J. Dairy Sci. 98:3508–3513
http://dx.doi.org/10.3168/jds.2014-8874

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

---

Received September 18, 2014.
Accepted January 30, 2015.

[1] The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the US Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

[2] Corresponding authors: George.Wiggans@ars.usda.gov and Guosheng.Su@agrsci.dk

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