Short communication: Improving accuracy of Jersey genomic evaluations in the United States and Denmark by sharing reference population bulls[1]

G. R. Wiggans,*[2] G. Su,†[2] T. A. Cooper,* U. S. Nielsen,‡ G. P. Aamand,§ B. Guldbrandtsen,† M. S. Lund,† and P. M. VanRaden*

*Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350
†Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark
‡Knowledge Center for Agriculture, Danish Agricultural Advisory Service, DK-8200 Aarhus, Denmark
§Nordic Cattle Genetic Evaluation, DK-8200 Aarhus, Denmark

ABSTRACT

The effect on prediction accuracy for Jersey genomic evaluations of Danish and US bulls from using a larger reference population was assessed. Each country contributed genotypes from 1,157 Jersey bulls to the reference population of the other. Data were separated into reference (US only, Danish only, and combined US–Danish) and validation (US only and Danish only) populations. Depending on trait (milk, fat, and protein yields and component percentages; productive life; somatic cell score; daughter pregnancy rate; 14 conformation traits; and net merit), the US reference population included 2,720 to 4,772 bulls and cows with traditional evaluations as of August 2009; the Danish reference population included 635 to 996 bulls. The US validation population included 442 to 712 bulls that gained a traditional evaluation between August 2009 and December 2013; the Danish validation population included 105 to 196 bulls with multitrait across-country evaluations on the US scale by December 2013. Genomic predicted transmitting abilities (GPTA) were calculated on the US scale using a selection index that combined direct genomic predictions with either traditional predicted transmitting ability for the reference population or traditional parent averages (PA) for the validation population and a traditional evaluation based only on genotyped animals. Reliability for GPTA was estimated from the reference population and August 2009 traditional PA and PA reliability. For prediction of December 2013 deregressed daughter deviations on the US scale, mean August 2009 GPTA reliability for Danish validation bulls was 0.10 higher when based on the combined US–Danish reference population than when the reference population included only Danish bulls; for US validation bulls, mean reliability increased by 0.02 when Danish bulls were added to the US reference population. Exchanging genotype data to increase the size of the reference population is an efficient approach to increasing the accuracy of genomic prediction when the reference population is small.

Key words: Jersey, genomic evaluation, Denmark, reference population, reliability

Short Communication

An important factor that affects the accuracy of genomic evaluations is the size of the reference population (Daetwyler et al., 2008; Goddard, 2009). In dairy cattle, reference populations are composed primarily of progeny-tested bulls because they have reliable phenotypic information from a large group of daughters. However, for a single country and for breeds other than Holstein, the number of progeny-tested bulls may be too small to achieve reliabilities for genomic evaluations near those for Holsteins. One effective approach to increase the size of the reference population has been to share animal genotypes (Schenkel et al., 2009; VanRaden et al., 2009a,c; Jorjani et al., 2011; Lund et al., 2011; VanRaden et al., 2012; Zhou et al., 2013; Lund et al., 2014). Canada and the United States have shared genotypes for all dairy cattle breeds since 2007 (Wiggans et al., 2009). Since then, those North American collaborators have also shared Holstein genotypes with the United Kingdom and Italy (VanRaden et al., 2012) and Brown Swiss genotypes with Germany, Austria, and Switzerland (Wiggans et al., 2010a, 2011) and later with other Interbull participants through the InterGenomics project (Jorjani et al., 2011). EuroGenomics was formed to facilitate sharing Holstein genotypes for

predictor populations among European partners (Lund et al., 2011).

Danish Jersey is a small dairy cattle population. Only about 1,200 to 1,400 progeny-tested bulls (depending on trait) are used as a reference population for genomic evaluation (Su et al., 2014). Because of the small size of the reference population, accuracy of genomic prediction for Danish Jerseys (Thomasen et al., 2012) is much lower than that for Danish Holstein (Su et al., 2010) and Nordic Red Cattle populations (Su et al., 2012). At the beginning of December 2013, the reference population in the United States for predicting Jersey SNP effects included 3,041 bulls and 15,662 cows from the United States and Canada, which is substantially smaller than the corresponding Holstein reference population (24,547 bulls and 60,658 cows). Later in December 2013, genotypes for Jersey bulls were exchanged between Denmark, the United States, and Canada to create larger reference populations for genomic prediction in each country (Wiggans et al., 2014b). The objective of this study was to determine the effect on prediction accuracy of US and Danish genomic evaluations for performance on the US scale from using the larger Jersey reference population.

