Genome-Wide Association Of The Ratio Of Saturated To Unsaturated Milk Fatty Acids In Dutch Dairy Cattle

A.C. Bouwman[*], H. Bovenhuis, M.H.P.W. Visker and J.A.M. van Arendonk

Introduction

Milk fat is an important substance of the human diet; it provides energy, fat-soluble nutrients (including essential nutrients) and bioactive lipids (German and Dillard, 2006). However, the consumption of saturated fatty acids (SFA) is associated with increased risk of cardio-vascular diseases, while unsaturated fatty acids (UFA) are neutral or even seen to decrease this risk (Jensen, 2002; Mensink et al., 2003). Dairy products have high amounts of SFA, especially C14:0 and C16:0, and relatively low amounts of (poly)UFA (Maijala, 2000; Mensink et al., 2003).

Stoop et al. (2008) as well as Soyeurt et al. (2008) showed that ratio SFA/UFA has a heritability of 0.20 and 0.27, respectively. The identification of genomic regions and individual genes responsible for this variation will enhance the understanding of biological pathways involved in FA synthesis and offers further opportunities to improve milk fat composition through breeding.

Recent developments in molecular genetics have made it possible to perform genome wide association studies using thousands of single nucleotide polymorphism (SNP) markers to detect quantitative trait loci (QTL). Schennink et al. (2009) performed genome wide linkage analysis for long-chain fatty acids, including ratio SFA/UFA. Using a denser marker map with linkage analysis provides little extra information about the position of a QTL, unless the number of recombinants is increased by increasing the population size (Darvasi et al., 1993). Genome wide association using a denser marker map has a higher power to detect QTL with a small effect on the trait and results in more precise locations. The aim of this study was to use a dense marker map to confirm previously detected QTL for ratio SFA/UFA in the linkage analysis by Schennink et al. (2009) and to position them to a more precise location, and to possibly identify additional QTL associated with ratio SFA/UFA.

Material and methods

Data. For this study phenotypic data was recorded from 1,905 first-lactation Dutch Holstein Friesian cows. Milk fat composition was measured by gas chromatography. More detailed information about the data and the gas chromatography can be found in Stoop et al. (2008). Individual SFA (C4:0-C20:0), mono-UFA (C10:1-C18:1), and poly-UFA (C18:2-C18:3) were measured. The fatty acids were expressed as weight-proportion of total fat weight. Ratio SFA/UFA was calculated as the sum of SFA divided by the sum of UFA. The cows were genotyped using a custom Infinium Assay (Illumina, USA). This approach resulted in 50,856 technically successful SNPs. Monomorphic SNPs, SNPs with a genotyping rate smaller than 80% and SNPs with a genotype frequency of less than 10 animals for one of the genotypes were discarded from the analysis resulting in 43,516 SNPs.

[*] Animal Breeding and Genomics Centre, Wageningen University, .P.O. Box 338, 6700 AH Wageningen, The Netherlands

Statistical analysis. A two-step single SNP association analysis was performed for ratio SFA/UFA. The first step involved a fast screen of the whole genome using a general linear model. In this first step the R package SNPassoc (González et al., 2007) was used. The phenotypes analyzed were pre-corrected for four systematic environmental effects: days in milk, age at first calving, calving season and herd. The correction factors were estimated on the phenotypes of all 1,905 cows using an animal model in ASReml:

$$y_{ijklmn} = \mu + b_1 * dim_i + b_2 * e^{-0.05*dim_i} + b_3 * afc_j + b_4 * afc_j^2 + season_k + scode_l + herd_m + A_n + e_{ijklmn} \quad (1)$$

where y was the (uncorrected) phenotype; μ was the overall mean; dim was the covariate describing the effect of days in lactation; afc was the covariate describing the effect of age at first calving; season was the fixed effect of the class of calving season (June-Aug '04, Sept-Nov '04, or Dec '04-Jan '05); scode was the fixed effect accounting for differences in genetic level between groups of proven bull daughters and young bull daughters; herd was the random effect of herd; A was the random additive genetic effect of the individual; and e was the random residual. For the SNPassoc analysis only the 1,706 individuals which had both phenotypic and genotypic information were used. Sire effect was included in the SNPassoc model to account for half sib-relations. The general linear model used was:

$$y_{ij}^{*} = \mu + sire_i + SNP_j + e_{ij}, \quad (2)$$

where y^* was the phenotype adjusted for the systematic environmental effects; sire was the fixed effect of sire; SNP was the fixed effect of SNP genotype; and e was the random residual. The genome-wise false discovery rate (FDR) was controlled according to the method of Storey et al. (2003), by calculating q-values based on the *P* values for each marker. Associations with a genome-wise q-value smaller than 0.05 were called significant.

