Including overseas performance information in genomic evaluations of Australian dairy cattle

M. Haile-Mariam,*†¹ J. E. Pryce,*†§ C. Schrooten,‡ and B. J. Hayes*†§

*Biosciences Research Division, Department of Environment and Primary Industries, Agribio, and

†Dairy Futures Cooperative Research Centre, La Trobe University, Bundoora, Victoria 3083, Australia

‡CRV, PO Box 454, 6800 AL Arnhem, the Netherlands

§La Trobe University, Bundoora, Victoria 3086, Australia

ABSTRACT

In dairy cattle, the rate of genetic gain from genomic selection depends on reliability of direct genomic values (DGV). One option to increase reliabilities could be to increase the size of the reference set used for prediction, by using genotyped bulls with daughter information in countries other than the evaluating country. The increase in reliabilities of DGV from using this information will depend on the extent of genotype by environment interaction between the evaluating country and countries contributing information, and whether this is correctly accounted for in the prediction method. As the genotype by environment interaction between Australia and Europe or North America is greater than between Europe and North America for most dairy traits, ways of including information from other countries in Australian genomic evaluations were examined. Thus, alternative approaches for including information from other countries and their effect on the reliability and bias of DGV of selection candidates were assessed. We also investigated the effect of including overseas (OS) information on reliabilities of DGV for selection candidates that had weaker relationships to the current Australian reference set. The DGV were predicted either using daughter trait deviations (DTD) for the bulls with daughters in Australia, or using this information as well as OS information by including deregressed proofs (DRP) from Interbull for bulls with only OS daughters in either single trait or bivariate models. In the bivariate models, DTD and DRP were considered as different traits. Analyses were performed for Holstein and Jersey bulls for milk yield traits, fertility, cell count, survival, and some type traits. For Holsteins, the data used included up to 3,580 bulls with DTD and up to 5,720 bulls with only DRP. For Jersey, about 900 bulls with DTD and 1,820 bulls with DRP were used. Bulls born after 2003 and genotyped cows

that were not dams of genotyped bulls were used for validation. The results showed that the combined use of DRP on bulls with OS daughters only and DTD for Australian bulls in either the single trait or bivariate model increased the coefficient of determination [(R^2) (DGV,DTD)] in the validation set, averaged across 6 main traits, by 3% in Holstein and by 5% in Jersey validation bulls relative to the use of DTD only. Gains in reliability and unbiasedness of DGV were similar for the single trait and bivariate models for production traits, whereas the bivariate model performed slightly better for somatic cell count in Holstein. The increase in R^2 (DGV,DTD) as a result of using bulls with OS daughters was relatively higher for those bulls and cows in the validation sets that were less related to the current reference set. For example, in Holstein, the average increase in R^2 for milk yield traits when DTD and DRP were used in a single trait model was 23% in the least-related cow group, but only 3% in the most-related cow group. In general, for both breeds the use of DTD from domestic sources and DRP from Interbull in a single trait or bivariate model can increase reliability of DGV for selection candidates.

Key words: genomic prediction, daughter trait deviation, deregressed proof, reliability

INTRODUCTION

The reliability of direct genomic values (DGV) is a function of the heritability of the trait, the proportion of genetic variance explained by the markers, the genetic diversity of the population, and the number of animals in the reference population where SNP effects are estimated (Daetwyler et al., 2008; Goddard, 2009). The reliability of DGV is also affected by relatedness of the candidate animals to the reference population (e.g., Habier et al., 2010), at least with 50,000 SNP. Broadening the diversity of the reference set by including more animals can improve the reliability of DGV because for each candidate, the likelihood of having related animals in the reference set increases (Clark et al., 2012; Pszczola et al., 2012).

Received August 26, 2014.

Accepted January 23, 2015.

¹Corresponding author: Mekonnen.HaileMariam@depi.vic.gov.au

For national dairy cattle genomic evaluations, one way to diversify and increase the size of the reference set would be to use genotype of animals with performance records in other countries, for example bulls with daughters in other countries. Currently several countries exchange genotype data to increase their reference set for genomic evaluation and use deregressed multiple trait across-country evaluation (**MACE**) proofs as a response variable. Including information from other countries for genomic predictions has resulted in small to substantial increases in reliability of genomic evaluations (Brøndum et al., 2011; Lund et al., 2011; VanRaden et al., 2012). Including overseas (**OS**) information for genomic prediction is more likely to improve reliability when genetic correlations among countries are reasonably high, that is limited genotype by environment interaction (**G**×**E**) exists between the countries. For example, the genetic correlation between the Eurogenomics consortium countries is high (~0.9 for most traits; Interbull, 2013), this means there is weak **G**×**E** and treating performance in the Eurogenomics consortium countries as a single trait works well (Lund et al., 2011). Studies in Europe (Lund et al., 2011) and the United States (VanRaden et al., 2012) have shown that the increase in reliability of bull proofs from using data from other countries was greater for traits with higher genetic correlations among countries (e.g., protein) than for traits that have lower correlations between countries (e.g., fertility or survival). Furthermore, VanRaden et al. (2012) suggested that countries with more diverse populations and where trait definitions are more variable would be expected to benefit more from implementing multitrait genomic evaluation models (with performance in different countries treated as different traits) rather than single-trait models. The genetic correlation of the same trait measured in Australia with most of the northern hemisphere countries is typically about 0.8 or less (Interbull, 2013) for Jersey and Holstein. This is likely to be because the majority of Australian (**AUS**) herds practice pasture-based seasonal calving, whereas most herds in northern hemisphere largely rely on TMR. Thus, simply including OS information in the reference set for AUS genomic evaluations in a single-trait model could potentially create biases or reduce the value of including the information. Australia is an interesting test case to investigate the best approach to include OS daughter information in the reference population, as the **G**×**E** between Australia and other countries is considerable.

The main aim of our study was to evaluate alternative approaches (including multitrait) for using OS information in AUS genomic evaluations. The main criteria for evaluation were how well DGV from these alternative approaches predict a bull's AUS daughter

performance. This is done by predicting DGV either (1) using domestic daughter trait deviations (**DTD**) as response variables, using OS information by including deregressed proofs (**DRP**) from MACE for bulls with only OS daughters as well as domestic DTD as the response variable in a single trait model; or, (2) using a multitrait approach where DTD for bulls with AUS daughters and **DRP** for bulls with only OS daughters were considered as different traits. The second objective was to test if improvement in reliability of DGV as a result of using OS information was relatively larger if the candidate animals were less related to the current reference set. In addition to bulls, cows were used to test this hypothesis, as the degree of relatedness of the genotyped cows available to us was more diverse than that of bulls. Separate analyses were conducted using Holstein and Jersey data.

MATERIALS AND METHODS

Genotype Data

Bulls in the AUS reference set were genotyped for the 50,000 panel (Matukumalli et al., 2009); quality filters were imposed as described in Hayes et al. (2009) and Nieuwhof et al. (2010). Genotypes were also available for about 8,478 Holstein and 3,917 Jersey cows that were used in the reference set for AUS genomic evaluations, but here we used a subset of them as validation animals. For both Holsteins and Jerseys, 43,990 SNP were retained for genomic prediction.

Raw genotype data files were obtained for 5,720 bulls from the Cooperative Cattle Improvement Organization [**CRV**; Arnhem, the Netherlands (**NLD**)] and 3,072 bulls from Livestock Improvement Corporation [**LIC**; Hamilton, New Zealand (**NZL**)]. Of the 5,720 bull genotypes received from CRV, 4,228 Holstein and 561 Jersey were not genotyped in AUS and were included in this study. Of the 3,072 bull genotypes received from LIC, 1,472 Holstein and 1,182 Jersey bulls were not genotyped in AUS and none had AUS daughters for inclusion in the Australian Dairy Herd Improvement Scheme (**ADHIS**) routine genetic evaluation. The genotype data of these bulls, which were evaluated using the commercial Illumina BovineSNP50 chip (Illumina, San Diego, CA), were edited for imputation up to the same number of SNP (43,990) as the AUS data. When imputing these Holstein and Jersey bulls, the genotype data of AUS Holstein and Jersey bulls, respectively, were used as a reference set, and Beagle (Browning and Browning, 2009) was used for imputation.

The genotype data provided by CRV were carried out using 1 of the 2 versions of a custom 50,000 SNP chip. The CRV custom 50,000 SNP panel and the commercial

Illumina BovineSNP50 chip used to genotype the AUS bulls shared from 10,000 to 17,000 SNP (Lund et al., 2011). The CRV genotype data were imputed to have 43,990 SNP based on the AUS bull genotype data using Beagle (Browning and Browning, 2009). The quality of imputation was checked by comparing genotype data of bulls that were genotyped twice (i.e., AUS vs. NZL or AUS vs. NLD). For bulls genotyped more than once, the imputation showed very high concordance between imputed genotypes (>99%) from OS data and those that were provided by ADHIS.

