J. Dairy Sci. 97:1139–1149
http://dx.doi.org/10.3168/jds.2013-7197

Fine mapping of a quantitative trait locus for bovine milk fat composition on *Bos taurus* autosome 19

Aniek C. Bouwman,[1] Marleen H. P. W. Visker, Johan A. M. van Arendonk, and Henk Bovenhuis
Animal Breeding and Genomics Centre, Wageningen University, PO Box 338, 6700 AH Wageningen, the Netherlands

ABSTRACT

A major quantitative trait locus (QTL) for milk fat content and fatty acids in both milk and adipose tissue has been detected on *Bos taurus* autosome 19 (BTA19) in several cattle breeds. The objective of this study was to refine the location of the QTL on BTA19 for bovine milk fat composition using a denser set of markers. Opportunities for fine mapping were provided by imputation from 50,000 genotyped single nucleotide polymorphisms (SNP) toward a high-density SNP panel with up to 777,000 SNP. The QTL region was narrowed down to a linkage disequilibrium block formed by 22 SNP covering 85,007 bp, from 51,303,322 to 51,388,329 bp on BTA19. This linkage disequilibrium block contained 2 genes: coiled-coil domain containing 57 (*CCDC57*) and fatty acid synthase (*FASN*). The gene *CCDC57* is minimally characterized and has not been associated with bovine milk fat previously, but is expressed in the mammary gland. The gene *FASN* has been associated with bovine milk fat and fat in adipose tissue before. This gene is a likely candidate for the QTL on BTA19 because of its involvement in de novo fat synthesis. Future studies using sequence data of both *CCDC57* and *FASN*, and eventually functional studies, will have to be pursued to assign the causal variant(s).
Key words: *Bos taurus* autosome 19, fine mapping, milk fatty acid, quantitative trait loci

INTRODUCTION

Many linkage and genome-wide association studies (**GWAS**) have been performed to identify QTL in cattle. These studies have detected numerous chromosomal regions affecting traits of interest (e.g., http://www.animalgenome.org/cgi-bin/QTLdb/index; Khatkar et al., 2004; Hu et al., 2010). Typically, the location of the QTL is estimated inaccurately and the confidence interval contains several candidate genes. To fine-map QTL, researchers have genotyped additional markers on the same animals, genotyped additional animals, and sometimes sequenced candidate genes (e.g., Grisart et al., 2002; Meuwissen et al., 2002; Blott et al., 2003; Cohen-Zinder et al., 2005; Gautier et al., 2006; Druet et al., 2008; Kim et al., 2009; Karim et al., 2011). Currently, opportunities for fine-mapping QTL in cattle are provided by the availability of high-density SNP panels with up to 777,000 SNP. Genotyping at higher density, or imputation of genotypes to higher densities, increases the power of GWAS, gives a more detailed view of associated regions, and increases the chance of one of the SNP being in strong linkage disequilibrium (**LD**) with the causal variant of the QTL (Marchini et al., 2007; Spencer et al., 2009; Marchini and Howie, 2010).

For milk FA, a few genome-wide linkage studies and GWAS have been performed. Those studies showed that 3 regions exist with major effects on milk FA, located on BTA14, BTA19, and BTA26 (Schennink et al., 2009b; Stoop et al., 2009; Bouwman et al., 2011, 2012). The region on BTA14 has been studied extensively and a dinucleotide polymorphism in diacylglycerol-O-acyltransferase 1 (*DGAT1*) has been suggested as the causal variant (Grisart et al., 2002; Schennink et al., 2007). For the region on BTA26, a polymorphism in stearoyl-CoA desaturase 1 (*SCD1*) has been suggested as causal variant (Taniguchi et al., 2004; Schennink et al., 2008). For the region on BTA19, the causal variant has not been identified.

The QTL on BTA19 shows association with several FA in both milk and adipose tissue (Morris et al., 2007; Bouwman et al., 2011, 2012; Ishii et al., 2013). Fatty acid synthase (***FASN***) has been suggested as a candidate gene and several SNP in *FASN* are significantly associated with FA in both beef and dairy cattle (Roy et al., 2006; Morris et al., 2007; Ordovás et al., 2008; Zhang et al., 2008; Abe et al., 2009; Schennink et al., 2009a; Oh et al., 2012), but the QTL region is rather large and shows much higher significance levels than those observed in the candidate gene studies (Bouwman et al., 2011). Therefore, the objective of this study was to refine the location of the QTL on BTA19 for bovine milk fat composition previously reported by Bouwman et al. (2011), using a denser set of markers.

