



Genomic predictions for crossbred dairy cattle

P. M. VanRaden,^{1*} M. E. Tooker,¹ T. C. S. Chud,^{2†} H. D. Norman,³ J. H. Megonigal Jr.,³ I. W. Haagen,^{3‡} and G. R. Wiggins³

¹USDA, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD 20705-2350

²Departamento de Ciências Exatas, Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo CEP 14884-900, Brazil

³Council on Dairy Cattle Breeding, Bowie, MD 20716

ABSTRACT

Genomic evaluations are useful for crossbred as well as purebred populations when selection is applied to commercial herds. Dairy farmers had already spent more than \$1 million to genotype over 32,000 crossbred animals before US genomic evaluations became available for those animals. Thus, new tools were needed to provide accurate genomic predictions for crossbreds. Genotypes for crossbreds are imputed more accurately when the imputation reference population includes purebreds. Therefore, genotypes of 6,296 crossbred animals were imputed from lower-density chips by including either 3,119 ancestors or 834,367 genotyped animals in the reference population. Crossbreds in the imputation study included 733 Jersey × Holstein F₁ animals, 55 Brown Swiss × Holstein F₁ animals, 2,300 Holstein backcrosses, 2,026 Jersey backcrosses, 27 Brown Swiss backcrosses, and 502 other crossbreds of various breed combinations. Another 653 animals appeared to be purebreds that owners had miscoded as a different breed. Genomic breed composition was estimated from 60,671 markers using the known breed identities for purebred, progeny-tested Holstein, Jersey, Brown Swiss, Ayrshire, and Guernsey bulls as the 5 traits (breed fractions) to be predicted. Estimates of breed composition were adjusted so that no percentages were negative or exceeded 100%, and breed percentages summed to 100%. Another adjustment set percentages above 93.5% equal to 100%, and the resulting value was termed breed base representation (BBR). Larger percentages of missing alleles were imputed by using a crossbred reference population rather than only the

closest purebred reference population. Crossbred predictions were averages of genomic predictions computed using marker effects for each pure breed, which were weighted by the animal's BBR. Marker and polygenic effects were estimated separately for each breed on the all-breed scale instead of within-breed scales. For crossbreds, genomic predictions weighted by BBR were more accurate than the average of parents' breeding values and slightly more accurate than predictions using only the predominant breed. For purebreds, single-trait predictions using only within-breed data were as accurate as multi-trait predictions with allele effects in different breeds treated as correlated effects. Crossbred genomic predicted transmitting abilities were implemented by the Council on Dairy Cattle Breeding in April 2019 and will aid producers in managing their breeding programs and selecting replacement heifers.

Key words: genomic evaluation, multiple breed, crossbreeding, imputation

INTRODUCTION

Breed composition of the dairy herd has changed during the past decade, and numbers of crossbred cows have increased to more than 200,000 (greater than 5%) of the 3.9 million US milk-recorded cows in 2017 (Norman et al., 2018). The changes are due to increased emphasis on fertility and health traits, shifts in demand and pricing for various milk products (e.g., increased cheese consumption), and active promotion of additional European breeds. Several US dairy breed organizations have developed programs to enroll cattle not traditionally eligible for registry within their breed. Genomic tools can help verify or discover the ancestry and predict the genetic merit of both purebreds and crossbreds. Dairy herd owners had already spent over \$1 million genotyping more than 32,000 crossbred animals before US genomic evaluations were provided for crossbreds.

Most genomic evaluations of dairy cattle are separate by breed, and crossbreds usually are not included

Received March 15, 2019.

Accepted October 14, 2019.

*Corresponding author: Paul.VanRaden@usda.gov

†Current address: Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada.

‡Current address: 335 Henning, Department of Animal Science, Pennsylvania State University, University Park, PA 16802.

except for the multibreed evaluation in New Zealand (Winkelman et al., 2015). Some genetic evaluations include combinations of closely related breeds (Zhou et al., 2014) or subpopulations within breeds that were separate for many generations (Thomassen et al., 2013). Many research studies have combined data from different pure breeds to improve marker estimates, but most gains were small and have not yet been implemented in routine national evaluations (Karoui et al., 2012; Olson et al., 2012; Makgahlela et al., 2013; Hozé et al., 2014; Lund et al., 2014; Kemper et al., 2015).

Genomic breed composition (**GBC**) of crossbreds can be estimated accurately using all SNP markers instead of selecting a set of specific markers (VanRaden et al., 2011; Hulsege et al., 2013). In 2016, an adjusted GBC termed as breed base representation (**BBR**) was developed using all US genotyped animals (VanRaden and Cooper, 2015; Norman et al., 2016). The BBR estimates percentages of DNA contributed to a crossbred animal by each of the pure breeds.

Previous US genomic evaluations excluded the genotypes of many crossbred animals. The crossbred genotypes were removed based on counts of breed check markers that were monomorphic only in 1 breed and had fewer than 30% of animals homozygous for that allele in another breed (Wiggans et al., 2010). Reasons to exclude crossbred genotypes were that imputation of crossbred genotypes from low marker densities may not be accurate using only a single-breed reference population, marker effects estimated within a pure breed may not be appropriate for crossbreds, and parent averages (**PA**) are incomplete or incorrect if PTA of both parents are not on the same scale.

Multibreed genomic models that include crossbreds require first computing multibreed traditional pedigree-based genetic evaluations. Routine US multibreed pedigree-based evaluations were implemented for calving ease in 2005 (Cole et al., 2005); yield traits, productive life, SCS, and daughter pregnancy rate in 2007

(VanRaden et al., 2007); heifer and cow conception rates in 2013 (VanRaden et al., 2014); cow livability in 2016 (VanRaden et al., 2016); and gestation length in 2017 (Wright and VanRaden, 2017). Multibreed animal models are not applied to conformation traits because breed associations measure different traits using different definitions and scales. Thus, genomic predictions for conformation traits of crossbreds could use only data from the nearest pure breed.

The goals of this study were to develop genomic predictions for crossbreds and to analyze breed composition categories of the genotyped crossbreds by (1) computing BBR for all animals with genotypes in the national database, (2) comparing imputation strategies for crossbreds, (3) evaluating crossbreds using marker effects for each pure breed weighted by BBR, (4) testing accuracy of crossbred predictions using truncated data to predict later data, and (5) testing accuracy of single-trait versus multi-trait predictions with marker effects for different breeds considered to be correlated.

