits were derived from ratios of components to milk. VanRaden et al. (2009) found that R^2 for PA was 0.10 less for fat yield compared with fat percentage; the difference was 0.19 when genomic information was included. In a Brown Swiss population with a validation set as small as in our study, Wiggans et al. (2011b) showed R^2 of PA for fat (0.075) and protein (0.062) yield around 4 times smaller than for those for fat and protein percentages. Similar trends were observed in our study. VanRaden et al. (2009) showed that R^2 values increase as the number of genotyped individuals increases. When comparing US, New Zealand, and Australian genomic predictions, Hayes et al. (2009) found higher R^2 for US and New Zealand, because of their greater numbers of genotypes.

Predictions based on genomic information are expected to have higher R^2 than traditional evaluations. However, R^2 for BayesC DGV were smaller than R^2 for PA for fat percentage in the second parity and were the same for fat percentage in the first and third parities and milk in third parity. The GBLUP DGV had smaller R^2 than PA for milk in second parity, fat percentage in all parities, and protein percentage in the first 2 pari-

ties. For all traits, BayesC DGV had higher or same R^2 compared with GBLUP DGV, except for protein yield in third parity. According to a simulation study by Meuwissen (2009), GBLUP requires more records to reach a high accuracy than does a Bayesian method (BayesB), and when the number of records is small and the marker density is high, Bayesian methods seem superior to GBLUP. For $GEBV_i$, R^2 was higher than PA R^2 for all traits except fat percentage for the GBLUP method in all parities. In a study with simulated data, Guo et al. (2010) obtained better predictive ability with a nonlinear Bayesian model compared with GBLUP using a simplified index that did not account for PA from the subset of genotyped animals. They also reported low PA reliabilities, because of the small number of animals as well as small gains in genomic predictions by including PA information. When applying an index for predicted bulls that combines DGV, published PA, and PA from the subset of genotyped animals, VanRaden et al. (2012) found that R^2 decreased slightly when weight was removed from DGV and transferred either to PA or to PA from the genotyped subset, and that bias of prediction was reduced.

Blending PA information with DGV had a positive effect on R^2 and improved some poor genomic predictions relative to traditional predictions. For all parities, R^2 were higher for $GEBV_i$ compared with DGV, except for

Table 3. Heritabilities for production traits by parity and total breeding value (BVT)[1] index, and coefficients of determination (R^2) for regressions of deregressed evaluations in 2011 on traditional parent average (PA) and genomic predictions[2] in 2006 from different evaluation methods[3] for 135 Israeli Holstein validation bulls

| | | | | | R^2 | | | | |
| | | | | BayesC | | GBLUP | | | |
Trait	Parity	h^2	PA	DGV	$GEBV_i$	DGV	$GEBV_i$	ssGBLUP GEBV	WssGBLUP GEBV
Milk yield	1	0.39	0.18	0.23	0.25	0.20	0.24	0.24	0.23
	2	0.29	0.13	0.14	0.17	0.12	0.16	0.15	0.16
	3	0.27	0.11	0.11	0.15	0.11	0.13	0.14	0.14
	BVT	—	0.14	0.17	0.20	0.15	0.18	0.18	0.18
Fat yield	1	0.42	0.07	0.15	0.12	0.11	0.11	0.14	0.11
	2	0.38	0.06	0.12	0.11	0.11	0.11	0.12	0.09
	3	0.34	0.06	0.12	0.10	0.11	0.10	0.09	0.07
	BVT	—	0.06	0.13	0.11	0.11	0.10	0.11	0.09
Protein yield	1	0.34	0.07	0.21	0.19	0.19	0.18	0.20	0.20
	2	0.29	0.01	0.09	0.07	0.07	0.06	0.09	0.10
	3	0.27	0.01	0.07	0.05	0.08	0.06	0.09	0.09
	BVT	—	0.02	0.13	0.10	0.11	0.09	0.13	0.13
Fat percentage	1	0.42	0.37	0.37	0.43	0.26	0.36	0.40	0.43
	2	0.38	0.35	0.34	0.41	0.26	0.34	0.38	0.41
	3	0.34	0.33	0.33	0.39	0.25	0.32	0.36	0.39
	BVT	—	0.36	0.35	0.41	0.26	0.35	0.39	0.42
Protein percentage	1	0.34	0.33	0.40	0.44	0.32	0.37	0.41	0.46
	2	0.29	0.30	0.36	0.40	0.28	0.34	0.39	0.44
	3	0.27	0.27	0.33	0.38	0.27	0.33	0.37	0.42
	BVT	—	0.31	0.37	0.42	0.31	0.36	0.40	0.44