Phenotype Data

The traits analyzed were milk, fat and protein yield, SCC, fertility, survival, mammary system, overall type, stature, and pin set. The phenotype data for AUS animals were DTD for bulls and trait deviations (TD) for cows, which were calculated by ADHIS for routine genomic evaluation (Nieuwhof et al., 2010). The TD were calculated from raw phenotype data by fitting a model that included the fixed herd-year-season (class) and age of cow (covariate) and the random permanent environment effect. The DTD for bulls were then calculated from the TD of their daughters after correcting for the genetic merit of their mate.

For NLD, NZL, and AUS bulls whose data were submitted to Interbull, MACE proofs were available. Both the DTD and MACE proof that were used for the current study were those used for the official national and MACE evaluation run of April 2013, respectively. The MACE proofs obtained from Interbull on the Australian scale were deregressed using the full sire-dam Interbull pedigree for Holstein and Jersey separately. The deregression software developed by Liu (2009) produced DRP and effective daughter contributions for

each bull with MACE proof on Australian scale taking into account the genetic correlation matrix among Interbull member countries. The deregression procedure uses a Gauss-Seidel algorithm to solve the equation system (Liu, 2009).

For the selected traits, the number of bulls with DTD and with MACE DRP included in our study is shown in Table 1 for Holsteins and Jerseys. For Holsteins, 3,020 bulls with daughters in AUS and 3,943 bulls with daughters in NZL or NLD were included in the reference set, and 1,574 bulls were included in the validation set (bulls born after 2004). Of the Holstein bulls, about 270, mostly born before 1986, were included in the reference set when DTD was the response variable but not used when DRP was the response variable. This is because the routine evaluation by Interbull uses data of bulls born after 1985 (Table 1). For Jerseys, there were 784 bulls with daughters in AUS and 1,216 bulls with daughters in NZL (LIC, CRV), and 609 validation bulls (born after 2004). For milk yield traits, 1,553 Holstein and 2,077 Jersey genotyped cows with 2 or more repeated records were used as a validation set to assess the effect of genomic relationship on reliability of DGV.

Data Analyses

The imputed 43,990 SNP data of all animals were used to form the genomic relationship matrix (**G**) following Yang et al. (2010), separately for each breed. The **G** constructed using the method of Yang et al. (2010) standardizes the genotype calls (x_{ij}) that have mean = 0 and variance = 1 by the transformation $z = (x_{ij} - 2p_i) / [2p_i(1 - p_i)]^{1/2}$, in which x_{ij} are genotypes of the i th SNP of the j th individual scored as 0 or 2 for homozygotes and 1 for heterozygotes and p_i is the allele frequency of the i th SNP.

Table 1. Number of reference and validation bulls with daughter trait deviation (DTD) or deregressed proofs (DRP) for the analyzed traits based on the data available in April 2013

Item	Reference bulls		Validation bulls ¹	
	with DTD	with DRP	with DTD	with DRP
Holstein				
Milk yield traits	3,020 (2,749) ²	6,963	534 (513)	1,574
Cell count	2,892 (2,658)	6,848	534 (513)	1,574
Survival	3,004 (2,721)	6,931	520 (471)	1,479
Fertility	2,990 (2,723)	6,945	520 (471)	1,479
Overall type	2,057 (1,059)	4,970	317 (177)	1,552
Jersey				
Milk yield traits	784 (728)	2,000	190 (173)	609
Cell count	756 (636)	1,824	174 (135)	602
Survival	776 (713)	2,000	190 (111)	344
Fertility	756 (685)	1,824	135 (135)	426
Overall type	440 (231)	1,476	38 (38)	421

¹Bulls born in 2004 and after are in the validation set.

²Bulls with both DTD and DRP are in parentheses.

Table 2. Models and response variables [daughter trait deviation (DTD) and deregressed proofs (DRP)] used for analyses of data of bulls

Model	Response variable	Overseas daughter information used	Model	Additional effect fitted	Weighting ¹
1	DTD	No	Univariate	μ	Weighted
2	DRP	Yes	Univariate	μ	Weighted
3	DTD or DRP ²	Yes	Univariate	Type of response variable	Weighted
4	DTD or DRP ²	Yes	Bivariate	Type of response variable	No
5	DTD and DRP ³	Yes	Bivariate	Type of response variable	No

¹Calculated as Garrick et al. (2009) assuming the proportion of variance not explained by markers (SNP) is 0.2.

²DTD for bulls with domestic daughters and DRP for bulls with overseas daughters only.

³Both DTD and DRP of bulls used if available.

First we estimated the proportion of variance explained by the marker data (i.e., $\mathbf{G}$) using the DTD (model 1; Table 2) and DRP of AUS and OS bulls (model 2; Table 2). This is done because the proportion of variance explained in DTD or DRP by $\mathbf{G}$ relative to the total DTD or DRP is a parameter affecting the reliability of the DGV (e.g., Daetwyler et al., 2012). Genomic correlations between DRP and DTD were also calculated using 2 bivariate models (models 4 and 5; Table 2) to find out if the 2 sources of information are similar or different.

To achieve the objective of our study, either the DTD or DRP of bulls, or combinations of both, were used as response variables and analyzed using 5 models described in Table 2 and below. The models vary in whether or not they use data from bulls with only OS daughter information (models 1 vs. models 2–5), and whether they are single trait or bivariate (models 1, 2, 3 vs. models 4 and 5). All data analyses were carried out using ASReml (Gilmour et al., 2009).

The single-trait model (i.e., models 1–3) was

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

where $\mathbf{y}$ is a vector of response variables, $\mathbf{b}$ is a vector of the effect of the type of the response variable (DTD or DRP) fitted when both DTD and DRP are used as shown in Table 2 (model 3), $\mathbf{X}$ is a design matrix relating $\mathbf{y}$ to type of response variable, $\mathbf{Z}$ is a design matrix relating records to DGV of animals, $\mathbf{g}$ is a vector of DGV $\sim N(0, \mathbf{G}\sigma_g^2)$, where $\mathbf{G}$ is the genomic relationship matrix, σ_g^2 is the genomic variance, and $\mathbf{e}$ is a vector of random residuals, $\sim N(0, \mathbf{D}\sigma_e^2)$, where σ_e^2 is residual variance and $\mathbf{D}$ is diagonal matrix derived from Garrick et al. (2009) to account for reliability of DTD and DRP of bulls assuming that 80% of the total variance is explained by marker data (Haile-Mariam et al., 2013b). However, when estimating the proportion of variance explained by $\mathbf{G}$ and genetic correlation, DTD and DRP were not weighted because the weighting formula as-

sumes that the $\mathbf{G}$ explained 80% of the variance, which would in turn alter the proportion of variance explained.

For the bivariate models (i.e., models 4 and 5),

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_1 & 0 \\ 0 & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & 0 \\ 0 & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}, \quad [2]$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are the vector of response variables (i.e., trait 1 is the DTD of AUS bulls and trait 2 is the DRP of bulls with OS daughters), $\mathbf{I}_1$ and $\mathbf{I}_2$ are identity matrices, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the vector of intercepts of DTD and DRP, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are the design matrices that relate DGV with the response variables, $\mathbf{g}_1$ and $\mathbf{g}_2$ are the vector of DGV for DTD and DRP, and $\mathbf{e}_1$ and $\mathbf{e}_2$ are vectors of random residuals for DTD and DRP, respectively. It was assumed that $\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} \sim MVN(0, \mathbf{G} \otimes \mathbf{T})$,

where $\mathbf{T} = \begin{bmatrix} \sigma_{g1}^2 & \sigma_{g12} \\ \sigma_{g12} & \sigma_{g2}^2 \end{bmatrix}$, the variance-covariance matrix

of DTD and DRP, and $\begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \sim MVN(0, \mathbf{I} \otimes \mathbf{R})$, where

$\mathbf{R} = \begin{bmatrix} \sigma_{e1}^2 & \sigma_{e12} \\ \sigma_{e12} & \sigma_{e2}^2 \end{bmatrix}$, the residual variance-covariance matrix

of DTD and DRP for model 5, and MVN = multivariate normal. Model 4 is the same as model 5, except the off-diagonals of the $\mathbf{R}$ matrix were by definition zero.

Criteria for Comparison of Models 1 to 5

The different models (Table 2) were used to predict DGV for a set of Holstein and Jersey validation bulls born in 2004 and later (Table 1). The following statistics were assessed for validation bulls: (1) coefficient of determination between DTD and DGV in the validation animals (R^2 , the proportion of variance in DTD explained by DGV); and (2) the slope of the regression (b) of DTD on DGV. The DGV used for validation from the bivariate models were the DTD components.

When calculating R^2 and the slope (b), the number of effective daughters per bull was used as weight in ASReml (Gilmour et al., 2009).