Received June 29, 2013.
Accepted October 13, 2013.
[1]Corresponding author: Aniek.Bouwman@wur.nl

MATERIALS AND METHODS

Population

Detailed fat composition was measured in milk samples from 1,905 first-lactation Dutch Holstein-Friesian cows, which were housed on 398 commercial farms throughout the Netherlands. At least 3 cows were sampled per herd. Milk samples were taken in winter (February to March 2005), when Dutch cows are mainly kept indoors and fed silage. The cows were between 63 and 282 DIM at the day of sampling. About half of the sampled cows were descendants from 5 proven sires (101–200 daughters per sire); the other half of the sampled cows descended from 50 test sires or 45 other sires (1–30 daughters per sire). The pedigree of the cows was provided by the Cooperative Cattle Improvement Organization (**CRV**, Arnhem, the Netherlands) and consisted of 26,300 animals.

Phenotypes

Milk FA were measured by gas chromatography and were expressed in terms of weight-proportion of total milk fat weight (wt/wt%). More details about the phenotypes can be found in Stoop et al. (2008).

This study focuses on C14:0, because in previous studies, this milk FA showed the strongest association with the region on BTA19 (Morris et al., 2007; Stoop et al., 2009; Bouwman et al., 2011). Milk fat of the 1,905 cows contained, on average, 11.61% C14:0, the phenotypic standard deviation of C14:0 was 0.78, the heritability of C14:0 was 0.62, and the QTL on BTA19 explained approximately 13.8% of the genetic variation in C14:0 (Bouwman et al., 2011).

Genotypes

Initially, 1,810 cows and 55 of their sires (all proven and test sires) were genotyped with a custom 50,000 (**50K**)-SNP array (Illumina Inc., San Diego, CA) designed by CRV. The 55 sires were regenotyped using the BovineHD BeadChip (Illumina Inc.) with 777,000 (**777K**) SNP. The high-density (**HD**) genotypes of these 55 sires were combined with HD genotypes of other Dutch Holstein-Friesians available at CRV to form a reference population for imputation of, in total, 1,333 HD genotyped Dutch Holstein-Friesian animals.

Animals with pedigree inconsistencies (71 cows) were removed before imputation. Pedigree inconsistencies were assumed when more than 0.5% of the 50K SNP for which both sire and daughter were homozygous, were homozygous for the opposite allele. The software BEAGLE 3.3 (Browning and Browning, 2009) was used to phase and impute missing genotypes for the HD reference animals. These phased genotypes were then used to impute the 50K genotypes of the cows to HD genotypes. The assumed map positions of the SNP were based on the bovine genome assembly UMD 3.1 (Zimin et al., 2009).

In the 50K data, 1,454 SNP were located on BTA19, with an average distance of 44,888 bp. The number of SNP on BTA19 increased to 18,893, using the HD imputed data, with an average distance between SNP of 3,386 bp. We found 998 SNP overlapping between the 50K SNP panel and the HD SNP panel. A total of 1,572 monomorphic SNP in the HD-imputed data were excluded, and in addition 2,659 SNP were excluded because they had a low genotype frequency (i.e., 1–9 individuals within 1 of the 3 genotype classes), resulting in 14,662 SNP used in this study.

Single SNP Association Analysis

For 1,640 cows with both C14:0 phenotypes and HD-imputed genotypes, a single SNP analysis was performed for BTA19 using the following mixed model in ASReml software (VSN International Ltd., Hemel Hempstead, UK):

$$y_{ijklmno} = \mu + b_1 \times dim_i + b_2 \times e^{-0.05 \times dim_i} + b_3 \times afc_j + b_4 \times afc_j^2 + season_k + scode_l + herd_m + genotype_n + animal_o + e_{ijklmno}, \quad [1]$$

where $y_{ijklmno}$ was the phenotype; μ was the overall mean; b_1 to b_4 were regression coefficients of corresponding covariates; dim_i was the covariate describing the effect of DIM; afc_j was the covariate describing the effect of age at first calving; $season_k$ was the class variable accounting for calving season (June–August 2004, September–November 2004, or December 2004–January 2005); $scode_l$ was the class variable accounting for differences in genetic level between proven-sire daughters and test-sire daughters; $herd_m$ was the random effect of herd, distributed as $N\left(0, \mathbf{I}\sigma^2_{herd}\right)$, with identity matrix $\mathbf{I}$ and herd variance σ^2_{herd}; $genotype_n$ was the class variable accounting for the genotype of the SNP; $animal_o$ was the random additive genetic effect, distributed as $N\left(0, \mathbf{A}\sigma^2_a\right)$, with additive genetic relationship matrix $\mathbf{A}$ based on the full pedigree and additive genetic variance σ^2_a; and $e_{ijklmno}$ was the random residual, distributed as $N\left(0, \mathbf{I}\sigma^2_e\right)$, with identity matrix $\mathbf{I}$ and residual variance σ^2_e. To speed up the single SNP association analysis of all 14,662 SNP on BTA19, the genetic and herd variances were fixed to the variances estimated using model 1 without the genotype effect.