J. Dairy Sci. 98:1–4
http://dx.doi.org/10.3168/jds.2014-8868
© American Dairy Science Association®, 2015.

Technical note: Rapid calculation of genomic evaluations for new animals[1]

G. R. Wiggans,[2] P. M. VanRaden, and T. A. Cooper
Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

A method was developed to calculate preliminary genomic evaluations daily or weekly before the release of official monthly evaluations by processing only newly genotyped animals using estimates of single nucleotide polymorphism effects from the previous official evaluation. To minimize computing time, reliabilities and genomic inbreeding are not calculated, and fixed weights are used to combine genomic and traditional information. Correlations of preliminary and September official monthly evaluations for animals with genotypes that became usable after the extraction of genotypes for August 2014 evaluations were >0.99 for most Holstein traits. Correlations were lower for breeds with smaller population size. Earlier access to genomic evaluations benefits producers by enabling earlier culling decisions and genotyping laboratories by making workloads more uniform across the month.
Key words: genomic evaluation, genotype, continuous evaluation

Technical Note

The dairy industry would like to have a shorter time between collecting a DNA sample and receiving the genomic evaluation, and genotyping laboratories would like a more uniform submission of samples throughout the month. Many animals are genotyped as calves; therefore, earlier culling decisions can reduce rearing costs. Although the monthly system for dairy cattle could be processed more frequently, several days of computing time are needed, and that time increases each month because of the addition of newly genotyped animals. A system that processes only the new animals while using the results from the last official monthly evaluation could take considerably less time and have fairly stable time requirements. It also could be a testing ground for techniques that might be needed if the time for processing a monthly evaluation became excessive.

More frequent release of genomic evaluations also should result in more uniform submission of DNA samples over the month. Currently, sample submission is concentrated in the period preceding the genotype due date, which creates a spike in the workload for genotyping laboratories.

The Netherlands and Germany currently provide weekly genomic evaluations (Alkhoder et al., 2014; Stoop et al., 2014). The objective of this research was to develop a system that would generate preliminary US genomic evaluations that were similar enough to official monthly evaluations to be acceptable to the dairy industry but that could be calculated quickly enough that even daily evaluations would be practical and, whenever possible, use programs from the monthly evaluation system.

Genotype Extraction and Imputation for Preliminary Evaluations

The primary basis for computational savings is to extract only genotypes from the North American database that have become usable since the last genotype extraction. A new extraction program was developed that relies on the date and time (timestamp) of the last extract and checks a transaction table to determine if a usable genotype was not in the table or if the most recent entry for the genotype before the timestamp was not usable. Separate genotype files are created by breed, and subsequent processing is separate by breed as in the monthly evaluations. Although breeds with small populations would benefit from the contribution of genotype information from breeds with large populations, an across-breed genomic evaluation system currently is not available. Genotypes from the Interbull Centre (Uppsala, Sweden) and the Associazione Nazionale Allevatori di Razza Frisona (Cremona, Italy) are used only in official monthly genomic evaluations and, therefore, are excluded from the extract. Almost all the extracted genotypes are from low-density genotyping chips (up to 13,218 usable SNP).

Received September 17, 2014.
Accepted November 28, 2014.
[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 author: George.Wiggans@ars.usda.gov

In the monthly system, the next step is to harmonize genotypes for animals that should have identical genotypes, such as identical twins, split embryos, and clones. This same program sets SNP with a parental conflict to missing and fills other missing SNP with values that are obvious from parents. This step is omitted from the preliminary system to save time and because the necessary information is not readily available because not all genotypes are extracted. The extraction of pedigree information is done as in the monthly system but includes only newly genotyped animals and their ancestors. The findhap imputation program (VanRaden, 2011) was upgraded to version 3 to support monthly and weekly processing, and genotypes are imputed to 60,671 SNP. Monthly, it uses the haplotype library from the previous month and updates it to reflect information from added genotypes. In the preliminary system, only new genotypes are processed, and the haplotype library is not updated. To aid in the determination of imputation accuracy, the findhap program was also modified to access a count of moderate- to high-density genotypes that are represented in the monthly haplotype library.

Calculation of Preliminary Genomic Evaluations

The steps in generating preliminary evaluations are shown in Figure 1. The file of traditional evaluations, parent averages, and their reliabilities is created as in the monthly system but only for newly genotyped animals. The program that estimates SNP effects is run using solutions from the previous monthly evaluation as priors, but the number of rounds of iteration to perform is set to zero, which causes it to apply the previous month's SNP effects to the new genotypes. That program also calculates the polygenic effects. The allele frequencies from the previous monthly evaluation are used because those from the findhap imputation in the preliminary system reflect only newly genotyped animals and their ancestors. A program specific to the preliminary system replaced the selection index step and instead applies constant weights to genomic and traditional information. That approximation is used because reliabilities and the traditional subset evaluation are not calculated to save time.