In the second step regions with significant SNP effects were further analyzed using an animal model in ASReml (Gilmour et al., 2006), to account for all family relationships. For this analysis model (1) was extended with a SNP effect.

Results and discussion

A genome wide association study of ratio SFA/UFA was performed using 43,516 SNPs (figure 1). The SNPassoc analysis detected 209 genome-wise significant SNPs which were located on BTA 2, 3, 4, 5, 9, 12, 14, 19, 24 and 26. Associations on BTA 3, 5, 14 and 19 remained significant when accounting for all family relationships using model 2. The two most significant regions were BTA 14 and BTA 19.

The most significant SNP associated with ratio SFA/UFA was located at 0.4 Mbp on BTA 14 and explained 38.3% of the genetic variation of ratio SFA/UFA (table 1). This QTL confirms the genome-wise significant QTL detected by Schennink et al. (2009) at 0 cM for ratio SFA/UFA in their linkage study with 1,341 SNPs, on the same population, but with half the number of animals. The two most significant SNPs on BTA 14 were the two SNPs of the dinucleotide substitution of acylCoA:diacylglycerol acyltransferase (DGAT1) causing a K to A amino acid substitution, which has an effect on milk fat (Grisart et al., 2002) and milk fat composition, including ratio SFA/UFA (Schennink et al., 2007).

On BTA 19 the most significant SNP associated with ratio SFA/UFA was located at 37.3 Mbp and explained 5.3% of the genetic variation of ratio SFA/UFA (table 1). Previous studies have reported QTL for several milk fatty acids between 51 and 98 cM on BTA 19, but not for ratio SFA/UFA (Morris et al., 2007; Schennink et al., 2009; Stoop et al., 2009).

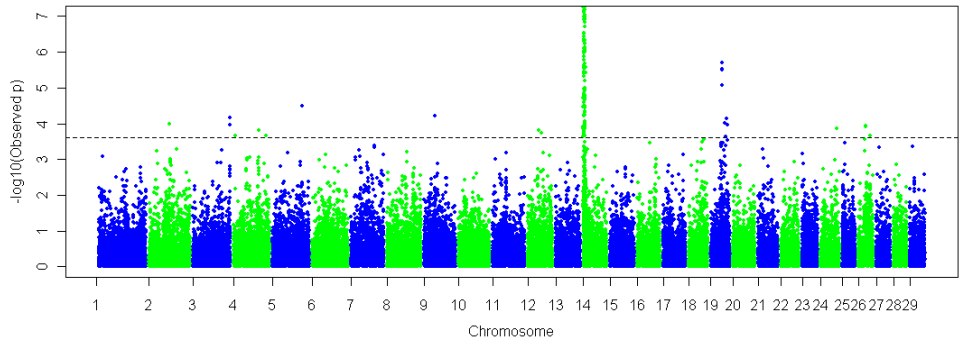

Figure 1: Manhattan plot for the association of 43,516 SNPs with ratio SFA/UFA. The dashed line corresponds to a threshold level of FDR = 0.05. All –log10(*P* values) > 7 are not shown

On BTA 19 there are many genes related to fat metabolism, such as sterol regulatory element-binding protein 1 (SREBP1), signal transducer and activator of transcription 5A (STAT5A), ATP citrate lyase (ACLY), growth hormone (GH) and fatty acid synthase (FASN). Therefore, this region was of further interest.
The most significant SNP on BTA 19 was located, together with 2 other significant SNPs, in the calcium channel, voltage-dependent, T type, alpha 1G subunit (CACNA1G). A SNP in the GH gene also showed a very significant association with ratio SFA/UFA. The SNP in the FASN gene showed no significant effect. None of the SNPs investigated were located in the other before mentioned genes related to fat metabolism.