Adjusted R^2 [$R(\text{DGV, true breeding value})$]² of validation bulls were calculated by dividing the $R^2(\text{DGV, DTD})$ by the average reliability of the DTD (calculated from h^2 and number of effective daughters) of validation bulls. The adjusted R^2 are likely to be overestimated because we assumed that the correlation between DGV and DTD is through their correlation to the true breeding values (**BV**), which may not be true (e.g., Amer and Banos, 2010). These were compared with the theoretical or expected reliabilities (e.g., Hayes et al., 2009) that were calculated for the i th bull as $1 - \frac{PEV_i}{\sigma_g^2}$, where PEV_i is the prediction error variance of DGV for the i th bull and σ_g^2 is the genomic variance.

Effect of Relatedness of Validation Animals with the Current Reference set on R^2

One of the objectives of using OS information was to improve reliability of DGV for young bulls less related to the current AUS reference bull population. According to Clark et al. (2012) and Daetwyler et al. (2013), the average of the 10 highest genomic relationship coefficients between a candidate animal and animals in reference set is strongly related to reliability of DGV. To assess the effect of relatedness on empirical reliabilities for selected traits, validation bulls were divided into 2 groups of about equal size.

1. Bulls that were least related to the current reference population: validation bulls were grouped into this category if the average genomic relationship of the top 10 related bulls in the AUS reference set to a candidate bull was below 0.225 in case of the Holstein (mean = 0.204, $n = 250$) and below 0.20 in the case of Jersey (mean = 0.176, $n = 69$).
2. Bulls that were most related to the current reference population: validation bulls were grouped into this category if the average genomic relationship of the top 10 related bulls in the AUS reference set to a candidate bull was 0.225 or above in case of the Holstein (mean = 0.275, $n = 264$) and is 0.20 or above in the case of Jersey bulls (mean = 0.243, $n = 110$).

Similarly, 1,553 Holstein and 2,077 Jersey cows that were not dams of genotyped bulls with at least 2 lactations were categorized into groups based on the strength of their relationship to the current reference set (i.e.,

the mean of the top 10 genomic relationship coefficients of cows with the bulls in the current AUS reference set) to examine if their relationship is associated with the observed $R^2(\text{DGV, TD})$ for milk yield traits. For both breeds, cows were classified by their mean of the top 10 genomic relationship into 5 groups as <0.15 , ≥ 0.15 and <0.20 , ≥ 0.20 and <0.25 , ≥ 0.25 and <0.30 , and ≥ 0.30 . For each group the overall mean genomic relationship between validation cows and the top 10 highest related bulls in the reference set was 0.132, 0.162, 0.245, 0.261, and 0.327 for Holstein and the number of cows in each group was 351, 187, 172, 623, and 211, respectively. The overall genomic relationship mean of each group for Jersey was 0.128, 0.178, 0.220, 0.273, and 0.332, respectively, and the corresponding numbers of Jersey cows were 268, 846, 746, 124, and 92. For milk yield traits, DGV were predicted for these cows using the different models (Table 2).

Validation in Bulls Using DRP as Criteria for Selected Traits

Preliminary analyses showed that for some traits, such as fertility and some type traits, $R^2(\text{DGV, DTD})$ were lower than expected. One possible reason for these could be when we used DRP as a response variable by itself (model 2) or in addition to DTD (models 3–5) we still used DTD as validation criteria, so we assessed the usefulness of DRP for validation for selected traits including fertility and milk yield. We did this by calculating $R^2(\text{DGV, DRP})$ in validation Holstein and Jersey bulls. When using DRP for validation, the $R^2(\text{DRP, DGV})$ and the regression of DRP on DGV were calculated separately based on validation bulls with DTD (i.e., bulls with AUS daughters only) and based on all bulls with DRP (i.e., bulls with both AUS and OS daughters) to ensure that the results are not affected by the sample of bulls used. The second situation also allowed us to use a larger validation set because there were more bulls with DRP but no DTD (Table 1). Bulls born in 2004 or after were used for validation, as was the case when DTD was used for validation. In this case, for brevity, results from the use of DRP for validation bulls are presented based on DGV obtained from model 2 (DRP only), model 3 (DTD + DRP), and DRP component from model 5 (bivariate model).

RESULTS

Proportion of Variance in DTD or DRP Explained by SNP

For most traits analyzed in our study, the SD of DRP was higher than that of DTD when data from all bulls

were considered. However, for bulls that had both DTD and DRP the difference in mean and SD between their DTD and DRP was small (results not tabulated). In Holsteins, the proportion of variance explained by SNP varied from 0.30 (survival) to 0.77 (cell count), and was slightly higher for DTD than DRP (Table 3). For traits with low h^2 , such as fertility and survival, and traits subjectively scored, such as overall type (overall conformation score) and mammary system (overall udder score), the proportion of variance explained was lower (~ 0.4). The proportion of variance explained by the SNP for individual type traits such as pin set and stature was similar to that for milk yield traits (Table 3).

For Jerseys, the proportion of variance explained by the SNP was higher when DRP was used as a response variable than when DTD was used for fertility, survival, and all the type traits (Table 3). For most traits, particularly when DRP was used as response variable, the proportion of variance explained by the SNP was higher in Jerseys than in Holsteins.

Genetic Correlation Between DTD and DRP

Genetic correlations between DTD and DRP for most traits were close to 1.0 from model 5 (Table 4). The lowest genomic correlation was for mammary system (0.93) in Holstein and for survival (0.79) in Jersey when

model 5 was used. For model 5 some bulls had both DTD and DRP, so residual correlations were also estimated (Table 4). Residual correlations between DTD and DRP varied from as low as 0.61 for survival to as high as 0.94 for overall type in Holsteins. In Jerseys, these ranged from 0.36 for survival to 0.96 for pin set. However, the estimates for Jerseys were associated with larger sampling errors than those in Holsteins.

Genomic correlations between DTD and DRP from model 4 (i.e., where DRP was used only for bulls with no DTD) was the lowest for survival (0.72) and the highest (0.95) was for stature in Holstein. In Jersey, the lowest was for overall type (0.56) and the highest was for fertility (0.99; Table 4).

Validation in Bulls Using DTD

For milk production traits, $R^2(\text{DGV}, \text{DTD})$ and regression (DTD on DGV), adjusted R^2 , and expected reliabilities for validation Holstein and Jersey bulls are shown in Table 5. As the main objective of genomic prediction in AUS is to predict performance in AUS, and because correlations between DGV from DTD and DRP of the bivariate models were above 0.98 (e.g., for milk yield traits), for brevity results (i.e., DGV) based on DTD are presented for model 4 and 5.

Difference in R^2 between the models was relatively small in Holstein. For example, using milk yield in Hol-

Table 3. Total variance (variance) of daughter trait deviation (DTD) and deregressed proof (DRP) in Holstein and Jersey and proportion (Prop) explained by genomic relationship matrix (**G**)

Trait	DTD (model 1) ²			DRP (model 2) ²		
	No.	Total variance	Prop	No.	Total variance	Prop
Holstein						
Milk yield	3,501 ¹	243,630 $\pm$ 5,838	0.73 $\pm$ 0.02	8,882	315,230 $\pm$ 6,288	0.60 $\pm$ 0.01
Fat yield	3,501	372.5 $\pm$ 9.1	0.69 $\pm$ 0.02	8,882	517.7 $\pm$ 10.3	0.56 $\pm$ 0.01
Protein yield	3,501	140.1 $\pm$ 3.4	0.73 $\pm$ 0.02	8,882	180.2 $\pm$ 3.6	0.55 $\pm$ 0.01
Survival	3,419	33.4 $\pm$ 0.9	0.39 $\pm$ 0.03	8,379	50.3 $\pm$ 0.9	0.30 $\pm$ 0.02
Cell count	3,362	476.6 $\pm$ 15.2	0.77 $\pm$ 0.02	8,741	505.1 $\pm$ 10.7	0.67 $\pm$ 0.01
Fertility	3,456	186.2 $\pm$ 5.2	0.40 $\pm$ 0.03	8,439	240.1 $\pm$ 4.6	0.43 $\pm$ 0.02
Overall type	2,374	75.4 $\pm$ 2.5	0.43 $\pm$ 0.04	6,522	172.5 $\pm$ 3.6	0.36 $\pm$ 0.02
Mammary system	2,374	72.9 $\pm$ 2.4	0.43 $\pm$ 0.04	6,522	97.5 $\pm$ 2.1	0.47 $\pm$ 0.02
Stature	2,374	151.0 $\pm$ 5.3	0.54 $\pm$ 0.03	6,522	187.9 $\pm$ 4.4	0.62 $\pm$ 0.02
Pin set	2,296	195.6 $\pm$ 7.2	0.60 $\pm$ 0.4	6,522	195.9 $\pm$ 4.7	0.70 $\pm$ 0.01
Jersey						
Milk yield	920	228,570 $\pm$ 12,841	0.84 $\pm$ 0.05	2,609	246,520 $\pm$ 8,827	0.77 $\pm$ 0.02
Fat yield	920	390.1 $\pm$ 20.4	0.87 $\pm$ 0.03	2,609	405.2 $\pm$ 14.3	0.80 $\pm$ 0.02
Protein yield	920	149.0 $\pm$ 8.1	0.85 $\pm$ 0.04	2,609	163.3 $\pm$ 7.7	0.80 $\pm$ 0.02
Survival	860	58.2 $\pm$ 3.1	0.44 $\pm$ 0.06	1,956	71.2 $\pm$ 2.8	0.68 $\pm$ 0.03
Cell count	808	475.0 $\pm$ 29.7	0.88 $\pm$ 0.05	2,426	525.4 $\pm$ 19.5	0.78 $\pm$ 0.02
Fertility	776	103.7 $\pm$ 5.4	0.14 $\pm$ 0.06	2,363	102.2 $\pm$ 3.7	0.60 $\pm$ 0.03
Overall type	478	166.4 $\pm$ 12.3	0.54 $\pm$ 0.09	1,897	341.1 $\pm$ 14.1	0.79 $\pm$ 0.03
Mammary system	477	192.0 $\pm$ 14.3	0.59 $\pm$ 0.09	1,874	127.9 $\pm$ 5.2	0.77 $\pm$ 0.03
Stature	478	145.1 $\pm$ 11.2	0.62 $\pm$ 0.10	1,897	123.5 $\pm$ 5.2	0.74 $\pm$ 0.03
Pin set	478	115.1 $\pm$ 8.8	0.77 $\pm$ 0.11	1,897	91.9 $\pm$ 3.8	0.84 $\pm$ 0.03