The program from the monthly system is used to combine evaluations across traits and calculate the net,

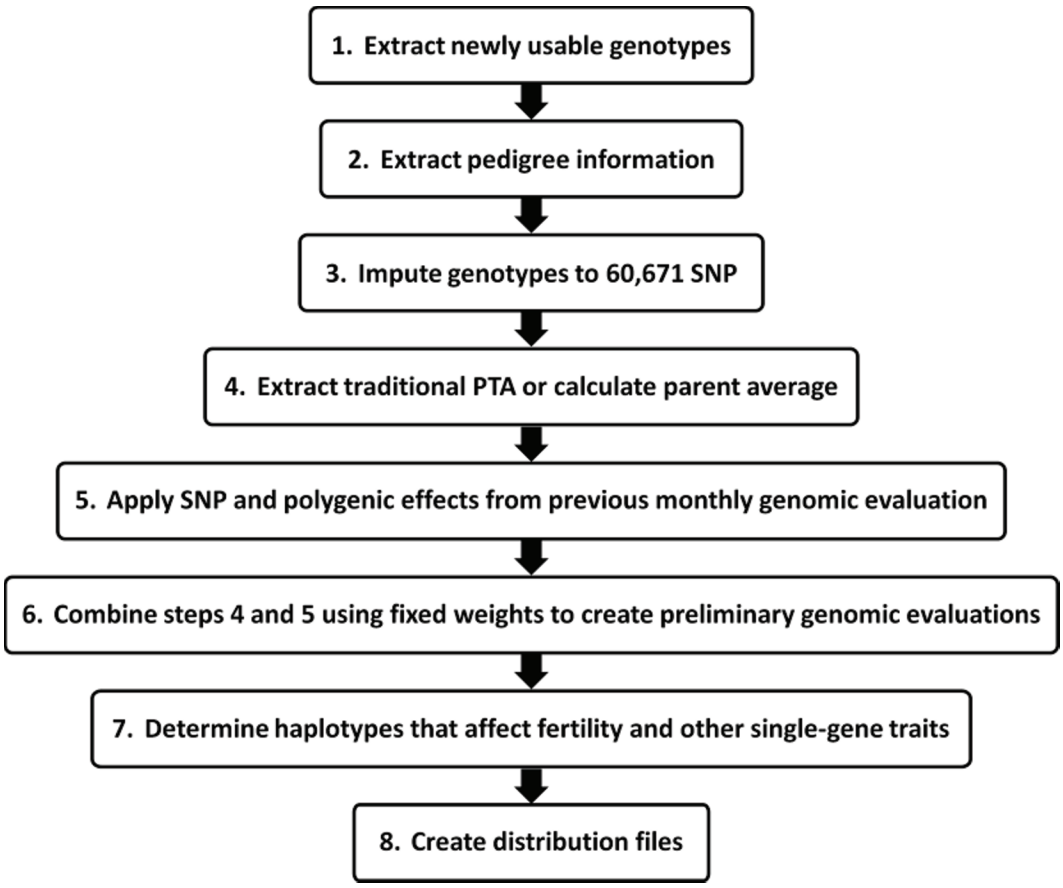

Figure 1. Data processing steps for US preliminary genomic evaluation system.

cheese, and fluid merit indexes. A preliminary-specific version of the program for determining predicted recessive indicators was developed to bypass discovery of new haplotypes that affect fertility. The program to create distribution files was adapted to detect if a preliminary or monthly evaluation is being processed. That program sets reliabilities and inbreeding values to missing for preliminary evaluations and changes file names to differentiate preliminary from monthly evaluations.

Testing

Test files from the preliminary system were generated from genotypes that became usable after the extraction of genotypes for August 2014 evaluations. Genotypes were added for 23,691 Holsteins (2,102 bulls), 2,568 Jerseys (350 bulls), 87 Brown Swiss (30 bulls), and 30 Ayrshires (8 bulls) in September 2014. The total number of genotypes that were usable for evaluation in September 2014 was 596,469 for Holsteins, 77,725 for Jerseys, 17,293 for Brown Swiss, and 3,239 for Ayrshires. Computing time for weekly preliminary evaluations was 1.5 h compared with >3 d for official monthly genomic evaluations.