MATERIALS AND METHODS

Data

Genotypes of crossbred animals were examined for potential inclusion in the routine US genomic evaluation with new methods and edits. Previously, animals had been excluded if they had (1) a pedigree sire or dam of another breed, (2) more than 40% of breed check markers not from the evaluation breed for medium- or high-density genotypes ($\geq 50,000$ markers) or more than 20% for low-density genotypes, or (3) a genotype identified as from a completely different breed. Numbers of crossbreds that were included in or excluded from the evaluation and numbers of animals in each evaluation breed as of August 2016 are shown in Table 1. Genomic predictions were computed separately for Ayrshires (**AY**), Brown Swiss (**BS**), Guernseys (**GU**), Holsteins

Table 1. Numbers of bulls and cows in reference populations for the August 2016 US genomic evaluation and total numbers of genotyped animals by breed

Breed	Reference population		All genotyped animals
	Bulls	Cows	
Holstein	34,566	271,110	1,376,340
Jersey	4,898	55,944	187,496
Brown Swiss	6,494	2,056	26,093
Ayrshire	763	204	5,941
Guernsey	458	626	2,722
Crossbred with breed base representation <90%			
Included in genomic evaluation	323	3,345	34,678
Excluded from genomic evaluation	0	0	19,490

(HO), and Jerseys (JE). Some imputation and breed composition tests for crossbreds used an earlier 2015 subset of the 2016 data.

Data set 1 included information from 834,367 purebred and crossbred animals that had been genotyped as of March 2015 to test whether the 6,296 crossbreds that had been excluded from routine evaluation should be evaluated together with all purebreds (full data) or separately (partial data). The partial pedigree file of the crossbreds included 79,235 animals, but only 3,119 ancestors of the crossbreds had genotypes, and those were included to improve imputation accuracy. From the full data, percentages of imputed alleles based on purebred or all-breed reference populations were compared for approximately 6,000 crossbreds that had already been evaluated as purebreds. Testing the accuracy of imputation was difficult because few crossbreds had full genotypes observed. Instead, percentages of missing alleles that were not imputed were reported. For genomic prediction, allele frequencies are substituted for any alleles that remain missing after imputation.

Data set 2 included information from 1,562,760 animals that had been genotyped as of August 2016, as well as their truncated phenotypes from before December 2012, to test accuracy of predictions for crossbred cows with phenotypes in August 2016 but not in 2012. Numbers of each breed and crossbreds in data set 1 and in the truncated data were proportional to those in data set 2, except that GU marker effects were not available in data set 1. Numbers of genotyped crossbreds have increased rapidly to a total of 74,619 as of May 2019 (Figure 1) but are still only about 2% of all genotyped animals, which now total more than 3 million.

Genomic Breed Composition

For each animal in data set 1, GBC was estimated using 60,671 markers after imputation with Findhap, version 3 (VanRaden et al., 2015). Only 33 of the excluded crossbred animals had genotypes from chips with 50,000 markers or more; the other 99.5% had genotypes from a variety of lower-density chips with 3,000 to 13,000 usable markers. Marker effects to predict breed percentage as a trait were estimated using genotypes of progeny-tested purebred bulls from each of 4 breeds (AY, BS, HO, and JE). To predict the breed contributions, phenotypes were assigned as 100% or 0% according to the reported breed; for example, to predict percentage of JE genes, JE purebred bulls received 100 as phenotype, and their AY, BS, and HO contributions were set to 0 (VanRaden et al., 2011). Marker effects for GBC were then estimated using the same software as for all other traits. A simpler analysis used only 4 phenotypic observations (100 or 0) and each breed's allele frequencies instead of individual reference bulls to estimate marker effects for predicting GBC.

For BBR, the purebred reference populations are each breed's AI bulls that have daughters with milk evaluations. Any bull with another breed recorded in its 5-generation pedigree was excluded from the purebred reference population. An exception was that bulls included in the AY reference population were permitted to have ancestors from other red dairy cattle breeds, because the AY analysis included all red dairy cattle breeds (except red and white HO) as if they were 1 breed. After initial tests with the 4 breeds, GU was added as a fifth breed in 2016, when sufficient GU

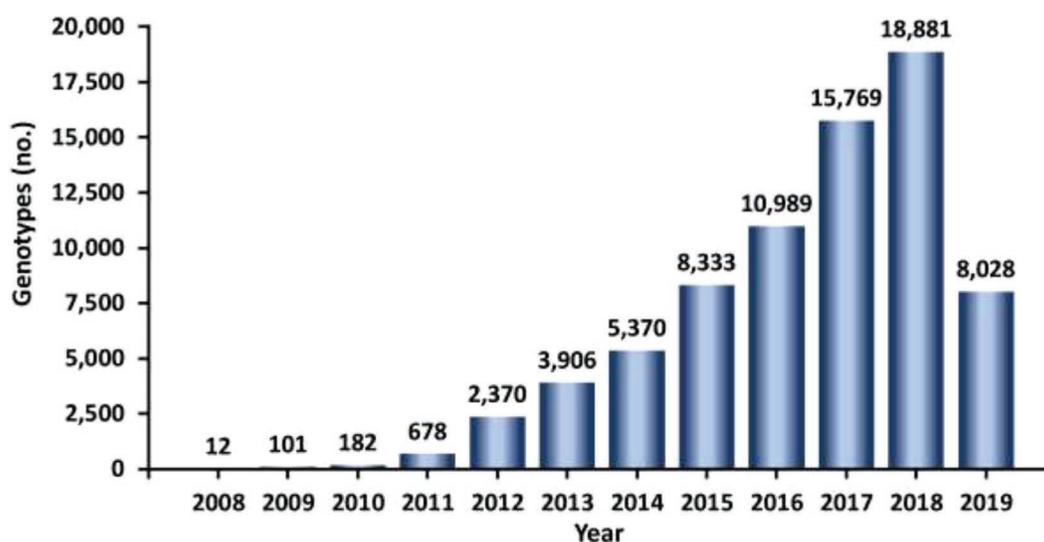


Figure 1. Frequency of crossbred genotypes received by year from 2008 through May 2019.

genotypes to form a purebred reference population were available in the national collaborator database maintained by the Council on Dairy Cattle Breeding (CDCB; Bowie, MD).

Genomic estimates of breed composition from multiple regressions can exceed 100% for a given breed or be negative for some breeds. For purebreds, expected and average GBC equal 100% for the declared breed and 0% for other breeds, but some individual animals fluctuate around those values because of genomic variation. This is similar to genomic inbreeding becoming negative for animals less related than the average for base animals. Generally, the most popular animals within a breed may exceed 100% GBC because they have more of the alleles used to differentiate the breeds.