¹To improve accuracy of estimation data of both reference and validation bulls were used.

²See Table 2 for model description.

Table 4. Genomic and residual correlation between daughter trait deviation (DTD) and deregressed proof (DRP) in Holstein and Jersey from model 5 and genomic correlation from model 4

Trait	Correlation		
	Model 5 (bivariate) ¹		Model 4 (bivariate) ²
	Genomic	Residual	Genomic
Holstein			
Milk yield	0.99 ± 0.01 ³	0.66 ± 0.01	0.92 ± 0.02
Fat yield	1.00 ± 0.01	0.68 ± 0.01	0.93 ± 0.02
Protein yield	0.99 ± 0.01	0.67 ± 0.01	0.88 ± 0.03
Survival	0.94 ± 0.02	0.61 ± 0.01	0.72 ± 0.07
Cell count	0.96 ± 0.03	0.73 ± 0.01	0.85 ± 0.02
Fertility	0.99 ± 0.01	0.76 ± 0.02	0.84 ± 0.05
Overall type	0.98 ± 0.02	0.94 ± 0.01	0.77 ± 0.05
Mammary system	0.93 ± 0.01	0.79 ± 0.02	0.79 ± 0.07
Stature	0.99 ± 0.01	0.91 ± 0.01	0.95 ± 0.03
Pin set	0.99 ± 0.01	0.87 ± 0.01	0.93 ± 0.03
Jersey			
Milk yield	0.99 ± 0.01	0.74 ± 0.03	0.90 ± 0.06
Fat yield	0.99 ± 0.01	0.72 ± 0.03	0.85 ± 0.05
Protein yield	0.99 ± 0.01	0.60 ± 0.02	0.90 ± 0.05
Survival	0.79 ± 0.06	0.36 ± 0.05	0.70 ± 0.11
Cell count	0.99 ± 0.01	0.69 ± 0.04	0.91 ± 0.06
Fertility	0.99 ± 0.01	0.55 ± 0.03	0.99 ± 0.02
Overall type	0.99 ± 0.01	0.96 ± 0.01	0.56 ± 0.13
Mammary system	0.84 ± 0.07	0.45 ± 0.10	0.77 ± 0.13
Stature	0.99 ± 0.01	0.96 ± 0.01	0.92 ± 0.07
Pin set	0.99 ± 0.01	0.96 ± 0.03	0.75 ± 0.07

¹Using both DTD and DRP when available and see Table 2 for model description.²Using DTD for all bulls with Australian daughters and DRP for overseas bulls with no Australian data.³Standard errors of correlations close to 1.0 are possibly underestimated.

stein when only DTD of AUS bulls were used to predict DGV of validation bulls (i.e., the current practice) the $R^2(\text{DGV}, \text{DTD})$ was 0.41, which increased to 0.43 (Table 5) when DRP was used (model 2). The number of bulls in the reference set when DTD (model 1) was used was 3,020, which increased to 6,963 when DRP (model 2) was used. When both DTD and DRP of bulls were used in a single-trait model (model 3) to predict DGV, the R^2 increased to 0.45. In this case, 7,234 bulls were included in the reference set (as an extra 270 bulls with DTD only were included). The $R^2(\text{DGV}, \text{DTD})$ for the 2 bivariate models that also used 7,234 bulls were only marginally better than that using DRP only (Table 5). In Jerseys, the use of DRP as a response variable (model 2) results in the highest $R^2(\text{DGV}, \text{DTD})$ for milk yield traits (Table 5). Also for milk yield traits, the regression of DTD on DGV was close to 1.0 for most models for both breeds. However, the regression was slightly below 1.0 (DGV overestimated DTD) for fat yield in Jersey (Table 5) when model 2 was used. The adjusted R^2 for milk and fat yield were close to the expected reliability for Holstein. However, they were higher than the expected reliability for protein yield in Holstein and for milk and protein yield in Jersey (Table 5).

For functional traits, the $R^2(\text{DGV}, \text{DTD})$ and adjusted R^2 were considerably lower than for milk yield traits for both breeds (Tables 5 and 6), except for cell count. Among the models in Holstein, model 3 gave slightly higher $R^2(\text{DGV}, \text{DTD})$ than the other models for all traits except for cell count. For SCC, the bivariate models gave slightly higher $R^2(\text{DGV}, \text{DTD})$ than the single-trait models (Table 6). In Jersey, there was little difference among models in their ability to predict DGV of validation bulls for functional traits (Table 6). In general, the expected reliabilities increased when Jersey bulls with DRP were added into the reference set by about 7% (i.e., models 2, 3, 4, and 5) compared with model 1 (Tables 5 and 6). Also, for survival and fertility for both breeds, the adjusted R^2 were lower than the expected reliabilities. For cell count in Holstein, adjusted R^2 were higher than expected reliabilities by up to 0.18 in some cases (Table 6). For fertility and survival, the regression values were all below 1.0 in Jersey, particularly when model 2 was used for prediction; however, the standard errors of these coefficients were large.

Among type traits, the $R^2(\text{DGV}, \text{DTD})$ was the highest for pin set, the lowest for overall type and mammary system, and intermediate for stature (Table 7).

Table 5. Reliability of prediction [expressed as R^2 (direct genomic values, daughter trait deviations)]¹ and regression (b) of daughter trait deviations (DTD) on direct genomic values (DGV) in validation bulls and adjusted R^2 (adj. R^2)² and expected reliability (Exp. Rel.) for milk yield traits from different models

Trait	Holstein				Jersey			
	R^2	b	Adj. R^2	Exp. Rel.	R^2	b	Adj. R^2	Exp. Rel.
Milk yield								
Model 1 (DTD)	0.41	1.02 ± 0.05	0.52	0.47	0.33	1.30 ± 0.14	0.41	0.31
Model 2 (DRP) ³	0.43	0.91 ± 0.04	0.54	0.54	0.40	0.93 ± 0.09	0.50	0.38
Model 3 (DTD + DRP_Univ) ⁴	0.45	0.97 ± 0.05	0.57	0.56	0.38	1.25 ± 0.12	0.48	0.35
Model 4 (DTD and DRP_D_DTD) ⁵	0.44	1.03 ± 0.05	0.55	0.53	0.38	1.34 ± 0.13	0.47	0.37
Model 5 (DTD and DRP_M_DTD) ⁵	0.44	0.98 ± 0.05	0.55	0.56	0.39	1.41 ± 0.13	0.48	0.37
Fat yield								
Model 1 (DTD)	0.35	1.08 ± 0.06	0.45	0.44	0.30	1.27 ± 0.15	0.37	0.33
Model 2 (DRP) ³	0.37	1.00 ± 0.06	0.47	0.51	0.35	0.85 ± 0.09	0.44	0.41
Model 3 (DTD + DRP_Univ) ⁴	0.40	1.07 ± 0.06	0.51	0.53	0.34	1.24 ± 0.13	0.42	0.37
Model 4 (DTD and DRP_D_DTD) ⁵	0.40	1.13 ± 0.06	0.51	0.52	0.35	1.39 ± 0.14	0.44	0.38
Model 5 (DTD and DRP_M_DTD) ⁵	0.39	1.07 ± 0.06	0.50	0.54	0.32	1.28 ± 0.14	0.40	0.40
Protein yield								
Model 1 (DTD)	0.42	1.05 ± 0.05	0.57	0.47	0.42	1.12 ± 0.10	0.52	0.33
Model 2 (DRP) ³	0.41	1.01 ± 0.05	0.55	0.52	0.54	1.00 ± 0.07	0.67	0.41
Model 3 (DTD + DRP_Univ) ⁴	0.44	1.01 ± 0.05	0.58	0.54	0.49	1.15 ± 0.09	0.62	0.38
Model 4 (DTD and DRP_D_DTD) ⁵	0.42	1.04 ± 0.05	0.57	0.51	0.49	1.22 ± 0.09	0.61	0.36
Model 5 (DTD and DRP_M_DTD) ⁵	0.42	0.98 ± 0.05	0.56	0.54	0.52	1.24 ± 0.09	0.65	0.40

¹Approximate SE of the R^2 is 0.04 and 0.07 for Holstein and Jersey, respectively.