[1]Calculated from an index that combines breeding values for parities 1, 2, and 3 from a multitrait evaluation.

[2]DGV = direct genomic value; GEBV = genomic EBV; and $GEBV_i$ = GEBV calculated from an index that combines PA and DGV.

[3]BayesC = Bayesian regression model; GBLUP = genomic BLUP; ssGBLUP = single-step GBLUP; and WssGBLUP = weighted ssGBLUP.

fat and protein yield. The small number of genotyped animals in addition to the intense selection on those traits, which can cause very low R^2 for PA, most likely led to smaller R^2 after blending. The R^2 obtained for genomic predictions with ssGBLUP and WssGBLUP were always higher than PA R^2 for all parities and all traits. Comparing GEBV from ssGBLUP and $GEBV_i$ from BayesC in the first parity, BayesC $GEBV_i$ for milk yield was 0.01 higher than R^2 of ssGBLUP. The BayesC $GEBV_i$ R^2 was 0.03 higher than the ssGBLUP R^2 in first parity for fat percentage and protein percentage. However, including weights for SNP (WssGBLUP) resulted in R^2 values for percentage traits equal or higher than the ssGBLUP values. Weighting SNP resulted in increased R^2 values for milk and protein yield for parity 2. On average, ssGBLUP performed better than BayesC DGV, GBLUP DGV and $GEBV_i$, but had R^2 0.01 lower than BayesC for $GEBV_i$. Gao et al. (2012) found that the application of ssGBLUP increased validation reliability and reduced bias for 16 traits of Nordic Holsteins when compared with GBLUP in a validation population 10 times larger than in the present study. The R^2 for WssGBLUP GEBV averaged 0.01 and 0.04 higher than R^2 for BayesC and GBLUP $GEBV_i$, respectively, and 0.02 and 0.07 higher than the corresponding DGV. The advantage of WssGBLUP was greater for percentage traits, mainly because this method gives more weight to SNP with effects on the trait analyzed. Perhaps the difference between (W)

ssGBLUP and other methods would be higher if more animals were genotyped.

For the BVT index that combines all parities in a single total evaluation, DGV from BayesC had smaller R^2 than did PA for fat percentage, and DGV from GBLUP had smaller R^2 than did PA for both percentage traits. Fat percentage was the only trait with GBLUP $GEBV_i$ R^2 smaller than the PA R^2. The average BVT R^2 for ssGBLUP GEBV was 0.01 smaller than R^2 for BayesC $GEBV_i$ and 0.03 higher than R^2 for GBLUP $GEBV_i$, and 0.01 and 0.05 higher than corresponding DGV R^2. In all cases, WssGBLUP GEBV R^2 averaged 0.01 higher than ssGBLUP R^2.

Regression coefficients for all methods are in Table 4. The coefficients were within $\pm15\%$ of the optimal value ($\delta = 1.00$) for parity 1 for milk yield for all evaluation methods except for WssGBLUP GEBV, for fat percentage except for ssGBLUP GEBV, and for protein percentage for all methods. Coefficients for BayesC DGV and ssGBLUP GEBV for protein yield were also within the acceptable range. All other regression coefficients for milk, fat, and protein yields were outside of the acceptable range for all methods, especially for later parities. Almost all regression coefficients for fat percentage were in the acceptable range except for ssGBLUP GEBV for parity 1 ($\delta = 1.18$) and parity 2 ($\delta = 1.19$). However, including weights in ssGBLUP helped to reduce prediction bias for component percentages to within the acceptable range for all parities.