The United States and Denmark contributed genotypes from 1,157 Jersey bulls to the reference population of the other. Jersey bulls and cows with genotypes in the US genomic database had been genotyped with the Illumina BovineSNP50 (**SNP50**; Illumina Inc., 2011a), Illumina Bovine3K (Illumina Inc., 2011b), Illumina BovineHD (Illumina Inc., 2010), Illumina BovineLD (Illumina Inc., 2013), GeneSeek Genomic Profiler (versions 1 and 2; Neogen Corporation, 2013a), or GeneSeek Genomic Profiler HD (Neogen Corporation, 2013b) BeadChips; Danish Jersey bulls had been genotyped with the SNP50 chip. All genotypes were imputed using the findhap.f90 program (VanRaden, 2015) to the common set of 61,013 SNP (45,195 from the SNP50 chip plus 15,818 from the GeneSeek Genomic Profiler HD chip) described by Wiggans et al. (2014a). That SNP set had been chosen based on SNP performance criteria such as minor allele frequency, parent-progeny conflicts, call rate, and correlation with other SNP (Wiggans et al., 2010b, 2014a).

The data were separated into reference and validation populations based on the availability of a traditional evaluation as of August 2009 (Table 1). Not all bulls with exchanged genotypes had a traditional evaluation as of August 2009. Depending on trait (milk, fat, and protein yields and component percentages; productive life; SCS; daughter pregnancy rate; 14 conformation traits; and net merit), the US reference population included 2,720 to 4,772 bulls and cows with

traditional evaluations as of August 2009; the Danish reference population included 635 to 996 bulls. The US validation population included 442 to 712 bulls (depending on trait) that gained a traditional evaluation between August 2009 and December 2013 and had ≥10 daughters; the Danish validation population included 105 to 196 bulls with multitrait across-country evaluations from the Interbull Centre (Uppsala, Sweden) on the US scale by December 2013 and daughters in ≥10 herds.

Genomic predictions were calculated using an algorithm that solved directly for marker effects using the Bayes A approximation of VanRaden (2008). The dependent variable for analysis was the deregressed daughter deviation, where the deregression factor was computed from total daughter equivalents minus daughter equivalents from parent average (**PA**; VanRaden et al., 2009b). The SNP accounted for 90% of total additive genetic variance; 10% was assigned to residual additive genetic variance. Genomic PTA (**GPTA**) were calculated using a selection index that combined direct genomic predictions with either traditional PTA for the reference population or traditional PA for the validation population and a traditional evaluation based on only genotyped animals (VanRaden et al., 2009b). Maximum weights for the direct genomic predictions were limited to 0.80 for yield traits, 0.85 for health and fertility traits, and 0.90 for type traits to improve regression of later performance on earlier prediction.

Following VanRaden et al. (2009b), genomic reliabilities were calculated from coefficients of determination for 2013 daughter deviations with 2009 predictions after adjusting for error variance in the daughter deviations and for prior selection on pedigree. Coefficients for regression of December 2013 daughter deviations on August 2009 GPTA (VanRaden, 2008) were also calculated.

The addition of Danish bulls to the US reference population changed the mean coefficient for regression of deregressed daughter deviations on GPTA only slightly (from 0.94 to 0.95) for the US validation population (Table 2); deviation from the expected value of 1.00 increased from 0.10 to 0.11. However, the addition of US bulls to the Danish reference population reduced the mean regression coefficient for the Danish validation population from 1.11 to 1.06 and the deviation from expected from 0.29 to 0.19.

For August 2009 evaluations of US validation bulls, reliability for traditional PA ranged from 0.23 (udder depth) to 0.39 (milk, fat, and protein yields and fat and protein percentages), with a mean of 0.33 across traits (Table 3). Reliabilities of GPTA for those bulls based on the US reference population ranged from 0.37 [rear

Table 1. Numbers of Jersey bulls and cows in the US, Danish, and combined US–Danish reference populations and numbers of Jersey bulls in the US and Danish validation populations by trait

