J. Dairy Sci. 96:5376–5387
http://dx.doi.org/10.3168/jds.2013-6650

Application of a posteriori granddaughter and modified granddaughter designs to determine Holstein haplotype effects[1]

J. I. Weller,*[2] P. M. VanRaden,† and G. R. Wiggans†
*Institute of Animal Sciences, A.R.O., The Volcani Center, Bet Dagan 50250, Israel
†Animal Improvement Programs Laboratory, Agricultural Research Service, US Department of Agriculture (USDA), Beltsville, MD 20705-2350

ABSTRACT

A posteriori and modified granddaughter designs were applied to determine haplotype effects for Holstein bulls and cows with BovineSNP50 [~50,000 single nucleotide polymorphisms (SNP); Illumina Inc., San Diego, CA] genotypes. The a posteriori granddaughter design was applied to 52 sire families, each with ≥100 genotyped sons with genetic evaluations based on progeny tests. For 33 traits (milk, fat, and protein yields; fat and protein percentages; somatic cell score; productive life; daughter pregnancy rate; heifer and cow conception rates; service-sire and daughter calving ease; service-sire and daughter stillbirth; 18 conformation traits; and net merit), the analysis was applied to the autosomal segment with the SNP with the greatest effect in the genomic evaluation of each trait. All traits except 2 had a within-family haplotype effect. The same design was applied with the genetic evaluations of sons corrected for SNP effects associated with chromosomes besides the one under analysis. The number of within-family contrasts was 166 without adjustment and 211 with adjustment. Of the 52 bulls analyzed, 36 had BovineHD (high density; Illumina Inc.) genotypes that were used to test for concordance between sire quantitative trait loci and SNP genotypes; complete concordance was not obtained for any effects. Of the 31 traits with effects from the a posteriori granddaughter design, 21 were analyzed with the modified granddaughter design. Only sires with a contrast for the a posteriori granddaughter design and ≥200 granddaughters with a record usable for genetic evaluation were included. Calving traits could not be analyzed because individual cow evaluations were not computed. Eight traits had within-family haplotype effects. With respect to milk and fat yields and fat percentage, the results on *Bos taurus* autosome (BTA) 14 corresponded to the hypothesis that a missense mutation in the diacylglycerol O-acyltransferase 1 (*DGAT1*) gene is the main causative mutation, although other polymorphisms in that gene also modify fat yield and percentage. The positive allele for protein concentration was less frequent, which indicated that selection on that locus could be effective. Although the results can be used to determine causative polymorphisms for most of the analyzed traits, complete DNA sequencing of most of the analyzed sires probably will be required.

Key words: granddaughter design, genetic evaluation, genomic selection, haplotype

INTRODUCTION

Since 2008, genomic evaluation has become a reality chiefly because of the development of high-density SNP chips that allow for relatively inexpensive genotyping of individuals for tens of thousands of genetic markers. With thousands of genotyped bulls with progeny records, reliabilities of >0.7 can be obtained for genotyped animals without records or progeny records, compared with reliabilities of <0.4 based only on parent evaluations. The methods developed for genomic evaluations are based on population-wide linkage disequilibrium between closely linked markers and the actual QTL that determine phenotypes for the traits of interest (e. g., VanRaden et al., 2009). Because linkage disequilibrium is generally incomplete, even if a SNP has a major estimated effect on an economic trait, the SNP genotypes of individual animals do not necessarily correspond to their QTL genotypes. With the exception of the diacylglycerol O-acyltransferase 1 (*DGAT1*) and ATP-binding cassette sub-family G member 2 (*ABCG2*) genes (Grisart et al., 2002; Winter et al., 2002; Cohen-Zinder et al., 2005), quantitative trait nucleotides (**QTN**), which are the actual polymorphisms responsible for detected QTL, remain unknown. Determination of the actual polymorphisms responsible for the observed genetic variation should result in increased rates of genetic gain (Weller and Ron, 2011).