Table 4. Coefficients for regression of deregressed evaluations in 2011 on traditional parent average (PA) and genomic predictions[1] in 2006 from different evaluation methods[2] for production traits by parity and total breeding value (BVT)[3] index for 135 Israeli Holstein validation bulls

Trait	Parity	PA	BayesC		GBLUP		ssGBLUP GEBV	WssGBLUP GEBV
			DGV	$GEBV_i$	DGV	$GEBV_i$		
Milk yield	1	0.86	0.93	0.90	0.86	0.86	0.98	0.83
	2	0.75	0.77	0.76	0.68	0.71	0.84	0.72
	3	0.67	0.70	0.68	0.62	0.64	0.78	0.65
	BVT	0.77	0.84	0.81	0.76	0.77	0.89	0.75
Fat yield	1	0.47	0.77	0.59	0.61	0.54	0.68	0.47
	2	0.47	0.69	0.56	0.58	0.53	0.63	0.42
	3	0.45	0.68	0.54	0.58	0.52	0.57	0.39
	BVT	0.44	0.72	0.55	0.59	0.52	0.62	0.42
Protein yield	1	0.59	0.86	0.78	0.77	0.73	0.86	0.69
	2	0.27	0.65	0.50	0.50	0.44	0.69	0.53
	3	0.24	0.59	0.44	0.54	0.45	0.58	0.49
	BVT	0.36	0.74	0.60	0.63	0.56	0.73	0.58
Fat percentage	1	1.10	1.08	1.10	0.90	0.99	1.18	1.10
	2	1.13	1.08	1.10	0.91	0.99	1.19	1.09
	3	1.09	1.05	1.07	0.90	0.98	1.14	1.07
	BVT	1.11	1.09	1.10	0.92	1.00	1.18	1.09
Protein percentage	1	0.97	1.00	0.99	0.85	0.88	1.04	0.98
	2	0.95	0.99	0.97	0.82	0.87	1.04	0.97
	3	0.91	0.97	0.95	0.81	0.85	1.02	0.96
	BVT	0.96	1.00	0.98	0.86	0.89	1.04	0.97