The adjustments for BBR force genomic estimates for each animal to sum to 100%, with no estimates lower than 0% or higher than 100%. The adjustments limit estimates to be within the parameter space as in Martínez et al. (2018), simplify interpretation, and provide better weights to combine marker effects for traits with phenotypes; however, all ancestors are assumed to be from breeds that have a GBC reference population. The mathematical steps for adjusting GBC to obtain BBR are as follows: (1) sum GBC across breeds; (2) adjust GBC mean by subtracting from each GBC value the sum of GBC divided by number of breeds (N_b); (3) obtain the range of adjusted GBC from maximum adjusted breed GBC and minimum adjusted breed GBC; (4) compute an adjustment for standard deviation (SD) if any adjusted GBC are higher than 100 or lower than 0: maximum of

$$\frac{\text{largest adjusted GBC} - (100/N_b)}{100[1 - (1/N_b)]}$$

or

$$\frac{(100/N_b) - \text{smallest adjusted GBC}}{100 / N_b};$$

(5) calculate BBR as

$$\left(\frac{100}{N_b}\right) + \frac{[\text{adjusted GBC} - (100/N_b)]}{\text{SD}};$$

and (6) round BBR to 100% for animals with a BBR of 94% or higher and set contributions from all other breeds to 0%. Thus, animals with contributions of less than 6% from other breeds were assumed to be purebred.

For step 4, an example animal may have a largest adjusted GBC of 103 and smallest adjusted GBC of -3 with $N_b = 5$. Then, SD is calculated as $\max(83/80, 23/20) = 1.15$. Step 5 preserves the sum of 100% so that another mean adjustment is not needed after reducing the range of BBR values to be within the parameter space of 0 to 100%.

Genomic Prediction for Crossbreds

Marker effects that had been estimated using an approximate BayesA algorithm (VanRaden, 2008) were already available from the national genomic evaluation for each breed and phenotypic trait. These were computed using deregressed evaluations for each trait from the traditional pedigree-based evaluations after adding international evaluations and recomputing PA, and then a blending step combined genomic and traditional evaluations to reduce bias and capture information from non-genotyped ancestors (VanRaden et al., 2009).

To compute the genomic PTA (**GPTA**) for a crossbred, first a GPTA is obtained from each breed's marker effects and then those GPTA are weighted by BBR. Strandén and Mäntysaari (2013) proposed a similar random regression approach but used pedigree breed composition (**PBC**) rather than BBR to weight marker effects. These methods require all input data to be on the same all-breed scale, whereas national evaluations often define separate breed bases by setting average PTA to 0 for current animals within each breed. The GPTA were converted to the published within-breed scales as a postprocessing step, as in VanRaden et al. (2007). Estimates of marker effects included both domestic and foreign sire PTA from the multi-trait across-country evaluation of Interbull (Uppsala, Sweden). The PTA from foreign dams that had been converted from the within-breed to the all-breed base as a preprocessing step were used in PA, but only domestic cows had their PTA used in SNP effect estimation, to prevent selection bias.

Instead of weighting marker effects by BBR, GPTA were computed for each of the genomic breeds and then weighted by BBR. Within each breed, the SNP effects, intercept, and polygenic effects were summed, combined with the traditional PTA and subset PTA using a 3-part selection index as in VanRaden et al. (2009), and then weighted across breeds by BBR. An advantage for this approach was that polygenic effects from each breed were also weighted by BBR. As in the national evaluation, polygenic effects were assumed to contribute 10% of additive genetic variance. Also, imputation leaves about 1% of alleles missing, and those were filled using frequencies from each breed before weighting by BBR.

The genomic inbreeding and genomic future inbreeding (GFI) of crossbreds were obtained from the all-breed genomic relationship matrix initially, but later each crossbred's GFI was obtained only from its relationship to the breed of evaluation, because GFI measures half of its average genomic relationship to purebred mates. A crossbred whose sire and dam share some ancestors from the same breed has inbreeding from that breed.

Accuracy of GPTA was compared with that of PA using the crossbreds in data set 2. Regression was used to predict later deviations for yield (milk, fat, and protein), SCS, productive life, and fertility (daughter pregnancy, heifer conception, and cow conception rates) traits from earlier GPTA or PA. For milk yield, separate regressions and squared correlations were also estimated for 3 BBR categories of crossbreds based on BBR of the primary breed: 50 to 74%, 75 to 89%, and 90 to 100%. Very few crossbreds had a BBR of less than 50% for every breed. The truncation test included only US phenotypes (no multi-trait across-country evaluations) and excluded HO cows from the GPTA reference population to reduce computation and improve convergence of SNP effect estimates when testing multi-trait models. Because conformation traits and calving traits (calving ease and stillbirth) do not have an all-breed scale, crossbred GPTA for those traits used only marker effects from the animal's evaluation breed. A GPTA or PTA is needed for each trait to calculate selection index rankings.

Multi-Trait Modeling

Accuracy of multi-trait models that treated marker effects as correlated across all 5 breeds were compared with single-trait models that estimated marker effects independently within each breed. Correlations among marker effects for all pairs of different breeds were set to 0.3 or 0.5 as Olson et al. (2012) had performed for 3

breeds, and then multi-trait GPTA were compared with GPTA from the single-trait model with correlations set to 0. The multi-trait test used data set 2. Correlations of 2015 truncated predictions with 2017 deregressed evaluations were examined to verify whether purebred bull GPTA were more accurate when marker effects included correlated information from other breeds.

RESULTS

GBC

Breed composition is more accurately determined using genotypes than pedigrees because BBR estimates the differences caused by Mendelian sampling and because pedigrees often are partially missing or incorrect. For example, 1,024 animals had a GBC of greater than 90% purebred but also had 1 parent recorded as being from a different breed. For another 133 animals that were more than 90% purebred, breed code differed from genomic breed, which indicated possible sample switches. Crossbred categories were defined as F₁ animals with greater than 40% BBR from 2 breeds, backcrosses with greater than 67% but less than 90% BBR from 1 breed, and other crossbreds of various combinations. The crossbreds in data set 1 included 733 JE × HO F₁ animals, 55 BS × HO F₁ animals, 2,300 HO backcrosses, 2,026 JE backcrosses, 27 BS backcrosses, and 502 other crossbreds (Table 2). Many HO animals from Mexico and Chile were counted as HO backcrosses with about 10% contribution each from AY, BS, and JE, which may indicate local ancestors from other populations or breeds (such as *Bos indicus* cattle).