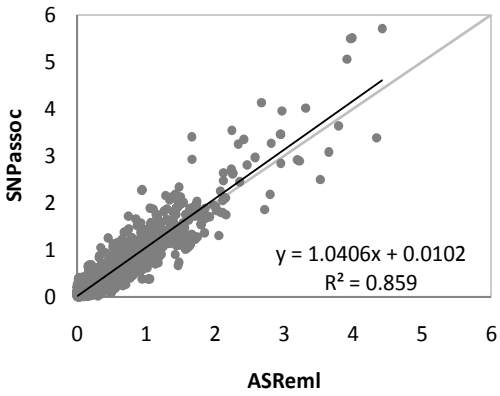

Figure 2: Regression plot of –log10(*P* values) from the SNPassoc and ASReml analysis for all 1,293 SNPs on BTA 19

All 1,293 SNPs on BTA 19 were re-analyzed in ASReml. The R^2 of the regression of the test-statistic of the SNPassoc analysis on the test-statistics of the animal model for all SNPs on BTA 19 was 0.86 (figure 2). This indicates that the SNPassoc analysis including a correction for sire gave similar results as the ASReml analysis accounting for all family relations. Therefore, SNPassoc can be considered as a good method for the determination of interesting regions. Of course the animal model in ASReml is a better model, but such an analysis would take too much time to run for many thousands SNPs.

These results show that there are markers significantly associated with ratio SFA/UFA, which can be exploited for selection in breeding programs. The genome wide association study also results in more precise locations, giving better opportunities to indicate candidate genes.

Table 1: Most significantly associated SNP (according to the SNPassoc analysis) per chromosome, their q-value, the –log10(*P* value) of the SNPassoc and ASReml analysis and the percentage of the genetic variance explained by the SNP

BTA	SNP Name	Position (Mbp)	q-value	-Log10(*P* value) SNPassoc	ASReml	Var[1] (%)
2	ULGR_BTA-101688	71.4	0.026	3.99	3.40	3.9
3	ULGR_BTA-122118	125.2	0.019	4.16	4.06	4.7
4	ULGR_rs29009595	87.5	0.035	3.81	2.21	2.8
5	ULGR_AAFC03122217_7089	99.9	0.009	4.50	3.89	5.0
9	ULGR_rs29013915	36.7	0.016	4.23	3.35	4.0
12	ARS-BFGL-NGS-82501	37.1	0.034	3.82	2.42	3.1
14	ULGR_SNP_AJ318490_1c	0.4	1.6E-26	30.10	30.61	38.3
19	ULGR_BTA-45363	37.3	0.001	5.71	4.42	5.3
24	ARS-BFGL-NGS-6396	56.5	0.032	3.85	3.56	2.1
26	ULGR_BTA-60969	24.6	0.029	3.93	3.35	3.9

[1] the percentage of the genetic variance explained by the SNP

References

Darvasi, A., Weinreb, A., Minke, V. *et al.* (1993). *Genetics* 134:943-951.

German, J.B., and Dillard, C.J. (2006). *Crit. Rev. Food Sci. Nutr.* 46:57-92.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. *et al.* (2006). ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, UK.

Grisart, B., Coppieters, W., Farnir, F. *et al.* (2002). *Genome Res.* 12:222-23.

Jensen, R.G. (2002). *J. Dairy Sci.* 85:295-350.

Maijala, K. (2000). *Livest. Prod. Sci.* 65:1-18.

Mensink, R.P., Zock, P.L., Kester, A.D.M. *et al.* (2003). *American Journal of Clinical Nutrition* 77(5):1146-1155.

Morris, C.A., Cullen, N.G., Glass, B.C. *et al.* (2007). *Mamm. Genome* 18:64-74.

Schennink, A., Stoop, W.M., Visker, M.H.P.W. *et al.* (2007). *Anim. Genet.* 38:467-473.

Schennink, A., Stoop, W.M., Visker, M.H.P.W. *et al.* (2009). *J. Dairy Sci.* 92:4676-4682.

Soyeurt, H., Dardenne, P., Dehareng, F. *et al.* (2008). *J. Dairy Sci.* 91:3611-3626.

Stoop, W.M., Van Arendonk, J.A.M., Heck, J.M.L. *et al.* (2008). *J. Dairy Sci.* 91:385-394.

Stoop, W.M., Schennink, A., Visker, M.H.P.W. *et al.* (2009). *J. Dairy Sci.* 92:4664-4675.

Storey, J.D., and Tibshirani, R. (2003). *PNAS* 100:9440-9445.