[1]DGV = direct genomic value; GEBV = genomic EBV; and $GEBV_i$ = GEBV calculated from an index that combines PA and DGV.
[2]BayesC = Bayesian regression model; GBLUP = genomic BLUP; ssGBLUP = single-step GBLUP; and WssGBLUP = weighted ssGBLUP.
[3]Calculated from an index that combines breeding values for parities 1, 2, and 3 from a multitrait evaluation.
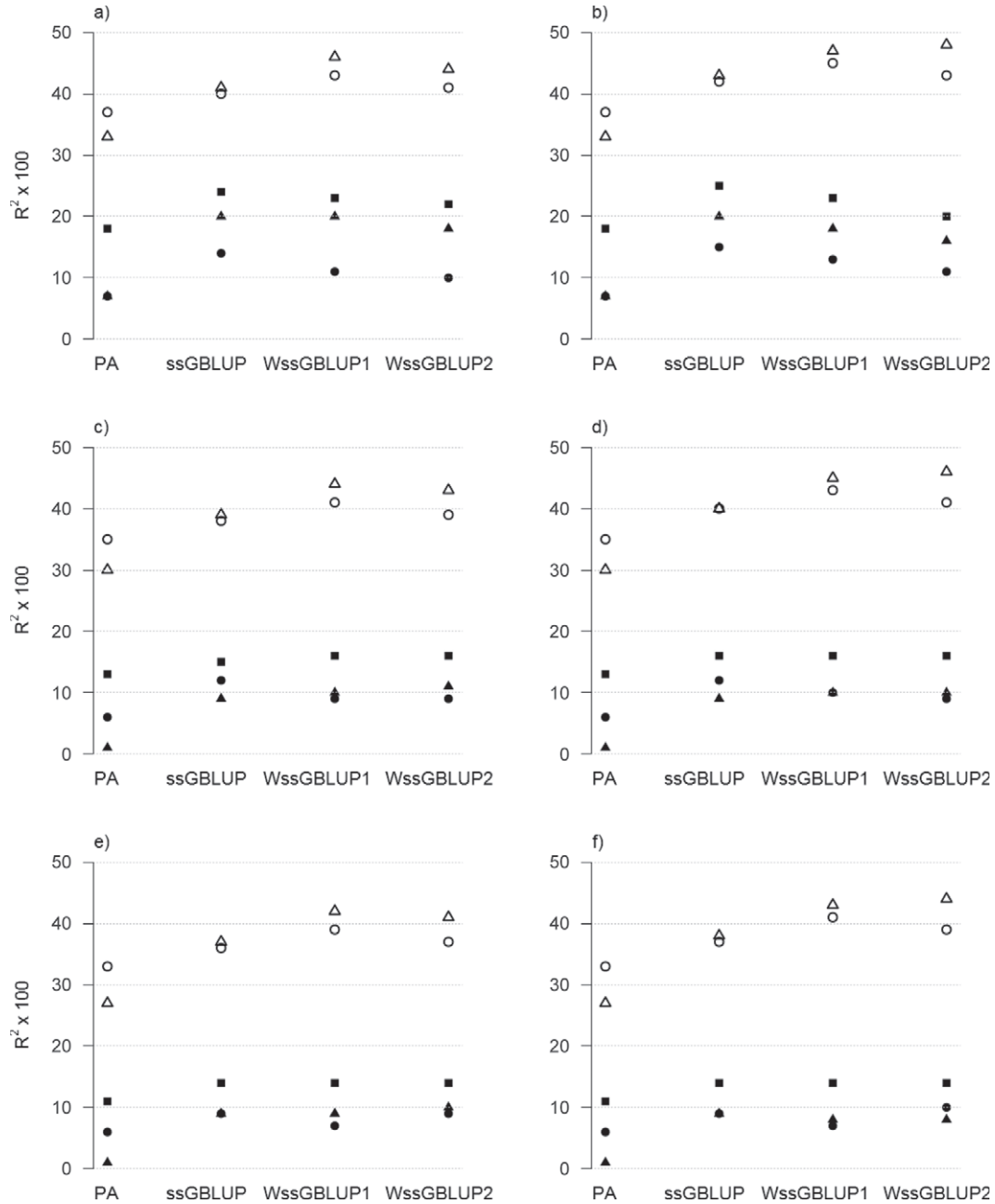
Figure 1. Reliabilities ($R^2 \times 100$) for parent average (PA) and genomic predictions from single-step genomic BLUP (ssGBLUP), and 2 iterations of weighted ssGBLUP (WssGBLUP1 and WssGBLUP2) for milk yield (■), fat yield (●), protein yield (▲), fat percentage (○), and protein percentage (△) for parity 1 (a, b), parity 2 (c, d), and parity 3 (e, f) when only 1,305 genotyped bulls were considered in the training population (a, c, e) and when 343 genotyped elite cows were also included in the training population (b, d, f) for evaluations of 135 Israeli Holstein validation bulls; elite cows were the top 2% for the Israeli breeding index at the time of selection for genotyping.

Predictive Ability with Cows Included in a Single-Step Method

Figure 1 shows coefficients of determination ($R^2 \times 100$) for PA and genomic predictions for evaluations both excluding and including cows' genotypes. The values are from ssGBLUP and the first 2 iterations of WssGBLUP. For most traits and parities, R^2 for WssGBLUP GEBV were higher after the first round of iteration than after the second round. This trend was very similar whether genotypes for cows were included or not. Consequently, first-round weights were chosen as the most appropriate weights for this study. In each round of iteration for WssGBLUP, the weights for SNP