Mean GBC for the 834,367 animals in data set 1 was 85.7% for HO, 11.5% for JE, 2.3% for BS, and 0.5% for AY; mean PBC was nearly the same as GBC (Table 3). Mean BBR for each breed was similar to breed means for GBC and PBC, but the BBR means had a slightly

Table 2. Crossbred categories as defined by breed base representation (BBR) and numbers of animals that had been excluded by breed check edits in data set 1¹

Crossbred category	BBR definition	Breed	Animals excluded (no.)
F ₁	>40% of both breeds	Jersey × Holstein	733
Backcross	>67% but <90%	Brown Swiss × Holstein	55
		Holstein	2,300
		Jersey	2,026
		Brown Swiss	27
Other cross	Not included above	—	502
Purebred	>90% of identified breed	—	1,024
Wrong breed	<20% of identified breed	—	133
All animals	Excluded by breed check	—	6,800

¹834,367 purebred and crossbred animals genotyped as of March 2015.

Table 3. Means, minimums, and maximums of genomic breed composition (GBC), pedigree breed composition (PBC), and breed base representation (BBR; adjusted GBC) across all animals in data set 1¹

Statistic	Breed composition measure	Breed composition			
		Holstein	Jersey	Brown Swiss	Ayrshire
Mean (%)	GBC	85.7	11.5	2.3	0.5
	PBC	85.7	11.5	2.2	0.5
	BBR	84.3	11.7	2.7	0.9
Minimum (%)	GBC	-8	-3	-4	-3
	BBR	0	0	0	0
	GBC	106	108	104	110
Maximum (%)	BBR	100	100	100	100
	GBC, PBC	0.996	0.996	0.998	0.990
	GBC, BBR	0.999	0.999	0.999	0.997

¹834,367 purebred and crossbred animals genotyped as of March 2015.

reduced range compared with the GBC means. Step 5 of the BBR adjustment causes this slightly reduced range by adjusting GBC coefficients toward $100/N_b$ so that BBR ranges between 0 and 100%. Lowest GBC were -3 to -8%, whereas highest GBC were 104 to 110%. Correlations of GBC with PBC were high (≥ 0.99), and correlations of GBC with BBR were even higher. Correlations of GBC with PBC for only crossbreds were much lower and ranged from 0.72 for AY to 0.87 for JE (data not shown).

Estimates of GBC from partial and full data set 1 were very similar, which implied that genotypes were imputed consistently even when fewer purebred animals were included. Thus, GBC could be computed fairly quickly after imputing from partial data or eventually with imputation of all animals together. Estimates of GBC using only breed allele frequencies gave similar overall statistics for crossbreds, but many individual purebred animals were not accurately estimated. Another possibility would be to estimate breed composition directly without imputation using only the 6,909 markers in common among most genotyping chips, but a previous study (VanRaden et al., 2011) found that

using 3,000 markers resulted in lower accuracy than using 43,000 markers.

Genomic Prediction for Crossbreds

Crossbred GPTA obtained as averages of purebred marker effects weighted by BBR were correlated by only about 0.91 to GPTA obtained from a common set of marker effects for all breeds. For young purebred animals, correlations of GPTA with official GPTA were also low (< 0.9) when all breeds had to share 1 common set of marker effects, as previously reported by Olson et al. (2012). In contrast, the weighted GPTA within each breed were correlated by 0.98 to 0.99 to the single-breed official GPTA for all breeds except AY, for which the correlation was 0.93. Thus, weighting by BBR retains information within each pure breed while allowing crossbreds to share that information.

Accuracy of GPTA was higher than that of PA (Table 4) for crossbred cows using truncated data from 2012 to predict later phenotypes in 2016 (data set 2) for all traits except productive life. Separate regressions for the 3 BBR categories of crossbreds (Table 5) suggest

Table 4. Regressions and squared correlations of later deviation on earlier prediction for crossbred cows in data set 2¹

Trait	Cows (no.)	Regression coefficient \pm SE		Squared correlation	
		Genomic PTA	Parent average	Genomic PTA	Parent average
Milk yield	10,124	0.95 \pm 0.01	0.83 \pm 0.01	0.46	0.41
Fat yield	10,124	1.04 \pm 0.02	0.91 \pm 0.02	0.23	0.18
Protein yield	10,124	0.97 \pm 0.02	0.84 \pm 0.02	0.26	0.22
Productive life	6,149	0.90 \pm 0.11	0.70 \pm 0.11	0.01	0.01
SCS	9,984	1.16 \pm 0.04	1.02 \pm 0.05	0.07	0.04
Daughter pregnancy rate	12,854	1.06 \pm 0.07	0.65 \pm 0.07	0.02	0.01
Cow conception rate ²	—	0.61 ³	0.70 ³	0.13	0.11
Heifer conception rate ²	—	0.39 ³	0.37 ³	0.12	0.09

¹1,562,760 animals genotyped as of August 2016 as well as their truncated phenotypes from before December 2012.

²Dependent variable was deregressed PTA instead of deviation.

³SE not available.

Table 5. Regressions and squared correlations of later milk yield deviation on earlier prediction by breed base representation (BBR) category for crossbred cows in data set 2 not previously evaluated¹

BBR category (% of primary breed)	Cows (no.)	Regression coefficient \pm SE		Squared correlation	
		Genomic PTA	Parent average	Genomic PTA	Parent average
50 to 74	2,303	1.07 \pm 0.03	0.78 \pm 0.03	0.35	0.22
75 to 89	3,412	0.94 \pm 0.02	0.83 \pm 0.02	0.52	0.18
90 to 100	4,370	0.91 \pm 0.02	0.82 \pm 0.02	0.45	0.39
All crossbreds	10,124	0.95 \pm 0.01	0.83 \pm 0.01	0.46	0.41

¹1,562,760 animals genotyped as of August 2016 as well as their truncated phenotypes from before December 2012.

that the methods perform equally well at 50% BBR, 75% BBR, and 90% BBR. Results in Tables 4 and 5 were based only on crossbreds that had failed breed checks.

For animals that had passed breed checks and already had genomic evaluations in the within-breed system, the use of either a crossbred reference population or the closest purebred reference population was compared. Accuracy of weighting GPTA by BBR was tested using the same set of genotypes imputed based on the all-breed reference population (Table 6). Only results for crossbreds with a BBR of less than 94% were compared, because animals with a higher BBR are treated as purebreds. Accuracy of genomic prediction slightly favored the use of a crossbred over purebred reference population for animals with a BBR of less than 94%.