²Calculated assuming average DTD reliability of 0.80 for milk yield and 0.78 for fat and protein yield for Holstein and 0.80 for all milk yield traits for Jersey.

³Deregressed proofs and see Table 2 for detailed model descriptions.

⁴DGV calculated using DTD for Australian and DRP for overseas bulls.

⁵For the bivariate models, the DGV that corresponds to DTD is used.

The discrepancy between the expected reliability and the adjusted R^2 (DGV,DTD) was also the highest for overall type and mammary system. For type traits, the use of DGV from DTD (models 1 and 3) or from DTD component of the bivariate models gave regression values closer to 1.0 compared with the use of DGV from DRP (i.e., model 2). For Jerseys, R^2 (DGV,DTD) for the other 3 type traits (mammary system, overall type, and stature) from all the models were effectively zero and the number of validation bulls was only 38, so these results were not tabulated.

Effect of Relatedness of Validation Animals with the Current Reference Set on R^2

The R^2 (DGV,DTD) for validation bulls least related to the reference set was lower than for the validation bulls most related for all traits except fat yield (Figure 1) in Holstein. For model 1, the increase in R^2 comparing least-related to most-related bulls varied from 3 (cell count) to 19% (milk yield). When DRP (model 2) was used as response variable instead of DTD (model 1), the R^2 (DTD,DGV) showed more increase in the least-related bulls category than in the most-related bulls category (Figure 1) for all traits except cell count.

In the least-related bulls, the R^2 averaged over 6 traits increased by 10% from 0.30 (model 1) to 0.33 (DTD plus DRP, model 3) compared with a 6% increase from 0.44 to 0.47 in the most related bulls. Similar results were observed in Jersey validation bulls, Figure 2.

When cows were used for validation in both breeds the inclusion of OS information either using single-trait models (models 2 and 3) or bivariate models (models 4 and 5) increased R^2 (DGV,TD) compared with model 1 (Figures 3 and 4). Averaged over the 3 milk yield traits, the relationship between relatedness and R^2 (DGV,TD) of cows was more in line with the expectation in Jersey than in Holstein when DTD (model 1) were used as response variable, for example.

Validation in Bulls Using DRP as Criteria for Selected Traits

Results using DRP for validation are presented in Tables 8 (Holstein) and 9 (Jersey) for traits where R^2 (DTD,DGV) were low (survival, fertility, overall type, and mammary system) and for comparison purposes (e.g., milk yield and pin set). In the case of survival, the use of DRP for validation instead of DTD did not show any increase in R^2 (i.e., Tables 6 and 8).

Table 6. Reliability of prediction [expressed as R^2 (direct genomic value, daughter yield deviation)]¹ and regression (b) of daughter yield deviation (DTD) on direct genomic value (DGV) in validation bulls and adjusted R^2 (adj. R^2)² and expected reliability (Exp. Rel.) for functional traits from different models

Trait	Holstein				Jersey			
	R^2	b	Adj. R^2	Exp. Rel.	R^2	B	Adj. R^2	Exp. Rel.
Survival								
Model 1 (DTD)	0.11	0.90 ± 0.12	0.20	0.40	0.03	0.84 ± 0.41	0.06	0.21
Model 2 (DRP) ³	0.11	1.01 ± 0.13	0.20	0.45	0.05	0.78 ± 0.31	0.10	0.25
Model 3 (DTD + DRP_Univ) ⁴	0.16	1.08 ± 0.11	0.29	0.42	0.03	0.51 ± 0.26	0.06	0.37
Model 4 (DTD and DRP_D_DTD) ⁵	0.11	0.98 ± 0.13	0.20	0.36	0.02	0.49 ± 0.32	0.04	0.26
Model 5 (DTD and DRP_M_DTD) ⁶	0.13	1.03 ± 0.12	0.24	0.39	0.02	0.50 ± 0.28	0.04	0.36
Cell count								
Model 1 (DTD)	0.49	0.98 ± 0.04	0.69	0.54	0.27	0.92 ± 0.11	0.41	0.40
Model 2 (DRP) ³	0.48	0.83 ± 0.04	0.68	0.58	0.30	0.71 ± 0.08	0.45	0.46
Model 3 (DTD + DRP_Univ) ⁴	0.52	0.93 ± 0.04	0.73	0.60	0.31	0.88 ± 0.10	0.46	0.50
Model 4 (DTD and DRP_D_DTD) ⁵	0.54	1.10 ± 0.04	0.75	0.57	0.29	0.96 ± 0.11	0.43	0.45
Model 5 (DTD and DRP_M_DTD) ⁵	0.53	1.15 ± 0.05	0.75	0.58	0.29	0.94 ± 0.11	0.43	0.47
Fertility								
Model 1 (DTD)	0.19	1.18 ± 0.11	0.35	0.38	0.01	0.72 ± 0.58	0.02	0.21
Model 2 (DRP) ³	0.22	1.01 ± 0.08	0.41	0.47	0.02	0.47 ± 0.28	0.04	0.37
Model 3 (DTD + DRP_Univ) ⁴	0.22	0.94 ± 0.08	0.40	0.48	0.02	0.56 ± 0.30	0.04	0.35
Model 4 (DTD and DRP_D_DTD) ⁵	0.19	1.08 ± 0.10	0.34	0.39	0.02	0.89 ± 0.63	0.04	0.25
Model 5 (DTD and DRP_M_DTD) ⁵	0.21	0.96 ± 0.08	0.38	0.45	0.02	0.52 ± 0.33	0.04	0.37

¹Approximate SE of the R^2 is 0.04 for Holstein and 0.08 for Jersey.

²Calculated assuming average DTD reliability of 0.50, 0.71 and 0.55 for survival, cell count, and fertility, respectively, in Holstein. and 0.4, 0.67 and 0.5 for survival, cell count, and fertility, respectively, in Jersey.

³Deregressed proofs and see Table 2 for detailed model descriptions.

⁴DGV calculated using DTD for Australian and DRP for overseas bulls.

⁵For the bivariate models, the DGV that corresponds to DTD is used.

The biggest advantage of the use of DRP for prediction both as response and validation variable was observed for fertility (Tables 6 and 8) in Holstein.

The use of DRP to predict DGV and then validate using DRP [i.e., R^2 (DRP,DGV)] was also better for mammary system than for overall type. The prediction for all bulls (i.e., bulls with AUS and OS daughters) was better than predicting for bulls with only AUS daughters. In the case of pin set, the R^2 between DRP and DGV in Table 8 is comparable to those in Table 7. For pin set, the use of DRP to predict DGV and the use of DRP to validate was even better when bulls with both AUS and OS daughters were considered for validation.

In Jersey, the use of DRP of milk yield for both prediction and validation increased R^2 by 8 to 14%, depending on the model (Table 9), compared with when DTD was the predicted variable (Table 5). For Jersey, prediction of DGV of milk yield for bulls with AUS daughters resulted in higher R^2 than for all bulls (i.e., AUS and OS daughters). The biggest advantage of the use of DRP for prediction and validation was for fertility and survival when predicting DGV for all bulls (bulls with both AUS and OS daughters; Table 9). In the case of pin set, the R^2 (DRP,DGV) in Table 9 was about twice the R^2 (DTD,DGV) in Table 7.

DISCUSSION

Proportion of Variance of DTD and DRP Explained by SNP

The proportions of variance explained by **G** in DTD and DRP of Holstein bulls observed in our study (Table 3) are lower than those reported by Jensen et al. (2012; in Nordic Holsteins) for most traits. The biggest differences between their results and ours are for low h^2 traits, such as fertility. Part of the reason for this may be associated with reliability of the phenotype (e.g., DTD) used. In AUS, the number of effective daughters per bulls is quite low compared with Nordic countries (Jensen et al., 2012). However, differences in the methods used to calculate **G** may also be an additional source of differences.