Table 5. Coefficients of determination (R^2) and coefficients of regression (δ) for deregressed evaluations in 2011 on genomic EBV in 2006 from single-step genomic BLUP (ssGBLUP) and weighted ssGBLUP (WssGBLUP) with genotypes for 1,305 bulls and 343 elite[1] cows included for production traits by parity and total breeding value (BVT)[2] index for 135 Israeli Holstein validation bulls

Trait	Parity	R^2		δ	
		ssGBLUP	WssGBLUP	ssGBLUP	WssGBLUP
Milk yield	1	0.25	0.23	0.98	0.81
	2	0.16	0.16	0.84	0.72
	3	0.14	0.14	0.77	0.66
	BVT	0.19	0.19	0.89	0.74
Fat yield	1	0.15	0.13	0.68	0.49
	2	0.12	0.10	0.63	0.43
	3	0.09	0.07	0.54	0.21
	BVT	0.12	0.12	0.62	0.42
Protein yield	1	0.20	0.18	0.84	0.67
	2	0.09	0.10	0.67	0.51
	3	0.09	0.08	0.59	0.48
	BVT	0.13	0.12	0.72	0.55
Fat percentage	1	0.42	0.45	1.23	1.12
	2	0.40	0.43	1.22	1.11
	3	0.37	0.41	1.17	1.10
	BVT	0.40	0.43	1.21	1.12
Protein percentage	1	0.43	0.47	1.06	1.00
	2	0.40	0.45	1.05	0.99
	3	0.38	0.43	1.04	0.99
	BVT	0.41	0.46	1.05	0.99

[1]Top 2% for the Israeli breeding index at the time of selection for genotyping.

[2]Calculated from an index that combines breeding values for parities 1, 2, and 3 from a multitrait evaluation.

with large effects are increased, whereas those with small effects are decreased. Continuous iteration leads to SNP weights near zero for SNP with very small effects. Sun et al. (2012) evaluated weighted GBLUP with a formula emulating the expectation-maximization (EM) algorithm and found that prediction accuracies reached a plateau after a few iterations. Use of a similar formula in this study did not increase the accuracy. Eventually, WssGBLUP may benefit from additional refinements.

The R^2 and regression coefficients for ssGBLUP and first round of WssGBLUP with both bulls and cows in the training population are in Table 5. The additional use of cow genotypes generally had a small effect on regression coefficients, but the inclusion of weights for SNP reduced regressions regardless of the use of cow genotypes. This resulted in greater prediction bias for the yield traits, for which all regressions were less than unity, and lower prediction bias for the percentage traits, for which regressions were close to unity. Almost no prediction bias was observed for protein percentage for all parities or for BVT when cows were included in the training population for WssGBLUP evaluations. The inclusion of genotyped cows generally had a small but positive effect on R^2 (increases of 0.01 for yield traits and 0.02 for component percentages).

A decrease in R^2 for GEBV was observed for all traits as parity increased, regardless of whether genotypes of cows were included in the training population (Tables 4 and 5), except that R^2 for ssGBLUP GEBV for protein yield were the same for parities 2 and 3. A greater difference between R^2 for ssGBLUP and WssGBLUP GEBV was observed for component percentages for all parities and also for BVT compared with yield, again regardless of the inclusion of cow genotypes. Overall, cow evaluations have lower accuracy compared with bull evaluations, but genomic evaluations of genotyped cows are required by the industry (Wiggans et al., 2011b). Reliability gains from their inclusion are expected to be minimal because of their low accuracy (Misztal et al., 2013).

In the present study, the single-step method had an advantage over multistep methods mainly because single-step uses phenotypes rather than pseudo-phenotypes and accounts for the entire population structure to estimate GEBV. This method is also able to account for preselection on genotyped animals, because the inclusion of genotypes on elite cows did not erode the predictive ability of future data of candidate bulls. The reasons for changing to single-step methodology for routine evaluation are reduction of bias due to preselection and simplicity (Ducrocq and Legarra, 2011).