Trait	Reference population			Validation population	
	US	Danish	US–Danish	US	Danish
Milk yield	4,772	996	5,768	712	196
Fat yield	4,772	996	5,768	712	196
Protein yield	4,758	996	5,754	712	196
Fat percentage	4,772	996	5,768	712	196
Protein percentage	4,758	996	5,754	712	196
Productive life	2,720	982	3,702	712	181
SCS	4,696	975	5,671	712	196
Daughter pregnancy rate	3,204	988	4,192	672	196
Final score	4,234	965	5,199	607	105
Stature	4,234	982	5,216	608	196
Strength	4,233	983	5,216	608	196
Dairy form	4,236	981	5,217	606	196
Foot angle	4,222	980	5,202	608	187
Rear legs (side view)	4,216	983	5,199	600	196
Rump angle	4,229	983	5,212	608	196
Rump width	4,233	982	5,215	605	178
Fore udder attachment	4,379	983	5,362	682	196
Rear udder height	4,254	635	4,889	593	196
Udder depth	4,432	981	5,413	442	196
Udder cleft	4,230	983	5,213	608	196
Front teat placement	4,234	983	5,217	607	196
Teat length	4,235	981	5,216	605	196
Net merit index	4,787	992	5,779	670	196

Table 2. Coefficients for regression of December 2013 deregressed daughter deviations on August 2009 genomic PTA of US and Danish validation bulls on the US scale by reference (Ref) and validation (Val) populations and trait

Trait	Regression coefficient			
	Ref: US	Ref: Danish	Ref: Combined US–Danish	
	Val: US	Val: Danish	Val: US	Val: Danish
Milk yield	0.81	0.96	0.81	1.00
Fat yield	0.82	0.41	0.82	0.61
Protein yield	0.75	0.57	0.75	0.69
Fat percentage	1.05	1.06	1.05	1.11
Protein percentage	1.00	1.17	0.99	1.17
Productive life	1.12	1.57	1.10	1.15
SCS	0.89	1.70	0.92	1.35
Daughter pregnancy rate	0.99	1.45	0.98	1.21
Final score	0.86	1.45	0.85	0.98
Stature	0.95	1.02	0.91	0.97
Strength	0.90	1.29	0.87	0.77
Dairy form	0.74	1.32	0.76	1.17
Foot angle	1.03	0.65	1.03	0.74
Rear legs (side view)	1.22	1.48	1.22	1.27
Rump angle	1.06	1.05	1.08	1.18
Rump width	0.89	1.19	0.86	1.02
Fore udder attachment	1.03	0.87	1.03	1.23
Rear udder height	0.83	0.76	0.83	0.89
Udder depth	1.07	1.15	1.09	1.35
Udder cleft	0.98	1.32	1.00	1.33
Front teat placement	0.98	1.18	1.01	1.15
Teat length	0.85	0.88	0.87	0.90
Mean	0.94	1.11	0.95	1.06

Table 3. Reliabilities (REL) of traditional parent average (PA) and genomic PTA (GPTA) of US and Danish validation bulls from predictions on the US scale by reference (Ref) and validation (Val) populations and trait

| Trait | PA REL | | GPTA REL | | | |
| | Ref: — | Ref: — | Ref: US | Ref: Danish | Ref: Combined US–Danish | |
	Val: US	Val: Danish	Val: US	Val: Danish	Val: US	Val: Danish
Milk yield	0.39	0.30	0.59	0.50	0.60	0.67
Fat yield	0.39	0.30	0.55	0.37	0.56	0.51
Protein yield	0.39	0.30	0.54	0.38	0.55	0.53
Fat percentage	0.39	0.30	0.73	0.70	0.74	0.80
Protein percentage	0.39	0.30	0.69	0.69	0.71	0.78
Productive life	0.29	0.22	0.54	0.35	0.53	0.35
SCS	0.35	0.25	0.51	0.44	0.55	0.51
Daughter pregnancy rate	0.30	0.25	0.44	0.46	0.47	0.43
Final score	0.31	0.18	0.42	0.43	0.42	0.41
Stature	0.35	0.28	0.58	0.49	0.59	0.59
Strength	0.33	0.24	0.49	0.47	0.49	0.35
Dairy form	0.33	0.24	0.50	0.34	0.52	0.56
Foot angle	0.30	0.21	0.50	0.23	0.51	0.29
Rear legs (side view)	0.28	0.25	0.37	0.42	0.41	0.47
Rump angle	0.32	0.27	0.46	0.42	0.49	0.55
Rump width	0.32	0.20	0.48	0.37	0.47	0.45
Fore udder attachment	0.25	0.24	0.47	0.37	0.48	0.63
Rear udder height	0.32	0.24	0.47	0.27	0.49	0.41
Udder depth	0.23	0.28	0.51	0.46	0.55	0.74
Udder cleft	0.31	0.24	0.41	0.39	0.41	0.49
Front teat placement	0.32	0.27	0.51	0.46	0.55	0.63
Teat length	0.32	0.29	0.48	0.43	0.53	0.51
Mean	0.33	0.26	0.51	0.43	0.53	0.53