The proportion of variance explained by **G** in DTD in the current study, mainly for milk yield traits, was lower than those estimated by Haile-Mariam et al. (2013b). This may be related to the increase in the effective population size due to the addition of bulls with a diverse genetic background since the analyses by Haile-Mariam et al. (2013b) was performed. For example, 140 Red and 255 NZ Holstein bulls with daughters in AUS but first tested in the NLD and NZL, respectively, were added to the current reference set. de Roos et al. (2008)

Table 7. Reliability of prediction [expressed as $R^2(\text{direct genomic value, daughter trait deviation})^1$, regression (b) of daughter trait deviation (DTD) on direct genomic value (DGV), adjusted R^2 (adj. R^2)² and expected reliability (Exp. Rel.) for type trait from different models in validation Holstein and Jersey bulls

Trait	Validation bulls			
	R^2	b	Adj. R^2	Exp. Rel.
Overall type in Holstein				
Model 1 (DTD) ³	0.10	0.85 ± 0.11	0.19	0.32
Model 2 (DRP) ⁴	0.05	0.75 ± 0.08	0.11	0.36
Model 3 (DTD + DRP_Univ) ⁵	0.09	0.70 ± 0.08	0.17	0.43
Model 4 (DTD and DRP_D_DTD) ⁶	0.08	0.86 ± 0.08	0.15	0.38
Model 5 (DTD and DRP_M_DTD) ⁶	0.08	0.81 ± 0.08	0.14	0.37
Mammary system in Holstein				
Model 1 (DTD) ³	0.10	0.96 ± 0.11	0.19	0.33
Model 2 (DRP) ⁴	0.06	0.73 ± 0.08	0.13	0.39
Model 3 (DTD + DRP_Univ) ⁵	0.08	0.79 ± 0.08	0.16	0.42
Model 4 (DTD and DRP_D_DTD) ⁶	0.08	0.86 ± 0.08	0.16	0.42
Model 5 (DTD and DRP_M_DTD) ⁶	0.05	0.81 ± 0.08	0.14	0.39
Stature in Holstein				
Model 1 (DTD) ³	0.17	0.83 ± 0.11	0.36	0.30
Model 2 (DRP) ⁴	0.17	0.76 ± 0.08	0.40	0.30
Model 3 (DTD + DRP_Univ) ⁵	0.23	0.79 ± 0.08	0.47	0.41
Model 4 (DTD and DRP_D_DTD) ⁶	0.22	0.86 ± 0.08	0.46	0.40
Model 5 (DTD and DRP_M_DTD) ⁶	0.22	0.81 ± 0.08	0.46	0.40
Pin set in Holstein				
Model 1 (DTD) ³	0.20	0.96 ± 0.11	0.36	0.38
Model 2 (DRP) ⁴	0.22	0.76 ± 0.08	0.40	0.48
Model 3 (DTD + DRP_Univ) ⁵	0.26	0.79 ± 0.08	0.47	0.54
Model 4 (DTD and DRP_D_DTD) ⁶	0.26	0.86 ± 0.08	0.46	0.50
Model 5 (DTD and DRP_M_DTD) ⁶	0.25	0.81 ± 0.08	0.46	0.52
Pin set in Jersey				
Model 1 (DTD) ³	0.19	0.78 ± 0.36	0.35	0.30
Model 2 (DRP) ⁴	0.32	1.13 ± 0.27	0.58	0.37
Model 3 (DTD + DRP_Univ) ⁵	0.30	1.18 ± 0.29	0.54	0.40
Model 4 (DTD and DRP_D_DTD) ⁶	0.25	0.95 ± 0.26	0.45	0.38
Model 5 (DTD and DRP_M_DTD) ⁶	0.29	1.33 ± 0.26	0.52	0.39

¹Approximate SE of the R^2 is 0.06 for Holstein and 0.16 for Jersey.

²Calculated assuming average DTD reliability of 0.50, 0.50, 0.55 and 0.55 for overall type, mammary system, stature, and pin set, respectively, in Holstein and 0.55 for pin set in Jersey.

³DTD = daughter trait deviation and see Table 2 for model description.

⁴DRP = deregressed breeding value.

⁵DGV calculated using DTD for Australian and DRP for overseas bulls.

⁶For the bivariate models the DGV that corresponds to DTD is used.

have shown that the linkage disequilibrium phase (i.e., association of markers with QTL) of NZL Holstein bulls differs from that of other Holstein populations (e.g., AUS). Reanalyses of the milk and fat yield DTD of bulls with AUS daughters excluding those first tested in NZL increased the proportion of variance explained by **G** from 0.73 and 0.69 (Table 3) to 0.77 and 0.72, respectively (data not shown). However, excluding Red Holstein bulls or further selection of bulls based on the number of effective daughters (i.e., more reliable bulls) did not increase the proportion of variance explained by **G**. When analyzing DRP of milk yield, excluding bulls first tested in NZL increased the proportion of variance explained by **G** to 0.69 (data not shown) compared with 0.60 (Table 3). Both the lower reliability of most OS bulls with DRP on the AUS scale and that they are of diverse genetic background may have contributed to the low proportion of variance explained by **G** in DRP.

Proportion of variance explained by **G** for type traits such as overall type and mammary system may be lower than that for pin set, for example, because of the introduction of some changes in the definition of these traits over the years in AUS (e.g., Battagin et al., 2013). The proportion of variance explained by SNP for milk yield traits and cell count for Jersey when DTD was used as a response variable were generally higher than for Holstein (Table 3), perhaps reflecting a reduced effective population size for this breed (e.g., Hayes et al., 2009).

Genetic Correlation Between DTD and DRP

The genetic correlation between DTD and DRP for milk yield traits in both Holstein and Jersey was effectively 1.0, suggesting the deregression worked as expected. The low genetic correlations for traits such

Figure 1. Squared correlation between direct genomic value and daughter trait deviation (DTD) in validation Holstein bulls which are the least-related (i.e., mean to top 10 bulls) and most-related groups to the Australian genomic reference set for different traits from different models (see Table 2 for description of the models). DRP = deregressed breeding value. Color version available online.

as survival, overall type, and mammary system in both Holstein and Jersey when model 4 was used may reflect difference in the definition of traits or in factors that may influence the traits. Clustering analyses of data on overall type and mammary system submitted to Interbull has shown that AUS bulls were different from those in other Interbull member countries (Battagin et al., 2012). Battagin et al. (2013) also showed that the genetic correlation between AUS and other Interbull member countries decreased over a 10-yr period for overall type and the cluster to which AUS belonged changed over the period.

Validation in Bulls Using DTD

The observed $R^2(\text{DGV}, \text{DTD})$ in validation bulls is associated with the proportion of variance explained by $\mathbf{G}$. This is in agreement, for example, with Daetwyler et al. (2012), who reported a positive relationship between the product of reference size and genomic h^2 (proportion of variance explained by $\mathbf{G}$), and the accuracy of genomic prediction. In our study, the increase in R^2 due to the increase in the reference size as the result of including DRP was not large when one considers that the doubling of the reference size increased R^2 from 0.41

Figure 2. Squared correlation between direct genomic value and daughter trait deviation (DTD) in validation Jersey bulls which are the least-related (i.e., mean to top 10 bulls) and most-related groups to the Australian genomic reference set for different traits from different models (see Table 2 for description of the models). DRP = deregressed breeding value. Color version available online.

to 0.45 (Table 5) for milk yield when validation was based on DTD. However, these results are consistent with the results from other studies (e.g., Habier et al., 2011) where it was shown that increases in R^2 due to increases in reference population size is slow once the reference size has reached 3,000 to 4,000 bulls. It may also reflect the effect of $G \times E$, such that information from bulls with OS daughters is of less value than bulls with AUS daughters to predict performance in AUS. That the R^2 for survival, overall type, and mammary system are relatively low and the discrepancy between the expected reliability and the adjusted R^2 for these traits are high suggests that this could be due to inconsistency in the data used for training and validation and the low overall reliability of the response variables used (Battaglin et al., 2012, 2013). For example, the ge-

netic correlation for mammary system and overall type of cows classified before 2007 and in 2007 and after (in 2007 the type classification scheme was modified by Holstein Australia) was low at 0.64 and 0.71, respectively. The genetic correlation for stature between the 2 periods was higher at 0.86 but still lower than that for pin set, which was effectively 1 (0.99), as would be expected (M. Haile-Mariam, unpublished data). However, some level of discrepancy between expected reliability and adjusted R^2 may be due to the fact that $R^2(DGV, DTD)$ is measured on a selected subset of the population, whereas the expected reliability is relevant to the whole population (Bijma, 2012).

In Jerseys, the increase in R^2 from the uses of DTD where only AUS bulls (i.e., 784 bulls) are included in the reference set to DRP where both AUS and OS bulls

Figure 3. Effect of relatedness (mean to top 10 bulls in the reference set) of validation Holstein cows to the Australian reference on squared correlation between direct genomic value and trait deviation for the average of milk yield traits from different models (see Table 2 for description of the models). DTD = daughter trait deviation; DRP = deregressed breeding value. Color version available online.

(i.e., 2,000 bulls) are used is consistent with the increase in the reference population. Relative to the reference populations size, the R^2 (DTD,DGV) observed is generally higher for Jersey than for Holstein which is in agreement to previous studies (Hayes et al., 2009; Erbe et al., 2012). This likely reflects the fact that the AUS Jersey bull reference set was quite small to begin with and that the Jersey bulls that were added were all from NZL, where the production environment is reasonably similar to that in AUS (i.e., mainly grazing systems). In the case of Holsteins, the OS bulls that were added were tested in diverse production environments, as production environments in NZL and the NLD are quite different.