Methods applied to plants and experimental animals to determine QTN cannot be applied to large farm

Received January 31, 2013.
Accepted April 22, 2013.
[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: weller@agri.huji.ac.il

animals. For dairy cattle, the most convincing proof that a QTN has been discovered is if concordance is obtained for a sample of animals with known QTL genotypes. Complete concordance is obtained only if all individuals heterozygous for the QTL are also heterozygous for the putative QTN and vice versa. In addition, among the individuals heterozygous for the QTL, the allele with the positive effect on the trait should be associated with the same allele of the putative QTN (Ron and Weller, 2007). Proof that the QTN has been determined requires that concordance be ascertained for a sufficiently large group of animals so that the hypothesis that concordance was obtained by chance can be statistically rejected with high power (Ron and Weller, 2007).

Unlike major genes for which genotypes can be determined directly from the phenotypes, QTL genotypes can only be determined for bulls with many progeny based on application of either a daughter or granddaughter design (Weller et al., 1990). This design was first applied to the US dairy cattle population by Georges et al. (1995), and has since been applied to almost all major commercial dairy cattle populations (Weller, 2007).

Weller and Ron (2011) proposed application of a posteriori granddaughter design (**APGD**), diagrammed in Figure 1A, to determine QTL genotypes for bulls from large populations of individuals genotyped using high-density SNP chips. Similar to the original granddaughter design, sires with many progeny-tested sons are analyzed. However, rather than genotype the sons specifically for application of a granddaughter design, the data generated by genotyping many bulls for high-density SNP chips are used. Thus, the design is considered a posteriori. The sons of each bull are divided into 2 groups based on which paternal haplotype was passed to each son for the chromosomal region with the putative QTL. If a contrast ($P < 0.05$) is obtained between the 2 progeny groups, then it can be deduced that the sire is heterozygous for the QTL. Otherwise, it can be assumed that the sire is homozygous for the QTL, provided that the experimental design has sufficient power to detect heterozygous sires with a high probability.

Compared with application of granddaughter designs based on microsatellites, the a posteriori design is more powerful for genomic analysis for several reasons. First, unlike individual microsatellites, which are homozygous for a significant fraction of the animals' genotypes, in the APGD, each haplotype is based on the genotypes of tens of tightly linked SNP (e. g., Druet et al., 2008). Many different haplotypes segregate in the population for each specific chromosomal segment, and almost all bulls are heterozygous for their haplotypes. Thus, the paternal haplotype of almost all sons can be determined. Second, the bulls available for analysis in the 1990s were a selected sample, as semen of inferior bulls was generally not retained by the AI institutes. Finally, if the whole genome is analyzed, then the number of individual comparisons is huge, and nominal significance levels of 5 or 1% are meaningless. Thus, much lower nominal significance levels are required to determine that a segregating QTL has been detected. In the current study, this multiple comparison problem was avoided by considering only specific chromosomal regions that were shown to harbor segregating QTL by genome scan results.

Weller et al. (2002) proposed that QTL allelic frequencies and the number of segregating QTL alleles in the population could be determined by application of a modified granddaughter design. In this design, diagrammed in Figure 1B, maternal granddaughters of a bull are divided into 3 groups based on which grandpaternal haplotype was passed to each granddaughter. Based on Mendelian sampling, 25% of the granddaughters should receive 1 of the 2 grandpaternal alleles, 25% should receive the other allele, and the remaining 50% should receive neither allele. Similar to the APGD for genomic evaluation, almost all grandsires will be heterozygous for their haplotypes, and haplotype determination of the granddaughters can be determined almost without error. The main advantage of the modified granddaughter design is that those granddaughters that received neither grandpaternal haplotype can be considered to be a random sample of the QTL alleles that are segregating in the population. Thus, by comparing the effects associated with the 2 grandpaternal haplotypes with the effect associated with the granddaughters that received neither grandpaternal allele, determining the relative frequencies of the 2 grandpaternal QTL alleles in the population should be possible (Weller et al., 2002). The closer the effect associated with 1 grandpaternal QTL allele is to the effect associated with neither grandpaternal allele, the greater the frequency of this allele in the general population. Determining if more than 2 effective QTL alleles are segregating in the population also should be possible (i.e., alleles with measurably different effects on the trait). This will be observed if the effects of the 2 paternal alleles relative to the progeny group that received neither paternal allele differ ($P < 0.05$) across families (Weller et al., 2002).

The primary objective of this study was to determine QTL genotypes for chromosomal regions determined previously to harbor segregating QTL for all traits that are genetically evaluated in the United States for a sample of approximately 50 bulls via application of the APGD. These data were then used to determine