legs (side view)] to 0.73 (fat percentage) and averaged 0.51, which was much higher than the mean reliability for PA (0.18). For individual traits, the largest reliability gain for GPTA compared with PA was 0.34 for fat percentage, and the smallest gain was 0.09 for rear legs (side view). The addition of Danish Jerseys to the US reference population increased mean GPTA reliability by 0.02; the magnitude of genomic reliability differences for individual traits generally was ≤0.02, with the exception of a reliability gain of 0.03 for daughter pregnancy rate and rump angle; 0.04 for SCS, rear legs (side view), udder depth, and front teat placement; and 0.05 for teat length.

For August 2009 evaluations of Danish validation bulls on the US scale (Table 3), reliability for traditional PA ranged from 0.18 (final score) to 0.30 (yield and component traits), with a mean of 0.26 across traits. Reliabilities of GPTA for those bulls based on the Danish reference population ranged from 0.23 (foot angle) to 0.70 (fat percentage) and averaged 0.43, which again was much higher than the mean reliability for PA (0.17). For individual traits, the largest reliability gain for GPTA compared with PA was 0.40 for fat percentage, and the smallest gain was 0.02 for foot angle. The addition of US Jerseys to the Danish reference population increased mean GPTA reliability by 0.10 for Dan-

ish validation bulls. Most traits had reliability gains [0.05 for rear legs (side view) to 0.28 for udder depth], productive life had no gain, and 3 traits had losses (0.02 for final score, 0.03 for daughter pregnancy rate, and 0.12 for strength). The reliability gain for GPTA on the Danish scale that was observed in Denmark after the exchange of US and Danish Jersey bull genotypes was lower than that shown in Table 3 for performance on the US scale and using a combined US–Danish reference population that included all US bulls and cows. Su et al. (2014) reported gains from 0.016 (fertility) to 0.125 (udder conformation) in genomic reliability over traditional PA when US Jersey bulls were added to the Danish reference population; only longevity had a reliability loss (0.056), and mean change in reliability over all traits was 0.040.

The substantial gain in reliability for Danish GPTA on the US scale after combining the US and Danish reference populations was likely the result of several factors. The Danish reference population was much smaller than the US reference population and, therefore, affected more by the addition of many new reference animals. The validation bulls selected may also have affected reliability gains. Although the number of Danish validation bulls was much smaller than the number of US validation bulls, the Danish bulls were

required to have daughters in ≥10 herds, whereas the US bulls were only required to have ≥10 daughters.

In January 2014, SNP50 genotypes for 1,168 progeny-tested Danish Jersey bulls that were provided by VikingGenetics (Randers, Denmark) were used in US monthly GPTA updates (Cooper et al., 2014). Correlations between GPTA for young animals with or without Danish genotypes included in the reference population were near 0.99 for most yield and type traits but slightly lower (0.97–0.98) for the less-heritable fitness traits. For all Jersey traits, US means and standard deviations for January 2014 GPTA changed little with the inclusion of Danish genotypes.

The exchange of 1,157 Jersey bulls doubled the size of the Danish Jersey reference population, led to a large improvement in accuracy of genomic prediction for Danish Jerseys on the US scale, and had a smaller effect for US Jerseys. Exchanging genotype data to increase the size of the reference population is an efficient approach to increasing the accuracy of genomic prediction when the reference population is small.

ACKNOWLEDGMENTS

REFERENCES

Cooper, T., G. Wiggans, P. VanRaden, and D. Null. 2014. Changes to monthly genomic evaluations (January 2014). Inclusion of Danish Jersey genotypes. Accessed Jan. 28, 2015. https://www.cdcb.us/reference/changes/eval1401.htm.

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.

Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica 136:245–257.

Illumina Inc. 2010. BovineHD Genotyping BeadChip. Accessed Jan. 28, 2015. http://res.illumina.com/documents/products/datasheets/datasheet_bovinehd.pdf.

Illumina Inc. 2011a. BovineSNP50 Genotyping BeadChip. Accessed Jan. 28, 2015. http://res.illumina.com/documents/products/datasheets/datasheet_bovine_snp5o.pdf.

Illumina Inc. 2011b. GoldenGate Bovine3K Genotyping BeadChip. Accessed Jan. 28, 2015. http://support.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_bovine3k.pdf.

Illumina Inc. 2013. BovineLD v1.1 Genotyping BeadChip. Accessed Jan. 28, 2015. http://res.illumina.com/documents/products/datasheets/datasheet_bovineld.pdf.