Effect of Relatedness of Validation Animals with the Current Reference Set on R^2

The fact that the R^2 (DTD,DGV) was higher for most traits in the most-related bull group than the least-

related bull group is in agreement with the literature (Habier et al., 2010; Daetwyler et al., 2013). The exception in our results was for fat yield in Holstein, in which the R^2 (DTD,DGV) was actually higher in the least-related bull group compared with the most-related. Conversely, this is consistent with Erbe et al. (2012), who found that the accuracy of DGV of bulls with their sire in the reference set is lower than bulls with no sire, suggesting that carrying or not carrying the DGAT1 mutation (Grisart et al., 2004) rather than relatedness may have a bigger influence on accuracy.

The relationship between the R^2 and the degree of relatedness of cows with the current reference bulls in Figure 4 in Jersey for milk yield traits is similar to the observation by Clark et al. (2012). The inclusion of more bulls into the reference set (i.e., OS bulls) by using their DRP narrowed the difference in R^2 between the most-related and least-related bulls or cows. The narrowing of the difference in R^2 between most-related and least-related candidate animals was more obvious

Figure 4. Effect of relatedness (mean to top 10 bulls) of validation Jersey cows to the Australian reference set on squared correlation between direct genomic value and trait deviation for the average of milk yield traits from different models (see Table 2 for description of the models). DTD = daughter trait deviation; DRP = deregressed breeding value. Color version available online.

when cows with diverse relationships to the bulls in the reference set were used, rather than the validation bulls. For example, when Holstein cows were used as validation animals for milk yield traits the average relationship to the new reference set that included OS bulls increased by 8, 9, 3, 7, and 1%, for the <0.15 , ≥ 0.15 and <0.20 , ≥ 0.20 and <0.25 , ≥ 0.25 and <0.30 , and ≥ 0.30 groups, respectively, showing the narrowing of the difference between groups.

Validation in Bulls Using DRP as Criteria for Selected Traits

In this study we observed that the use of DRP as response and validation variables improved R^2 more than using DTD for most traits and in both breeds, possibly due to the increase in reference size when DRP is used.

In many other countries, DRP is used as a response variable for genomic prediction. In AUS, DTD are used because they are considered to be more independent and accurate measures of phenotypic performance of a bull's daughters, though DRP may contain additional (VanRaden and Wiggans, 1991) source of information (e.g., grandsons and granddaughters). The DTD include both genetic and residual variations, and the latter may include major gene effects, which may not be fully included in DRP (Nieuwhof et al., 2010). Another practical advantage of the use of DTD is that it can be calculated relatively easily compared with DRP, where its calculation is difficult particularly for situations where a less complete pedigree exists, for example.

We are not aware of studies that compared DTD and DRP as response variables for genomic prediction. However, Thomsen et al. (2001) did not find a

Table 8. Reliability of prediction [expressed as R^2 (direct genomic value, deregressed breeding value)] and regression (b) of deregressed breeding value (DRP) on direct genomic value (DGV) in validation Holstein bulls

Trait	Prediction for bulls with DTD only		Prediction for all bulls with DRP	
	n = 513		n = 1574	
	R^2	b	R^2	b
Milk yield				
Model 2 (DRP used to predict DRP)	0.45	1.06 ± 0.05	0.39	0.95 ± 0.03
Model 3 (DRP and DTD used to predict DRP) ¹	0.47	1.00 ± 0.05	0.41	1.01 ± 0.03
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.46	1.00 ± 0.05	0.41	0.95 ± 0.03
Survival	n = 471		n = 1,479	
Model 2 (DRP used to predict DRP)	0.09	0.94	0.03	0.85
Model 3 (DRP and DTD used to predict DRP) ¹	0.15	1.07	0.09	0.95
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.12	0.85	0.06	0.71
Fertility	n = 471		n = 1,479	
Model 2 (DRP used to predict DRP)	0.35	1.13 ± 0.07	0.27	1.1 ± 0.04
Model 3 (DRP and DTD used to predict DRP) ¹	0.34	1.05 ± 0.06	0.28	1.04 ± 0.04
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.33	1.05 ± 0.07	0.27	1.02 ± 0.04
Overall type	n = 177		n = 1,552	
Model 2 (DRP used to predict DRP)	0.05	0.54 ± 0.17	0.18	1.02 ± 0.06
Model 3 (DRP and DTD used to predict DRP) ¹	0.11	0.80 ± 0.17	0.19	1.09 ± 0.06
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.08	0.55 ± 0.14	0.20	0.98 ± 0.05
Mammary system				
Model 2 (DRP used to predict DRP)	0.15	0.71 ± 0.12	0.30	0.95 ± 0.04
Model 3 (DRP and DTD used to predict DRP) ¹	0.18	0.88 ± 0.14	0.30	1.01 ± 0.04
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.16	0.64 ± 0.11	0.30	0.92 ± 0.03
Pin set				
Model 2 (DRP used to predict DRP)	0.21	0.73 ± 0.11	0.36	0.96 ± 0.03
Model 3 (DRP and DTD used to predict DRP) ¹	0.23	0.74 ± 0.10	0.36	0.97 ± 0.03
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.23	0.76 ± 0.10	0.35	1.01 ± 0.04

¹DGV calculated using daughter trait deviation (DTD) for Australian and DRP for overseas bulls.²For the bivariate models, the DGV that corresponds to DRP is used.

difference between using DTD and DRP as a response variable for detecting QTL. In the current study, the use of DRP as both response and validation variables was clearly advantageous in 3 cases. For fertility, the use of DRP was better than DTD in Holstein because the R^2 (DGV,DTD) using DRP as a response variable is even slightly higher than R^2 (DGV,DTD) using DTD as a response variable (Table 5). This may be because fertility BV in AUS is predicted from multitrait model that includes fertility traits and lactation length (Haile-Mariam et al., 2013a) and the resulting DRP may capture some of this information better than DTD, which is based on calving interval only. This was also observed by Haile-Mariam et al. (2014), where comparisons were made using the same number of reference bulls for both variables in Holstein. For all traits in Jerseys the increase in reliability as a result of using DRP justifies its use as response variable for genomic prediction because all the bulls included in the reference set were from NZL, where the production environments is similar to AUS and the number of bulls without DRP were fewer than 60 and the current AUS reference set is small (~800 bulls). Also, the use of DRP could be advantageous even for some bulls with AUS daughters if the majority of their daughters performed in another

country. Third, for overall type and mammary system where there was some change in the definition of traits over the years in AUS (e.g., Battagin et al., 2013) and consequently the genetic correlation between DTD and DRP was low, the use of DRP from both AUS and OS daughters results in higher R^2 because the OS data add more reliable information. Although high R^2 by itself may not be the only justification for choosing a response variable for genomic prediction, it is still advantageous to select a response variable that can be predicted with high degree of accuracy provided it is close enough to the true BV.

For survival, that proportion of variance explained by **G** in DRP was higher than that in DTD (Table 3), which resulted in higher R^2 (DRP,DGV) in Jersey than in Holstein despite the big difference in reference size between the 2 breeds. This may be because of the similarity in the definition of survival or factors affecting survival in Jersey (NZL and AUS) compared with that in Holstein, where genetic evaluation for survival is from survival analyses (Vollema and Groen,1998) in NLD (where most of the OS Holstein data originates) and linear models in AUS. Also, survival in a seasonal calving system in AUS is largely influenced by fertility (Haile-Mariam et al., 2013a) but not in NLD, which

Table 9. Reliability of prediction [expressed as R^2 (direct genomic value, deregressed breeding value)] and regression (b) of deregressed breeding value (DRP) on direct genomic value (DGV) in validation Jersey bulls