Truncation tests that included JE crossbreds with a BBR of less than 94% instead of only crossbred cows previously excluded from evaluation by breed check SNP showed that regression coefficients for weighted GPTA generally were closer to 1 than were those for single-breed GPTA (Table 6). Squared correlations for weighted GPTA ranged from 0.223 to 0.271 and were only slightly higher than squared correlations for single-breed GPTA, which ranged from 0.216 to 0.262 for the JE crossbreds with BBR of 50 to 87%. However, squared correlations showed no advantage when crossbreds with BBR of 75 to 94% were included, be-

cause most of those had small contributions from other breeds. Too few crossbreds with BBR of 50 to 74% had been evaluated as purebreds to test that subset separately. The squared correlations in both subsets were much higher for GPTA than for PA. Thus, the GPTA for currently evaluated crossbreds were quite reliable but could be improved slightly by weighting by BBR.

Multi-Trait Models

For yield traits, reliability gains over PA (Table 7) were similar for GPTA based on multi-trait and single-trait models for HO and JE, but some multi-trait gains were slightly larger for AY, BS, and GU. The reliability gains over PA for AY and BS were much lower than in routine evaluations, because most reference bulls for those breeds are foreign and this study included only US data.

DISCUSSION

More breeders now genotype their whole herds and may expect evaluations for all genotyped animals in the future. Numbers of genotyped F₁ crossbreds may still be too few to estimate crossbred performance, including interactions of marker effects contributed by different breeds as a trait separate from purebred performance, but theoretical studies have developed

Table 6. Regressions and squared correlations of later yield deviations on earlier genomic PTA predicted using weighted (crossbred) or purebred Jersey marker effects or on parent average for Jersey crossbreds in data set 2¹

Breed base representation category	Yield trait	Regression coefficient			Squared correlation		
		Genomic PTA		Parent average	Genomic PTA		Parent average
		Crossbred	Purebred		Crossbred	Purebred	
75 to 94% (n = 5,323)	Milk	0.76	0.72	0.73	0.217	0.215	0.129
	Fat	0.98	0.92	1.02	0.210	0.210	0.140
	Protein	0.82	0.77	0.89	0.169	0.169	0.115
50 to 87% (n = 1,089)	Milk	0.93	0.85	0.87	0.271	0.262	0.155
	Fat	1.09	0.97	1.06	0.247	0.242	0.146
	Protein	1.01	0.90	1.09	0.223	0.216	0.141

¹1,562,760 animals genotyped as of August 2016 as well as their truncated phenotypes from before December 2012.

Table 7. Gains (percentage points) in genomic reliability of purebred bulls¹ above parent average reliability using single-trait or multi-trait marker effects for breeds in data set 2²

Yield trait	Breed	Single-trait (correlation = 0.0)	Multi-trait	
			Correlation = 0.3	Correlation = 0.5
Milk	Holstein	27	26	26
	Jersey	16	16	15
	Brown Swiss	6	8	7
	Ayrshire	9	12	12
	Guernsey	18	20	21
Fat	Holstein	31	31	—
	Jersey	19	19	—
	Brown Swiss	8	8	—
	Ayrshire	2	6	—
	Guernsey	5	6	—
Protein	Holstein	25	25	—
	Jersey	15	15	—
	Brown Swiss	−4	−3	—
	Ayrshire	2	5	—
	Guernsey	13	10	—

¹Correlation of 2015 truncated predictions with 2017 deregressed evaluations.

²1,562,760 animals genotyped as of August 2016 as well as their truncated phenotypes from before December 2012.

methods that could be applied in populations where crossbreeding is more routine (Zeng et al., 2013; Christensen et al., 2014; Esfandyari et al., 2015). In tropical climates, composite breeds with both indicine and taurine ancestry are common, which increases the need for proper modeling of crossbreds (Cole and da Silva, 2016). Single-step evaluations could also be applied to crossbred populations (Pocrnic et al., 2019).

Other advantages of genomic testing are providing parentage verification and revealing pedigree errors. Earlier research showed that if the sire, maternal grandsire, or even maternal great-grandsire were not identified, these could be discovered in more than 99, 97, and 92% of HO cases, respectively, when tested with BovineSNP50 (Illumina Inc., San Diego, CA) markers (VanRaden et al., 2013). This also means that substantial information is available for identifying the breed composition of ancestors, even if previous on-farm record keeping was poor.

The BBR percentages can reveal the presence of either crossbreeding or outcross bloodlines, but in some cases it remains obscure which of these 2 alternatives caused the BBR to be lower than 100%. Precise percentages of breed sources are difficult to estimate because individual animal genotypes within a breed differ from the average allele frequencies for a breed. The BBR percentages for every individual are forced to sum to 100% (regardless of which breeds have reference populations) even if all genetic contributions are not from referenced breeds. Because breeds share some common DNA, the estimated percentage for the primary breed could differ from the true percentage by 5% or more.

Even animals that have true purebred ancestors for many generations often obtain a BBR percentage of less than 100% for the primary breed. Cases where the primary breed has a BBR of 90 to 97% can reveal the presence of outcross bloodlines, but a lower BBR is usually evidence of crossbreeding. Likewise, animals with 3 grandparents of 1 breed and 1 grandparent of another (expected breed alleles 75 and 25%) could easily have primary breed percentage in BBR be as low as 70% or as high as 80% because of estimation error, or primary breed percentage might deviate even more because of the random nature by which chromosomes are transmitted from grandparents through sire and dam.

Many producers want to see and use BBR despite the standard error of estimation. Some producers prefer to use bulls free of ancestors of another breed—that is, they want animals with pedigrees containing only members of the same breed. Some want bulls that will allow them to maintain what they consider their optimal mixed-breed composition; they want to capitalize on the benefits of heterosis. Still others select bulls to transform their herds to an alternate breed within a few generations. Regardless of the reason, most producers want to know an animal's breed composition when making choices.

