QTL Detection for Milk Fatty Acids in French Dairy Cattle

A. Govignon-Gion[1], **S. Fritz**[1,4], **H. Larroque**[2], **M. Brochard**[3], **C Chantry**[5], **F. Lahalle**[3, 6], **D. Boichard**[1]

[1]INRA, UMR1313 GABI, F78350, Jouy-en-Josas, [2] INRA, UMR1388 GenePhyse, F31326, Castanet, [3]Institut de l'Elevage, F75595 Paris, [4]UNCEIA, F75595 Paris, [5]LABOGENA, F78350, Jouy-en-Josas, [6]CNIEL, F75000, Paris

ABSTRACT: Milk fatty acid (FA) composition was estimated from mid infrared spectrometry (MIR) for more than 450,000 test-day from 86,458 cows in the three French Montbeliarde (MO), Normande (NO) and Holstein (HO) breeds, within the PhénoFinlait project. 23 FA expressed in percentage of milk or fat were analyzed. 6867 cows were genotyped with the 50K Illumina chip and QTL detection was carried out within breed by Linkage and Linkage Disequilibrium Analysis. In each breed, on average 11 QTL were detected per trait for saturated FA, vs 6 for unsaturated FA. Chromosomes 14, 5, and 19 affected many traits. Chromosomes 1, 11, 17, 26, and 27 were very influential on some traits. The same regions frequently affected many FA, as illustrated with chromosome 5. A high QTL conservation was observed across breeds.
Keyword: dairy cattle, milk fatty acids, QTL detection.

Introduction

Cow's milk is a major component of human food. Although the image of milk and dairy products is still very good, nutritional concern increases about the nature of milk fat. Milk fat contains a low proportion of unsaturated fatty acids (UNSAT) and a high proportion of saturated fatty acids (SAT) mainly C14:0 and C16:0. Long chain SAT have been associated with increased risk of cardiovascular disease (German et al. 2006).

One aim of the pluridisciplinary PhénoFinlait (PFL) project was to study milk composition in fatty acids (FA) in cattle. Equations were developed to predict FA (or FA combinations) from mid infrared spectrometry (MIR) (Ferrand et al., 2010). These equations were then used to build a database for milk fat composition from MIR spectra collected on a large scale in routine milk analysis laboratories. Blood samples were also collected by AI technicians for DNA extraction.

Genetic parameters were estimated by Gion et al (2011). In this paper, we present the results of QTL detection based on the analysis of FA data of about 7000 genotyped cows.

Materials and Methods

Animal and Milk samples. More than 425,000 milk samples were collected from 86,458 cows in 1,023 herds, between November 2009 and December 2010. Cows were distributed into the three main French dairy breeds, *i.e.* Montbéliarde (MO), Normande (NO) and Holstein (HO). Herds were selected according to their geographical locations, their history in milk recording, their size and their number of candidate cows with a first or second lactation during the experiment time frame and born from AI bulls in a preselected list. For each herd enrolled in the PFL project, in addition to conventional information of all test-day, the database contains MIR spectra for 6 PFL test-day on average. Calibration equations provide FA concentration in milk expressed in grams per 100g of milk. To avoid outliers, data deviating from more than 3 standard deviations were discarded. The FA, predicted with good accuracy (R^2>76%), were the following: C4:0, C6:0, C8:0, C10:0, C12:0, C14:0, C16:0, C18:0, SAT, C18:1cis9, C18:1cis12, C18:1trans11trans10, TotC18:1, TotC18:1cis, TotC18:1trans, C18:2cis9trans11, C18:2cis9cis12, C18:3n-3, TotC18:3, UNSAT, ω3, ω6, and total poly-unsaturated (POLY).

Data. Data were edited with the following additional rules: days in milk (DIM) between 8 and 350 days (370 in HO), at least 3 PFL test-day in the lactation, and sire known. For QTL detection and for each FA, one phenotype per cow was obtained by averaging the individual test-day records after adjustment for environmental effects (herd x test-day, month of calving and days in milk within year and parity). These effects were estimated in a linear model including also a polygenic and a permanent environment effect.

Cows were genotyped either with the Illumina 50K Beadchip® or with the Illumina LD Beadchip® (500) and then imputed. All genotypes were checked for quality control and phased in the French genomic evaluation system. QTL detection was carried out within breed with data from 2,186 HO, 2,029 MO and 2,652 NO cows.

Statistical analysis. Each of the 23 FA was analyzed using the combined linkage and linkage disequilibrium analysis method (LDLA) of Meuwissen et al. (2002). The whole genome was scanned with haplotypes defined with a sliding window of 6 consecutive markers. The likelihood ratio test (LRT) was used to detect significant QTL. It compares the likelihood between a model including a mean, a polygenic effect and a haplotype effect (H1 hypothesis) with a model without the haplotype effect (H0 hypothesis). A QTL was declared significant when the LRT exceeded 15, corresponding approximately to a 0.001 p-value. In addition to the LRT value, only the highest peak in an 8-Mb interval was considered as a QTL, resulting in a distance between two significant QTL of at least 4 Mb. The confidence interval of the QTL was defined with the lod drop-off method applying a 6.9 value in LRT difference.