Trait	Prediction for bulls with DTD only		Prediction for all bulls with DRP	
	n = 173		n = 609	
	R^2	b	R^2	B
Milk yield				
Model 2 (DRP used to predict DRP) ¹	0.53	1.20 ± 0.09	0.44	1.11 ± 0.05
Model 3 (DRP and DTD used to predict DRP) ¹	0.50	1.26 ± 0.10	0.42	1.17 ± 0.06
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.51	1.17 ± 0.09	0.42	1.10 ± 0.05
Survival		n = 111		n = 344
Model 2 (DRP used to predict DRP)	0.24	1.67 ± 0.28	0.19	1.48 ± 0.16
Model 3 (DRP and DTD used to predict DRP) ¹	0.17	1.15 ± 0.24	0.22	1.10 ± 0.11
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.17	0.88 ± 0.18	0.20	0.84 ± 0.09
Fertility		n = 135		n = 426
Model 2 (DRP used to predict DRP)	0.10	0.82 ± 0.21	0.26	1.22 ± 0.09
Model 3 (DRP and DTD used to predict DRP) ¹	0.10	0.90 ± 0.23	0.25	1.16 ± 0.08
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.08	0.64 ± 0.18	0.25	1.01 ± 0.07
Overall type		n = 38		n = 421
Model 2 (DRP used to predict DRP)	0.01	0.34 ± 0.17	0.22	0.54 ± 0.06
Model 3 (DRP and DTD used to predict DRP) ¹	0.02	0.48 ± 0.23	0.21	0.69 ± 0.07
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.01	0.35 ± 0.17	0.22	0.57 ± 0.05
Mammary system				
Model 2 (DRP used to predict DRP)	0.01	0.37 ± 0.24	0.22	0.64 ± 0.06
Model 3 (DRP and DTD used to predict DRP) ¹	0.0	0.38 ± 0.22	0.18	0.51 ± 0.05
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.01	0.46 ± 0.22	0.23	0.65 ± 0.06
Pin set				
Model 2 (DRP used to predict DRP)	0.40	1.46 ± 0.29	0.30	1.00 ± 0.07
Model 3 (DRP and DTD used to predict DRP) ¹	0.39	1.54 ± 0.31	0.31	1.02 ± 0.08
Model 5 (DRP and DTD Bi-DRP used to predict DRP) ²	0.36	1.32 ± 0.28	0.30	0.93 ± 0.07

¹DGV calculated using daughter trait deviation (DTD) for Australian and DRP for overseas bulls.

²For the bivariate models, the DGV that corresponds to DRP is used.

may have contributed to the low R^2 (DRP,DGV) observed in Holstein.

CONCLUSIONS

Our results demonstrate that the reference population for predicting DGV in a country can be expanded by including DRP from bulls with OS daughters. The R^2 (DTD,DGV) was increased for both Holstein and Jersey validation bulls when NLD and NZL bulls were added to the reference set, even though the $G \times E$ may be larger than between many other countries. Averaged over the 6 main traits (3 milk yield traits, survival, fertility, and cell count), in Holstein the R^2 (DGV,DTD) was the highest when DTD and DRP were used together as response variables (model 3) but in Jersey it was the highest when DRP (model 2) was used. In addition, validation using bulls and cows showed that the difference in R^2 of DGV between least-related and most-related animals can be narrowed by the addition of more bulls into the reference set by using their DRP in a bivariate or single-trait model. The proportion of variance explained by G in the DTD and DRP for different traits is related to the ability to predict the DGV, showing that choice of the response variable to

use for genomic prediction and the extent of achievable accuracy can be gauged by calculating it. For fertility for both breeds, the proportion of variance explained by G was the highest when DRP was used and the R^2 (DRP,DGV) was higher than R^2 (DTD,DGV). In Jersey, using DRP both as the response variable and as the validation criteria [e.g., R^2 (DGV,DRP)] generally gave significantly higher R^2 than when the DTD was used as the response variable. This, coupled with the fact that all OS bulls that were added into the reference set were from NZL, where production environment is similar to AUS, means there is an advantage to be gained from using DRP as response and validation variable.

ACKNOWLEDGMENTS

Australian Dairy Herd Improvement Scheme (AD-HIS) provided that data used for this study. We thank Gert Nieuwhof and Kon Konstantinov (Melbourne, Australia) for fruitful discussion. The Cooperative Cattle Improvement Organization (CRV, Arnhem, the Netherlands) provided the genotype data. We also thank Zengting Liu (VIT, Verden, Germany) for providing his deregression software and assistance in the implementation of the software.

REFERENCES

- Amer, P. R., and G. Banos. 2010. Implications of avoiding overlap between training and testing data sets when evaluating genomic predictions of genetic merit. *J. Dairy Sci.* 93:3320–3330.
- Battagin, M., F. Forabosco, J. H. Jakobsen, M. Penasa, T. J. Lawlor, and M. Cassandro. 2012. International genetic evaluation of Holstein bulls for overall type traits and body condition score. *J. Dairy Sci.* 95:4721–4731.
- Battagin, M., F. Forabosco, M. Penasa, and M. Cassandro. 2013. Cluster analysis on across-country genetic correlations for overall type traits and body condition score of Holstein bulls. *Livest. Sci.* 151:108–114.
- 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.
- Brøndum, R. F., E. Rius-Vilarrasa, I. Strandén, G. Su, B. Guldbrandtsen, W. F. Fikse, and M. S. Lund. 2011. Reliabilities of genomic prediction using combined reference data of the Nordic Red dairy cattle populations. *J. Dairy Sci.* 94:4700–4707.
- 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.
- Daetwyler, H. D., M. P. L. Calus, R. Pong-Wong, G. de los Campos, and J. M. Hickey. 2013. Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics* 193:347–365.
- Daetwyler, H. D., A. A. Swan, J. H. J. van der Werf, and B. J. Hayes. 2012. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. *Genet. Sel. Evol.* 44:33.
- Daetwyler, H. D., B. Villanueva, and J. A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE* 3:e3395.
- de Roos, A. P. W., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase on Holstein-Friesian, Jersey, and Angus cattle. *Genetics* 179:1503–1512.
- 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.
- Garrick, D. J., J. F. Taylor, and R. L. Fernando. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* 41:55.
- Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml User Guide Release 3.0 VSN International Ltd., Hemel Hempstead, UK.
- Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximization of long term response. *Genetica* 136:245–257.
- Grisart, B., F. Farnir, L. Karim, N. Cambisano, J. J. Kim, A. Kvasz, M. Mni, P. Simon, J. M. Frere, W. Coppieters, and M. Georges. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc. Natl. Acad. Sci. USA* 101:2398–2403.
- 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.
- Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12:186.
- Haile-Mariam, J., E. Pryce, and B. Hayes. 2014. Analyses of functional traits with indicator traits to improve genomic reliability in Australian Holstein cattle using domestic daughter trait deviation or de-regressed proofs as response variable. Accessed Nov. 2014. https://asas.org/docs/default-source/wcgalp-proceedings-oral/298_paper_10023_manuscript_1727_0.pdf.
- Haile-Mariam, M., P. J. Bowman, and J. E. Pryce. 2013a. Genetic analyses of fertility in Holstein herds with low and high mean calving intervals and in Jersey herds. *J. Dairy Sci.* 96:655–667.
- Haile-Mariam, M., G. J. Nieuwhof, K. T. Beard, K. V. Konstantinov, and B. J. Hayes. 2013b. Comparison of heritabilities of dairy traits in Australian Holstein-Friesian cattle from genomic and pedigree data and implications for genomic evaluations. *J. Anim. Breed. Genet.* 130:20–31.
- Hayes, B. J., P. J. Bowman, A. C. Chamberlain, K. Verbyla, and M. E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:51.
- Interbull. 2013. Appendix I. Sire standard deviations in diagonal and genetic correlations below diagonal. Accessed Oct. 2013. http://www.interbull.org/static/mace_evaluations_archive/eval/prod-appen-1-141.pdf.
- Jensen, J., G. Su, and P. Madsen. 2012. Partitioning additive genetic variance into genomic and remaining polygenic components for complex traits in dairy cattle. *BMC Genet.* 13:44.
- Liu, Z. 2009. Deregressing MACE proofs for genomic evaluations. Proteje Meeting, Brussels, Belgium.
- Lund, M. S., A. P. de Roos, A. G. 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.
- Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. L. Smith, T. S. Sonstegard, and C. P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* 4:e5350.
- Nieuwhof, G. J., K. T. Beard, K. V. Konstantinov, P. J. Bowman, and B. J. Hayes. 2010. Implementation of Genomics in Australia. *Interbull Bull.* 42:35–39.
- Pszczola, M., T. Strabel, H. A. Mulder, and M. P. L. Calus. 2012. Reliability of genomic selection for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389–400.
- Thomsen, H., N. Reinsch, N. Xu, C. Looft, S. Grupe, C. Kühn, G. A. Brockmann, M. Schwerin, B. Leyhe-Horn, S. Hiendleder, G. Erhardt, I. Medjugorac, I. Russ, M. Förster, B. Brenig, F. Reinhardt, R. Reents, J. Blümel, G. Averdunk, and E. Kalm. 2001. Comparison of estimated breeding values, daughter yield deviations and de-regressed proofs within a whole genome scan for QTL. *J. Anim. Breed. Genet.* 118:357–370.
- VanRaden, P. M., K. M. Olson, D. J. Null, M. Sargolzaei, M. Winters, and J.B.C.H.M. van Kaam. 2012. Reliability increases from combining 50,000- and 777,000-marker genotypes from four countries. *Interbull Bull.* 46:75–79.
- VanRaden, P. M., and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* 74:2737–2746.
- Vollema, A. R., and A. B. F. Groen. 1998. A comparison of breeding value predictors for longevity using a linear model and survival analysis. *J. Dairy Sci.* 81:3315–3320.
- Yang, J., B. Benyamin, B. McEvoy, S. Gordon, A. Henders, D. Nyholt, P. Madden, A. Heath, N. Martin, G. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565–569.