The more an animal's genetic makeup resembles its purebred reference population, the higher its BBR percentage for the primary breed. For example, if an animal is an outcross to the current population, its calculated percentage of the primary breed could be in the mid-90s. Because the reference populations for all breeds are updated each year, an animal's BBR would

Table 8. Example estimates of breed base representation (BBR, %) for various breeding systems¹

Breeding system	HO	JE	AY	BS	GU
1964 HO control line cow	93	2	1	3	1
Danish JE cow	1	96	0	2	1
F ₁ HO × JE	52	48	0	0	0
Backcross JE × (HO × JE)	28	71	1	0	0
Second backcross JE × [JE × (JE × HO)]	15	83	1	0	1
3-breed cross HO × (JE × AY)	47	23	28	0	2
Other breed with no reference population	18	21	20	22	19

¹Breeds: AY = Ayrshire; BS = Brown Swiss; GU = Guernsey; HO = Holstein; JE = Jersey. BBR before rounding to 100% for animals with $\geq 94\%$ and setting contributions from all other breeds to 0%.

change over time if BBR were recalculated. To minimize changes, an animal's BBR should be updated only for large changes that may be caused by re-genotyping at higher density, pedigree changes affecting imputation, or genotyping of additional close relatives.

Example BBR for animals from different breeding systems are shown in Table 8. The BBR presented are before rounding of animals with BBR of 94% or more to BBR of 100% and setting contributions from all other breeds to 0%. The HO cow from a 1964 control line had 1960s genetics from a University of Minnesota experimental selection project and a relatively low relationship to the current HO population because of changes in breed allele frequencies over the past half-century. The Danish JE cow has alleles that differ somewhat from the North American JE population. Other examples in the table show various breed crosses, and the example for an animal from a breed with no reference population shows that genetic contributions from some other breed may be evenly distributed among the included breeds so that BBR percentages sum to 100. These examples illustrate that GBC can be very effective at detecting significant percentages of DNA contributed by another breed. However, selection decisions should not be made without considering that GBC and BBR percentages could be off by as much as 5% (and occasionally more), especially if crosses involve breeds without a reference population. Since April 2016, BBR has been computed by CDCB and provided for all genotyped animals.

The development of BBR formed the basis for further development of genomic predictions for crossbred animals. Currently, some crossbreds with a high percentage of a single breed do receive predictions; however, those predictions can be improved using BBR. Over 30,000 additional animals already genotyped have a moderate percentage of 2 breeds or more but do not have genomic predictions. Producer interest in receiving genetic predictions is growing, especially from those with many multibreed animals. Crossbred animals were often excluded from receiving genomic evaluations based on their pedigree, but BBR provides a more precise method of including those animals. Evaluations

for all animals including crossbreds will reduce breeder need to guess whether an animal will pass breed check edits and be evaluated before paying for genotyping, thus simplifying management decisions.

In a more recent test with more data (not shown), differences among percentages of imputed genotypes indicated that crossbred genotypes should be imputed using a crossbred reference population that includes purebred animals as well as crossbreds instead of only the closest purebred reference population. Findhap software fully imputes a genotype if alleles in both haplotypes are found; if only 1 allele is found, partial imputation is performed, and allele frequency is assigned to missing alleles. Crossbreds had about 1.5 times more missing alleles using a purebred rather than crossbred reference population for imputation, and imputation with a crossbred reference population resulted in fewer missing alleles across all BBR groups of lower than 94% (not shown). Differences decreased as BBR increased. Accurate imputation of crossbreds may require a multibreed reference population, because imputing alleles from a different breed is difficult without reference animals from that breed.

Future strategies for evaluating crossbreds may include using haplotypes instead of marker regression to trace breed of origin (Vandenplas et al., 2016) or to predict breeding values (Cuyabano et al., 2015; Hess et al., 2017; Sevillano et al., 2017). Single-step evaluations can account for differing allele frequencies among breeds (Legarra et al., 2015; Lourenco et al., 2016) and breed of origin for each allele (Xiang et al., 2016). However, including interactions among breed metafounders may be a simpler and more effective strategy (Xiang et al., 2017). Adding variants selected using an all-breed reference population can increase prediction accuracy because of closer linkage to causative effects (Porto-Neto et al., 2015; van den Berg et al., 2016a,b). Reliability gains from all-breed evaluations of crossbreds can be predicted (Vandenplas et al., 2017) but were small in the present study, perhaps because most crossbreds were JE × HO and both breeds already had large reference populations. Gains from multi-trait evaluation

of multibreed data were very small for purebred animals, which was consistent with most previous studies. Genomic predictions of purebreds changed little when computed on an all-breed rather than a within-breed scale.

Crossbred GPTA computed as a weighted average of purebred marker effects were implemented by CDCB in April 2019. However, 79,276 SNP were used, instead of the 60,671 SNP used in this study and for official predictions from 2014 until December 2018. The final methods used by CDCB excluded animals with a BBR of lower than 94% from each breed's genomic reference population and applied BBR weighting only to animals with a BBR of lower than 90%. A slight revision to BBR rounded any contributions of 1 or 2% down to 0% and adjusted the remaining breeds up proportionally. In the selection index step that combines genomic and traditional PTA, additional weight was placed on an animal's traditional PTA in proportion to its traditional reliability because a crossbred's own data do not contribute directly to estimation of allele effects. Further edits ensured that an animal's evaluation breed matched its primary breed in BBR except for F₁ crossbreds, for which either parent breed was allowed. Genotypes of both parents and the maternal grandsire were included (if available) whenever imputing crossbred genotypes.

Separate marker effects for each breed instead of just 1 set of marker effects for all breeds should provide more accurate GPTA, especially when 1 breed (HO) provides nearly 90% of all genotyped animals. Each breed's data can then contribute more directly to imputation, marker effect estimation, and prediction, instead of being dominated by linkage disequilibrium within the major breed. Many previous crossbreeding studies simulated or evaluated breeds of equal size. Optimum strategies may differ depending on numbers of phenotypes, genotypes, and categories of crossbreds included. Models that use genomic relationship matrices can force off-diagonal blocks among breeds to 0 instead of solving directly for separate marker effects by breed (Stein et al., 2019). The methods we developed used only phenotypes of purebreds to estimate marker effects and may not be ideal for other populations or species with more phenotypes from crossbreds.

CONCLUSIONS

Breed composition was estimated from genotypes after imputing lower-density chips to 60,671 markers for all animals in the national database. Imputation for crossbreds was improved with a reference population that included genotypes from parents and from each

pure breed. Methods were developed to adjust the initial GBC to BBR that was limited to a range of 0 to 100% and summed to 100% across breeds. Correlations were high between GBC and BBR. Most of the genotyped crossbreds were JE × HO backcrosses. Advantages of genomic over pedigree estimates of breed composition are that pedigrees for crossbred animals are often incomplete or inaccurate and that actual breed contributions differ from expected because of Mendelian sampling. Genomic predictions for crossbreds required estimating marker effects on an all-breed base rather than within-breed bases, and the crossbred GPTA computed by weighting marker effects for separate breeds by BBR were more accurate than predictions from single-breed GPTA. A multi-trait model, with marker effects for each breed treated as correlated, gave some advantage only for smaller breeds. Many thousands of crossbred animals are being genotyped because many commercial producers now apply genomic selection to their whole herds.