Although a multiparity model was applied for traditional and single-step evaluations, which accounted for incomplete genetic correlations among parities, decreases in R^2 and increases in prediction bias were observed over parities, due to the reduction in number

of daughters with increase in parity and the resulting lower reliability of deregressed evaluations for candidate bulls. Lower predictive ability in single-step methods could be observed over parities if a single-trait or a repeatability model were used. In the same way, inferior predictivity of multistep methods might be observed if deregressed evaluations from a single-trait model were used. However, this topic remains to be addressed in additional studies.

CONCLUSIONS

Estimating accurate genomic predictions in populations with small number of genotyped animals is problematic for highly selected traits. The DGV from multistep methods based on pseudo-phenotypes may be less accurate than PA, and an optimal index is needed to improve accuracies over PA. On average, the most accurate yet simplest method is ssGBLUP, where computing of pseudo-phenotypes or an index is not needed and the same model as in traditional evaluations can be used. Additional gain in validation reliability can also be achieved by optimizing weights for SNP; however, weights must be computed separately for each trait. Addition of genotypes of elite cows in ssGBLUP is straightforward and has a generally positive, albeit small, effect.

ACKNOWLEDGMENTS

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752.

Aguilar, I., I. Misztal, A. Legarra, and S. Tsuruta. 2011a. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J. Anim. Breed. Genet. 128:422–428.

Aguilar, I., I. Misztal, S. Tsuruta, G. R. Wiggans, and T. J. Lawlor. 2011b. Multiple trait genomic evaluation of conception rate in Holsteins. J. Dairy Sci. 94:2621–2624.

Bijma, P. 2012. Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. J. Anim. Breed. Genet. 129:345–358.

Calus, M. P. L. 2010. Genomic breeding value prediction: Methods and procedures. Animal 4:157–164.

Calus, M. P. L., and R. F. Veerkamp. 2011. Accuracy of multi-trait genomic selection using different methods. Genet. Sel. Evol. 43:26.

Chen, C. Y., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2011. Effect of different genomic relationship matrices on accuracy and scale. J. Anim. Sci. 89:2673–2679.

Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2.

Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185:1021–1031.

Ducrocq, V., and A. Legarra. 2011. An iterative implementation of the single step approach for genomic evaluation which preserves existing genetic evaluation models and software. Interbull Bull. 44:138–142.

Fernando, R., and D. Garrick. 2009. GenSel—User manual for a portfolio of genomic selection related analyses. 3rd ed., version 2.14. Accessed Mar. 31, 2013. http://www.biomedcentral.com/content/supplementary/1471-2105-12-186-s1.pdf.

Gao, H., O. F. Christensen, P. Madsen, U. S. Nielsen, Y. Zhang, M. S. Lund, and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genet. Sel. Evol. 44:8.

Gray, K. A., J. P. Cassady, Y. Huang, and C. Maltecca. 2012. Effectiveness of genomic prediction on milk flow traits in dairy cattle. Genet. Sel. Evol. 44:24.

Guo, G., M. S. Lund, Y. Zhang, and G. Su. 2010. Comparison between genomic predictions using daughter yield deviation and conventional estimated breeding value as response variables. J. Anim. Breed. Genet. 127:423–432.

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.

Kizilkaya, K., R. L. Fernando, and D. J. Garrick. 2010. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. J. Anim. Sci. 88:544–551.

Koivula, M., I. Standén, G. Su, and E. A. Mäntysaari. 2012. Different methods to calculate genomic predictions—Comparisons of BLUP at the single nucleotide polymorphism level (SNP-BLUP), BLUP at the individual level (G-BLUP), and the one-step approach (H-BLUP). J. Dairy Sci. 95:4065–4073.

Legarra, A., A. Ricardi, and O. Filangi. 2013. GS3: Genomic Selection—Gibbs Sampling—Gauss Seidel (and BayesCπ). Accessed Mar. 31, 2013. http://snp.toulouse.inra.fr/~alegarra/manualgs3_last.pdf.

Luan, T., J. A. Wooliams, S. Lien, M. Kent, M. Svendsen, and T. H. E. Meuwissen. 2009. The accuracy of genomic selection in Norwegian Red Cattle assessed by cross-validation. Genetics 183:1119–1126.

