J. Dairy Sci. 96:2713–2726
http://dx.doi.org/10.3168/jds.2012-6008

Visualization of the transmission of direct genomic values for paternal and maternal chromosomes for 15 traits in US Brown Swiss, Holstein, and Jersey cattle

J. B. Cole[1] and D. J. Null
Animal Improvement Programs Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD 20705-2350

ABSTRACT

Haplotypes are available for 220,671 Brown Swiss, Holstein, and Jersey bulls and cows that received genomic evaluations in August 2012. Differences in least squares means of direct genomic values (DGV) for paternal and maternal haplotypes of [redacted] autosomes 1, 6, 14, and 18 for lifetime net merit were significant in all but one case. Those chromosomes were chosen to represent cases with and without known quantitative trait loci, and other chromosomes may differ as well. Paternal haplotypes had higher DGV than maternal haplotypes in most cases, and differences were larger when quantitative trait loci were present. Longer chromosomes generally accounted for more variance than shorter chromosomes, and differences among breeds were consistent with known mutations of large effect. For example, [redacted] autosome 18 accounted for 2.5, 7, and 2.6% of the variance in lifetime net merit for Brown Swiss (BS), Holsteins, and Jerseys, respectively. Distributions of the number of positive DGV inherited from sires and dams were negatively skewed in all breeds, and modes were slightly higher for paternally than maternally derived haplotypes in Holsteins and BS (22 vs. 20 and 22 vs. 21, respectively) and slightly lower in BS (17 vs. 19). Graphical representations of DGV are available to all users through a query on the Animal Improvement Programs Laboratory (ARS, USDA, Beltsville, MD) web site. Query results were also used to illustrate several quantitative genetic principles using genotype information from real animals. For example, offspring DGV can be compared with parental DGV to demonstrate that a parent transmits the average value of its 2 chromosomes to its progeny. The frequency of DGV with positive and negative values in animals of different ages can be used to show how selection affects allele frequencies. The effect of selection for alleles with large effects versus those with small effects is demonstrated using an animal with undesirable alleles for a marker with a large effect but many desirable alleles for markers with small effects. Strategies for the use of those data in selection programs are being studied, and work is underway to add data on conformation traits to the system.

Key words: direct genomic value, genomic selection, haplotype, visualization

INTRODUCTION

The success of genomic selection programs in dairy cattle populations has resulted in a growing demand for new tools to help understand the resulting data. Genotypes for more than 200,000 animals are available in the national dairy database (Beltsville, MD) and reliable haplotypes are available for most of those animals as well (VanRaden et al., 2011b). Several papers have addressed the effect of genomic selection in cattle improvement and conservation programs (Engelsma et al., 2011; Bastiaansen et al., 2012; Börner and Reinsch, 2012), and some studies have discussed the use of SNP genotypes for mate selection (Nishio et al., 2010; Toro and Varona, 2010; Cole and VanRaden, 2011; Pryce et al., 2012). Cole and VanRaden (2011) suggested that the greatest selection response possible on an overall economic index in a population could be achieved by constructing genotypes from desirable haplotypes, creating an animal whose genotype consisted of 2 copies each of the best haplotype for each chromosome. When the direct genomic values (**DGV**) of each haplotype were adjusted to account for inbreeding, the best genotypes consisted of 2 copies each of the same haplotype for 22 to 26 chromosomes, and differences between the best and next-best haplotypes generally were small (<$10). Kemper et al. (2012) recently have confirmed that such a strategy produces the largest long-term selection gains under some conditions, but they showed that selection on genomic breeding values with constraints on coancestry provided similar long-term gains, faster short-term gains, and greater flexibility. However, strategies for using haplotype data for mate selection still have not been described in detail in the literature.

Received July 31, 2012.
Accepted December 19, 2012.
[1]Corresponding author: john.cole@ars.usda.gov

Cole and VanRaden (2010) discussed several approaches for presenting data visually, including the display of chromosomal PTA. In August 2009, a publicly accessible query was added to the Animal Improvement Programs Laboratory's (Agricultural Research Service, USDA, Beltsville, MD) web site, which produced plots of chromosomal PTA for genotyped animals. However, those PTA were calculated as half the sum of the average effects of the genes carried by an individual, and suggested that animals transmit average chromosomes to their progeny rather than distinct haplotypes, which can vary dramatically. Due to growing demand from dairy farmers for genomic tools to assist in mate selection, the original query was modified to display DGV for maternal and paternal haplotypes, provide greater control over plot options, and include new traits. The purpose of this paper is to describe those new tools, and to discuss challenges related to the application of haplotypes to mating decisions.