ACKNOWLEDGMENTS

The Council on Dairy Cattle Breeding (Bowie, MD) provided the data for this research under USDA Agricultural Research Service Nonfunded Cooperative Agreement 58-1245-3-228-0-N. The authors thank G. C. Fok (Animal Genomics and Improvement Laboratory, USDA, Beltsville, MD) for assistance with data processing, S. M. Hubbard (Animal Genomics and Improvement Laboratory, USDA) for assistance with technical editing, the staff of the American Jersey Cattle Association (Reynoldsburg, OH) and Zoetis Genetics (Kalamazoo, MI) for helpful discussions, and journal reviewers for suggestions for manuscript improvement. Funding for T. C. S. Chud was provided through a scholarship from the CAPES Foundation, Ministry of Education of Brazil (Brasília, DF, Brazil). Funding for P. M. VanRaden and M. E. Tooker was solely from USDA Agricultural Research Service appropriated projects 8042-31000-101-00-D, "Improving Genetic Predictions in Dairy Animals Using Phenotypic and Genomic Information," and 8042-31000-002-00-D, "Improving Dairy Animals by Increasing Accuracy of Genomic Prediction, Evaluating New Traits, and Redefining Selection Goals."

REFERENCES

- Christensen, O. F., P. Madsen, B. Nielsen, and G. Su. 2014. Genomic evaluation of both purebred and crossbred performances. *Genet. Sel. Evol.* 46:23. <https://doi.org/10.1186/1297-9686-46-23>.
- Cole, J. B., and M. V. G. B. da Silva. 2016. Genomic selection in multi-breed dairy cattle populations. *Rev. Bras. Zootec.* 45:195–202. <https://doi.org/10.1590/S1806-92902016000400008>.