Lund, M. S., A. P. W. de Roos, A. G. de Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, Z. Liu, R. Reents, C. Schrooten, F. Seefried, and G. Su. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. Genet. Sel. Evol. 43:43.

Mäntysaari, E., Z. Liu, and P. VanRaden. 2010. Interbull validation test for genomic evaluations. Interbull Bull. 41:17–21.

Meuwissen, T. H. E. 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genet. Sel. Evol. 41:35.

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.

Misztal, I., I. Aguilar, A. Legarra, and T. J. Lawlor. 2010. Choice of parameters for single-step genomic evaluation for type. J. Dairy Sci. 93(Suppl. 1):533. (Abstr.)

Misztal, I., S. Tsuruta, I. Aguilar, A. Legarra, P. M. VanRaden, and T. J. Lawlor. 2013. Methods to approximate reliabilities in single-step genomic evaluation. J. Dairy Sci. 96:647–654.

Plummer, M., N. Best, K. Cowles, K. Vines, D. Sarkar, and R. Almond. 2013. Package 'coda'. Accessed Mar. 31, 2013. http://cran.r-project.org/web/packages/coda/coda.pdf.

Su, G., P. Madsen, U. S. Nielsen, E. A. Mäntysaari, G. P. Aamand, O. F. Christensen, and M. S. Lund. 2012. Genomic prediction for Nordic Red Cattle using one-step and selection index blending. J. Dairy Sci. 95:909–917.

Sun, X., L. Qu, D. J. Garrick, J. C. M. Dekkers, and R. L. Fernando. 2012. A fast EM algorithm for BayesA-like prediction of genomic breeding values. PLoS ONE 7:e49157.

Tsuruta, S., I. Misztal, I. Aguilar, and T. J. Lawlor. 2011. Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. J. Dairy Sci. 94:4198–4204.

Tsuruta, S., I. Misztal, and T. J. Lawlor. 2013. Short communication: Genomic evaluations of final score for US Holsteins benefit from the inclusion of genotypes on cows. J. Dairy Sci. 96:3332–3335.

Tsuruta, S., I. Misztal, and I. Strandén. 2001. Use of the preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. J. Anim. Sci. 79:1166–1172.

Van Grevenhof, E. M., J. A. M. Van Arendonk, and P. Bijma. 2012. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet. Sel. Evol. 44:26.

VanRaden, P. M., and G. Wiggans. 1991. Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737–2746.

VanRaden, P., G. Wiggans, C. Van Tassell, T. Sonstegard, and L. Walton. 2008. Genomic prediction. Changes to evaluation system (April 2008). Accessed Mar. 3, 2013. http://aipl.arsusda.gov/reference/changes/eval0804.html.

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

VanRaden, P. M., C. P. Van Tassel, 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.

VanRaden, P. M., J. R. Wright, and T. A. Cooper. 2012. Adjustment of selection index coefficients and polygenic variance to improve regressions and reliability of genomic evaluations. J. Dairy Sci. 95(Suppl. 2):520. (Abstr.)

Vitezica, Z. G., I. Aguilar, I. Misztal, and A. Legarra. 2011. Bias in genomic predictions for populations under selection. Genet. Res. (Camb.) 93:357–366.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. (Camb.) 94:73–83.

Weller, J. I., and E. Ezra. 2004. Genetic analysis of the Israeli Holstein dairy cattle population for production and nonproduction traits with a multitrait animal model. J. Dairy Sci. 87:1519–1527.

Wiggans, G. R., T. A. Cooper, P. M. VanRaden, and J. B. Cole. 2011a. Technical note: Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. J. Dairy Sci. 94:6188–6193.

Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2011b. The genomic evaluation system in the United States: Past, present, future. J. Dairy Sci. 94:3202–3211.

Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2012. Technical note: Adjustment of all cow evaluations for yield traits to be comparable with bull evaluations. J. Dairy Sci. 95:3444–3447.