MATERIALS AND METHODS

Overview

The flow of data through the US national genomic evaluation system is presented schematically in Figure 1. Inputs to the system include phenotypes, pedigrees, and genotypes that are provided by dairy farmers, data collection centers, artificial insemination firms, and breed associations. Pedigree and phenotypic data are combined using an animal model to produce traditional PTA that do not include genomic information. Independently of this process, the program findhap.f90 (VanRaden et al., 2011b) uses pedigree and genotypic data to construct haplotypes. The densemap.f90 program uses the traditional PTA and genotypes are then used to compute SNP effects, genomic PTA (**gPTA**), and DGV for paternal and maternal haplotypes. Finally, the DGV for paternal and maternal chromosomes are combined with animal identification and stored in a relational database following each official genetic evaluation. The query is implemented using a web application development framework and will run in all major web browsers.

Modifications to the Database

The DGV and associated information are stored in a relational database table (DB2 V9.5, IBM Corporation, Armonk, NY) for access by web query. As part of the routine data processing for each official genetic evaluation, an SAS program (SAS 9.3, SAS Institute Inc., Cary, NC) is used to drop the existing table, create an empty table in the database, and load the data. This ensures that only current data are included in the query, and that those data reflect the most recent information available for each animal. The addition of the haplotype data and inclusion of results for 2 more traits (heifer and cow conception rates) increased the size of the table by approximately a factor of 3.

The database table includes 3 rows for every genotyped animal-trait combination: 1 row each for paternal and maternal haplotypes, and a third row for the sum of the haplotypes. The overall DGV may be computed as the sum of the parental DGV, but storing it as a separate row simplifies the display logic for some queries. Rows are indexed by animal ID, trait, and DGV type (paternal DGV, maternal DGV, or total DGV).

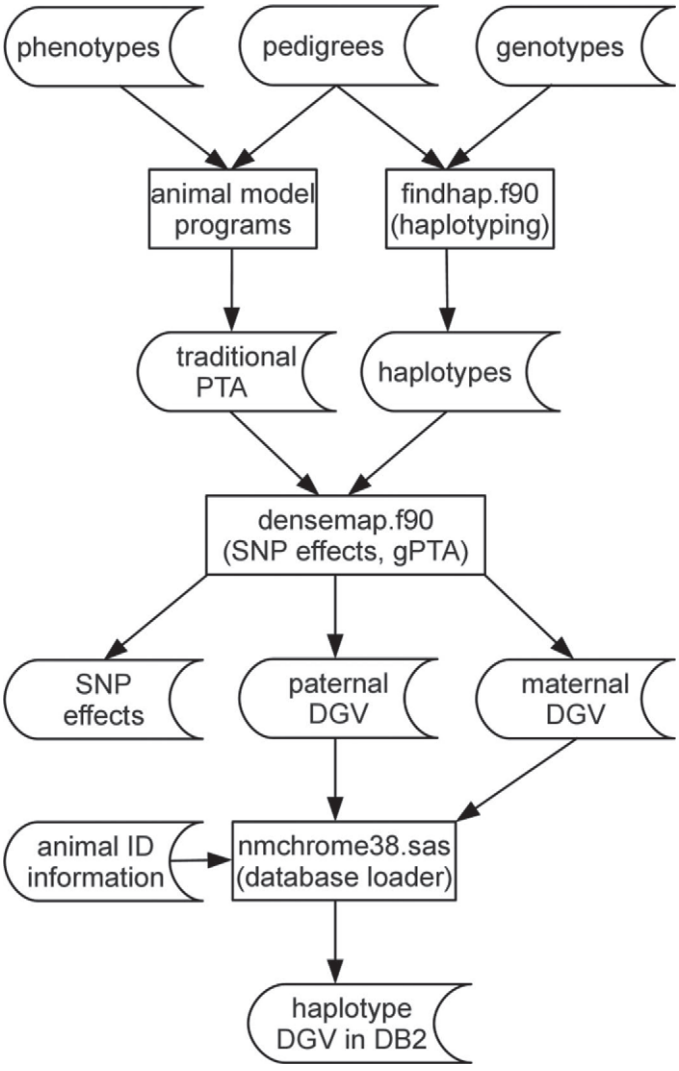

Figure 1. Flow of data and programs used to calculate direct genomic values (DGV) using traditional PTA and SNP. The resulting DGV for paternal and maternal haplotypes for each trait are combined with animal identification and stored in a table in a relational database.