- Cole, J. B., R. C. Goodling Jr., G. R. Wiggans, and P. M. VanRaden. 2005. Genetic evaluation of calving ease for Brown Swiss and Jersey bulls from purebred and crossbred calvings. *J. Dairy Sci.* 88:1529–1539. [https://doi.org/10.3168/jds.S0022-0302\(05\)72822-8](https://doi.org/10.3168/jds.S0022-0302(05)72822-8).
- Cuyabano, B. C. D., G. Su, G. J. M. Rosa, M. S. Lund, and D. Gianola. 2015. Bootstrap study of genome-enabled prediction reliabilities using haplotype blocks across Nordic Red cattle breeds. *J. Dairy Sci.* 98:7351–7363. <https://doi.org/10.3168/jds.2015-9360>.
- Esfandyari, H., A. C. Sørensen, and P. Bijma. 2015. Maximizing crossbred performance through purebred genomic selection. *Genet. Sel. Evol.* 47:16. <https://doi.org/10.1186/s12711-015-0099-3>.
- Hess, M., T. Druet, A. Hess, and D. Garrick. 2017. Fixed-length haplotypes can improve genomic prediction accuracy in an admixed dairy cattle population. *Genet. Sel. Evol.* 49:54. <https://doi.org/10.1186/s12711-017-0329-y>.
- Hozé, C., S. Fritz, F. Phocas, D. Boichard, V. Ducrocq, and P. Croiseau. 2014. Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. *J. Dairy Sci.* 97:3918–3929. <https://doi.org/10.3168/jds.2013-7761>.
- Hulsegge, B., M. P. L. Calus, J. J. Windig, A. H. Hoving-Bolink, M. H. T. Maurice-van Eijndhoven, and S. J. Hiemstra. 2013. Selection of SNP from 50K and 777K arrays to predict breed of origin in cattle. *J. Anim. Sci.* 91:5128–5134. <https://doi.org/10.2527/jas.2013-6678>.
- Karoui, S., M. J. Carabaño, C. Díaz, and A. Legarra. 2012. Joint genomic evaluation of French dairy cattle breeds using multiple-trait models. *Genet. Sel. Evol.* 44:39. <https://doi.org/10.1186/1297-9686-44-39>.
- Kemper, K. E., C. M. Reich, P. J. Bowman, C. J. vander Jagt, A. J. Chamberlain, B. A. Mason, B. J. Hayes, and M. E. Goddard. 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genet. Sel. Evol.* 47:29. <https://doi.org/10.1186/s12711-014-0074-4>.
- Legarra, A., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics* 200:455–468. <https://doi.org/10.1534/genetics.115.177014>.
- Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. *J. Anim. Sci.* 94:909–919. <https://doi.org/10.2527/jas.2015-9748>.
- Lund, M. S., G. Su, L. Janss, B. Gulbrandsen, and R. F. Brøndum. 2014. Genomic evaluation of cattle in a multi-breed context. *Livest. Sci.* 166:101–110. <https://doi.org/10.1016/j.livsci.2014.05.008>.
- Makgahlela, M. L., I. Strandén, U. S. Nielsen, M. J. Sillanpää, and E. A. Mäntysaari. 2013. The estimation of genomic relationships using breedwise allele frequencies among animals in multibreed populations. *J. Dairy Sci.* 96:5364–5375. <https://doi.org/10.3168/jds.2012-6523>.
- Martínez, C. A., K. Khare, and M. A. Elzo. 2018. BIBI: Bayesian inference of breed composition. *J. Anim. Breed. Genet.* 135:54–61. <https://doi.org/10.1111/jbg.12305>.
- Norman, H. D., P. M. VanRaden, J. H. Megonigal, J. W. Dürr, and T. A. Cooper. 2016. Breed base representation in dairy animals of five breeds. *J. Dairy Sci.* 99(E-Suppl. 1):151–152 (Abstr.).
- Norman, H. D., L. M. Walton, and J. W. Dürr. 2018. Reasons that cows in Dairy Herd Improvement programs exit the milking herd (2017). Accessed Mar. 4, 2019. <https://queries.uscdcb.com/publish/dhi/dhi18/cullall.html>.
- Olson, K. M., P. M. VanRaden, and M. E. Tooker. 2012. Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 95:5378–5383. <https://doi.org/10.3168/jds.2011-5006>.
- Pocrnic, I., D. A. L. Lourenco, C.-Y. Chen, W. O. Herring, and I. Misztal. 2019. Crossbred evaluations using single-step genomic BLUP and algorithm for proven and young with different sources of data. *J. Anim. Sci.* 97:1513–1522. <https://doi.org/10.1093/jas/skz042>.
- Porto-Neto, L. R., W. Barendse, J. M. Henshall, S. M. McWilliam, S. A. Lehnert, and A. Reverter. 2015. Genomic correlation: Harnessing the benefit of combining two unrelated populations for genomic selection. *Genet. Sel. Evol.* 47:84. <https://doi.org/10.1186/s12711-015-0162-0>.
- Sevillano, C. A., J. Vandenplas, J. W. M. Bastiaansen, R. Bergsma, and M. P. L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. *Genet. Sel. Evol.* 49:75. <https://doi.org/10.1186/s12711-017-0350-1>.
- Steyn, Y., D. A. L. Lourenco, and I. Misztal. 2019. Genomic predictions in purebreds with a multi-breed genomic relationship matrix. *J. Anim. Sci.* 97:4418–4427. <https://doi.org/10.1093/jas/skz296>.
- Strandén, I., and E. A. Mäntysaari. 2013. Use of random regression model as an alternative for multibreed relationship matrix. *J. Anim. Breed. Genet.* 130:4–9. <https://doi.org/10.1111/jbg.12014>.
- Thomassen, J. R., A. C. Sørensen, G. Su, P. Madsen, M. S. Lund, and B. Gulbrandsen. 2013. The admixed population structure in Danish Jersey dairy cattle challenges accurate genomic predictions. *J. Anim. Sci.* 91:3105–3112. <https://doi.org/10.2527/jas.2012-5490>.
- van den Berg, I., D. Boichard, and M. S. Lund. 2016a. Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. *J. Dairy Sci.* 99:8932–8945. <https://doi.org/10.3168/jds.2016-11073>.
- van den Berg, I., D. Boichard, and M. S. Lund. 2016b. Sequence variants selected from a multi-breed GWAS can improve the reliability of genomic predictions in dairy cattle. *Genet. Sel. Evol.* 48:83. <https://doi.org/10.1186/s12711-016-0259-0>.
- Vandenplas, J., M. P. L. Calus, C. A. Sevillano, J. J. Windig, and J. W. M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. *Genet. Sel. Evol.* 48:61. <https://doi.org/10.1186/s12711-016-0240-y>.
- Vandenplas, J., J. J. Windig, and M. P. L. Calus. 2017. Prediction of the reliability of genomic breeding values for crossbred performance. *Genet. Sel. Evol.* 49:43. <https://doi.org/10.1186/s12711-017-0318-1>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- VanRaden, P. M., and T. A. Cooper. 2015. Genomic evaluations and breed composition for crossbred U.S. dairy cattle. *Interbull Bull.* 49:19–23.
- VanRaden, P. M., T. A. Cooper, G. R. Wiggans, J. R. O’Connell, and L. R. Bacheller. 2013. Confirmation and discovery of maternal grandsires and great-grandsires in dairy cattle. *J. Dairy Sci.* 96:1874–1879. <https://doi.org/10.3168/jds.2012-6176>.
- VanRaden, P. M., K. M. Olson, G. R. Wiggans, J. B. Cole, and M. E. Tooker. 2011. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 94:5673–5682. <https://doi.org/10.3168/jds.2011-4500>.
- VanRaden, P. M., C. Sun, and J. R. O’Connell. 2015. Fast imputation using medium or low-coverage sequence data. *BMC Genet.* 16:82. <https://doi.org/10.1186/s12863-015-0243-7>.
- VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal Jr.. 2007. Genetic evaluations for mixed-breed populations. *J. Dairy Sci.* 90:2434–2441. <https://doi.org/10.3168/jds.2006-704>.
- VanRaden, P. M., M. E. Tooker, J. R. Wright, C. Sun, and J. L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. *J. Dairy Sci.* 97:7952–7962. <https://doi.org/10.3168/jds.2014-8489>.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16–24. <https://doi.org/10.3168/jds.2008-1514>.
- VanRaden, P. M., J. R. Wright, M. E. Tooker, and H. D. Norman. 2016. Value of selecting for cow and calf livability. *Interbull Bull.* 50:30–33.

- Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. *J. Dairy Sci.* 93:2287–2292. <https://doi.org/10.3168/jds.2009-2773>.
- Winkelman, A. M., D. L. Johnson, and B. L. Harris. 2015. Application of genomic evaluation to dairy cattle in New Zealand. *J. Dairy Sci.* 98:659–675. <https://doi.org/10.3168/jds.2014-8560>.
- Wright, J. R., and P. M. VanRaden. 2017. Genetic evaluation of gestation length as a trait of the service sire. *J. Dairy Sci.* 100(Suppl. 2):42 (Abstr.).
- Xiang, T., O. F. Christensen, and A. Legarra. 2017. Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. *J. Anim. Sci.* 95:1472–1480. <https://doi.org/10.2527/jas.2016.1155>.
- Xiang, T., B. Nielsen, G. Su, A. Legarra, and O. F. Christensen. 2016. Application of single-step genomic evaluation for crossbred performance in pig. *J. Anim. Sci.* 94:936–948. <https://doi.org/10.2527/jas.2015-9930>.
- Zeng, J., A. Toosi, R. L. Fernando, J. C. M. Dekkers, and D. J. Garrick. 2013. Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genet. Sel. Evol.* 45:11. <https://doi.org/10.1186/1297-9686-45-11>.
- Zhou, L., B. Heringstad, G. Su, B. Guldbbrandtsen, T. H. E. Meuwissen, M. Svendsen, H. Grove, U. S. Nielsen, and M. S. Lund. 2014. Genomic predictions based on a joint reference population for the Nordic Red cattle breeds. *J. Dairy Sci.* 97:4485–4496. <https://doi.org/10.3168/jds.2013-7580>.

ORCIDS

- P. M. VanRaden  <https://orcid.org/0000-0002-9123-7278>
- M. E. Tooker  <https://orcid.org/0000-0002-3890-3614>
- T. C. S. Chud  <https://orcid.org/0000-0001-7559-1165>
- I. W. Haagen  <https://orcid.org/0000-0003-1882-0798>
- G. R. Wiggans  <https://orcid.org/0000-0002-3604-6639>