Jorjani, H., J. Jakobsen, M. A. Nilforooshan, E. Hjerpe, B. Zumbach, V. Palucci, and J. Dürr. 2011. Genomic evaluation of BSW populations, InterGenomics: Results and deliverables. Interbull Bull. 43:5–8.

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

Lund, M. S., G. Su, L. Janss, B. Guldbrandtsen, and R. F. Brøndum. 2014. Genomic evaluation of cattle in a multi-breed context. Livest. Sci. 166:101–110.

Neogen Corporation. 2013a. GeneSeek Genomic Profiler for dairy cattle. Accessed Jan. 28, 2015. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP-LD_Dairy.pdf.

Neogen Corporation. 2013b. GeneSeek Genomic Profiler HD for dairy cattle. Accessed Jan. 28, 2015. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP_HD_Dairy.pdf.

Schenkel, F. S., M. Sargolzaei, G. Kistemaker, G. B. Jansen, P. Sullivan, B. J. Van Doormaal, P. M. VanRaden, and G. R. Wiggans. 2009. Reliability of genomic evaluation of Holstein cattle in Canada. Interbull Bull. 39:51–58.

Su, G., B. Guldbrandtsen, V. R. Gregersen, and M. S. Lund. 2010. Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. J. Dairy Sci. 93:1175–1183.

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

Su, G., U. S. Nielsen, G. Wiggans, G. P. Aamand, B. Guldbrandtsen, and M. S. Lund. 2014. Improving genomic prediction for Danish Jersey using a joint Danish-US reference population. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, BC, Canada, Comm. 060. Accessed Jan. 28, 2015. https://asas.org/docs/default-source/wcgalp-proceedings-oral/060_paper_8823_manuscript_396_0.pdf.

Thomasen, J. R., B. Guldbrandtsen, G. Su, R. F. Brøndum, and M. S. Lund. 2012. Reliabilities of genomic estimated breeding values in Danish Jersey. Animal 6:789–796.

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

VanRaden, P. M. 2015. findhap.f90, Find haplotypes and impute genotypes using multiple chip sets and sequence data. Accessed Feb. 20, 2015. http://aipl.arsusda.gov/software/findhap/.

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., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. Schenkel. 2009a. Genomic data and cooperation result in faster progress. ICAR Tech. Ser. 13:341–346.

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

VanRaden, P. M., G. R. Wiggans, C. P. Van Tassell, T. S. Sonstegard, and F. Schenkel. 2009c. Benefits from cooperation in genomics. Interbull Bull. 39:67–72.

Wiggans, G. R., T. A. Cooper, D. J. Null, and P. M. VanRaden. 2014a. Increasing the number of single nucleotide polymorphisms used in genomic evaluations of dairy cattle. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, Canada, Comm. 301.

Accessed Jan. 28, 2015. https://asas.org/docs/default-source/wcgalp-proceedings-oral/301_paper_9522_manuscript_742_0.pdf.

Wiggans, G. R., T. A. Cooper, P. M. VanRaden, D. J. Null, J. L. Hutchison, O. M. Meland, M. E. Tooker, and H. D. Norman. 2014b. Calculation and delivery of US genomic evaluations for dairy cattle. J. Dairy Sci. 92(E-Suppl. 1):77–78. (Abstr.)

Wiggans, G. R., T. A. Cooper, P. M. VanRaden, and M. V. Silva. 2010a. Increased reliability of genetic evaluations for dairy cattle in the United States from use of genomic information. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany, Comm. 0476. Gesellschaft für Tierzuchtwissenschaften e. V., Gießen, Germany. Accessed Jan. 28, 2015. http://www.kongressband.de/wcgalp2010/assets/pdf/0476.pdf.

Wiggans, G. R., T. S. Sonstegard, P. M. VanRaden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, J. P. Chesnais, F. S. Schenkel, and C. P. Van Tassell. 2009. Genomic evaluations in the United States and Canada: A collaboration. ICAR Tech. Ser. 13:347–353.

Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010b. Selection and management of DNA markers for use in genomic evaluation. J. Dairy Sci. 93:2287–2292.

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

Zhou, L., X. Ding, Q. Zhang, Y. Wang, M. S. Lund, and G. Su. 2013. Consistency of linkage disequilibrium between Chinese and Nordic Holsteins and genomic prediction for Chinese Holsteins using a joint reference population. Genet. Sel. Evol. 45:7.