J. Dairy Sci. 98:2039–2042
http://dx.doi.org/10.3168/jds.2014-8868

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,

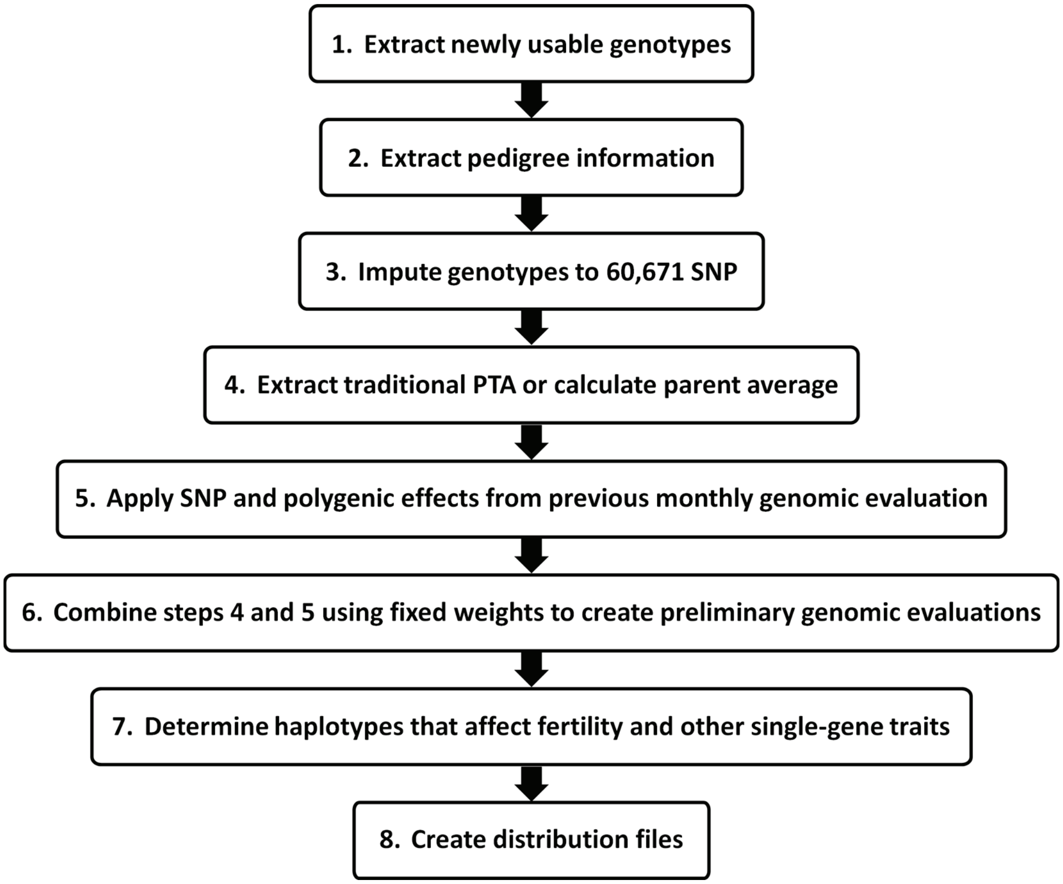

Figure 1. Data processing steps for US preliminary genomic evaluation system.