Correlations between the preliminary and the September 2014 official monthly genomic evaluations were calculated (Table 1). For Holsteins, correlations were ≥0.99 (0.992–0.999) for all traits. Correlations were slightly lower for Jerseys (0.984–0.997), but all were ≥0.99 except for heifer conception rate, which has a low heritability (0.016). Ayrshires and Brown Swiss each had <100 animals with newly added genotypes. Brown Swiss correlations (0.958–0.995) generally were slightly lower than Jersey correlations. Ayrshire correlations were the lowest (0.873–0.995), a reflection of decreased evaluation accuracy because of the much smaller size of the Ayrshire reference population. Mean differences between evaluations from preliminary and monthly evaluations were small. For Holsteins, the mean differences (preliminary minus monthly) were −2.6 kg for milk yield, −0.10 kg for fat yield, −0.08 kg for protein yield, −0.03 mo for productive life, 0.001 for SCS, 0.011 percentage points for daughter pregnancy rate, −0.010 for final score, and −$2.67 for net merit; mean absolute differences were 8.6 kg for milk yield, 0.32 kg for fat yield, 0.24 kg for protein yield, 0.06 mo for productive life, 0.005 for SCS, 0.040 percentage points for daughter pregnancy rate, 0.026 for final score, and $5.61 for net merit.

Differences between preliminary and official evaluations result primarily from the approximation of the selection index step. Differences are expected for cows but should not be evident for calves, which have no subset evaluation to include. Minor differences can result from the updated estimates for SNP effects. Those differences are likely to be greater for breeds with smaller numbers because their SNP estimates are more volatile. Because the objective of the preliminary evaluation is to assist in culling calves, its poorer performance for cows is probably acceptable. Accuracy of the preliminary evaluation could be improved if necessary, but processing time likely would increase.

Table 1. Correlations between preliminary and official September 2014 monthly genomic evaluations by breed

Trait	Holstein (23,691animals)	Jersey (2,568 animals)	Brown Swiss (87 animals)	Ayrshire (30 animals)
Milk yield	0.999	0.997	0.994	0.874
Fat yield	0.999	0.997	0.993	0.923
Protein yield	0.999	0.996	0.994	0.887
Fat percentage	0.998	0.997	0.992	0.873
Protein percentage	0.996	0.994	0.986	0.897
Productive life	0.999	0.994	0.974	0.929
SCS	0.999	0.995	0.978	0.921
Daughter pregnancy rate	0.998	0.993	0.990	0.952
Service-sire calving ease	0.996	—	0.995	—
Daughter calving ease	0.994	—	0.992	—
Service-sire stillbirth rate	0.992	—	—	—
Daughter stillbirth rate	0.994	—	—	—
Heifer conception rate	0.997	0.984	0.986	0.995
Cow conception rate	0.996	0.992	0.958	0.935
Stature	0.998	0.995	0.987	0.963
Strength	0.999	0.993	0.982	0.888
Udder depth	0.998	0.996	0.981	0.969
Final score	0.999	0.992	0.979	0.960
Net merit	0.999	0.996	0.988	0.912
Cheese merit	0.999	0.996	0.990	0.910
Fluid merit	0.999	0.996	0.987	0.908

Implementation of Preliminary Evaluations

The Council on Dairy Cattle Breeding (2014) began releasing weekly preliminary evaluations in November 2014. The Council provides the evaluations to nominators, dairy records processing centers, and breed associations to facilitate transfer to animal owners. Methods for rapid approximation of reliability and genomic inbreeding are being investigated. Currently, the lack of a reliability value assists in distinguishing the preliminary evaluations from official monthly genomic evaluations.

ACKNOWLEDGMENTS

The cooperation of the Council on Dairy Cattle Breeding (Reynoldsburg, OH) in supplying North American pedigree, performance, and genotypic data is acknowledged. The authors thank S. M. Hubbard (Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD) for technical review and assistance in manuscript preparation.

REFERENCES

Alkhoder, H., Z. Liu, T. Bach, E. Pasman, and F. Reinhardt. 2014. A continuous genomic evaluation system for German Holsteins. Interbull Bull. 48:100–104.

Council on Dairy Cattle Breeding. 2014. CDCB action items from CDCB board meeting August 5–6, 2014. Accessed Aug. 28, 2014. https://www.cdcb.us/News/CDCB%20Action%20Items%20Aug%204.pdf.

Stoop, W. M., E. Mullaart, and G. de Jong. 2014. Industry application of genomic predictions in The Netherlands. Communication 810 in Proc. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, BC, Canada. Am. Soc. Anim. Sci, Champaign, IL.

VanRaden, P. 2011. findhap.f90. Accessed Aug. 9, 2014. http://aipl.arsusda.gov/software/findhap/.