Expression of FA content. FA can be expressed in proportion of fat or milk. Whereas the first one just illustrates the FA proportion in fat, the second one also accounts for the total amount of fat in milk. FA in fat of the diet is often considered by the milk industry and nutritionists as the trait of interest for human health.

However, a milk sample produced by one cow is not representative of the diet and the question of the unit of interest remains open. The expression in percentage of fat has an important drawback as it increases negative covariances between important components, because all FA sum to 100%. Therefore this expression tends to generate common QTL due to these covariances, regardless of their biological effects. In this study we present results for FA expressed in both expressions and differences in results are discussed.

Results and Discussion

Many detected regions. In total, for FA expressed in milk concentration, 564 regions were significantly detected for all traits and in the three breeds, ie 8.2 on average per trait and breed. The largest QTL number (42%) was found in NO breed which can be explained by a greater power due to larger genotyped population and also a higher fat concentration, leading to more accurate MIR prediction.
More QTL were detected for FA expressed in fat concentration. 647 regions were significant, ie 9.4 on average per trait and breed. Again, the largest QTL number (39%) was found in NO.
In FA expressed in milk, there were more QTL for saturated fatty acids (11.2 per trait and breed) than for mono-unsaturated (MONO) ones (6.2) and POLY (7.2). This can be explained by their higher proportion in milk and their higher heritability. However, POLY displayed similar results as MONO ones, in spite of their low concentration in milk. In FA expressed in fat percent, results were quite similar for SAT and MONO, but higher for POLY (10.2).

Detailed description of NO results. Figure 1 presents Manhattan plots and illustrates the effect of the different QTL in NO breed for FA expressed in percentage of fat. It is important to notice that in NO breed the DGAT1 K232A mutation does segregate only at a low frequency (0.07).
A pattern of emerging and disappearing QTL was very clear along the chain of the saturated FA, showing that their variations are not influenced by the same genes all along the elongation pathway. BTA17 and 20 were most influential for short FA, especially for C4:0 to C8:0. BTA16 and 29 were especially important for C4:0 and C6:0, respectively. BTA19 strongly influenced C8:0 to C14:0, the main QTL being located in the region of FASN. BTA13 and 27 affected specifically C10:0 and BTA6 was found to significantly influence C12:0 and C14:0. Other QTL were also detected but cannot be described here. Surprisingly, QTL detected for C16:0 were substantially different, because the most significant ones were BTA14, 5, as well as BTA20 and 27. BTA14 and 5 did not appear important in the previous elongation steps. Indeed C16:0 is the final product and therefore an accumulating product and not a substrate, in contrast with C4:0-C14:0. MONO and POLY were mainly influenced by BTA5, 14, and 27. ω3 and ω6 shared BTA11 (close to BLG), 14 and 27 (close to AGPAT6) as major QTL, whereas BTA13 and BTA5 were more specific of ω3 and ω6, respectively.

Results on FA expressed in percentage of milk were quite different and much more homogeneous across traits. The major difference concerned BTA14 which affected all analyzed traits and was always found to be the most significant QTL. BTA5 was also found to have a broad spectrum, except on POLY which, in contrast, were dependent on BTA26 (SCD gene). As previously, C10:0-C14:0 were found to be influenced by BTA19, C4:0-C6:0 by BTA17 and 20, and C16:0 by BTA27. BTA6 also affected many POLY-.

Comparison with Holstein and Montbéliarde. On average, results in HO and MO were rather consistent with those in NO. The same major QTL on chromosomes 14, 5, 19 were shared across breeds, as well as chromosomes 11, 20, 27 for some FA, and confirmed the results of Bouwman et al. (2011).
As a major difference, for FA expressed in fat percentage, DGAT1 was always the major QTL for nearly all traits in HO whereas it was rarely observed in NO, the MO situation being intermediate. It should be noticed that the K232A mutation is rare in NO and absent from MO, although some other mutations close to the DGAT1 gene are suspected.
In HO and MO BTA19 also affected many UNSAT (expressed in fat percentage) whereas its effect was limited to medium-chain saturated FA in NO. BTA6 played an important role in unsaturated FA in MO with at least 3 different regions detected. Similarly BTA1 and 2 appeared especially important in MO breed.
As for NO, QTL were more diverse for percentages in fat than in milk. For instance in HO BTA1 was found to affect ω3 and not ω6. BTA2, 9, 17, 25 influenced C4:0 and C6:0 whereas they had limited effect on concentrations in milk. As a result, the rate of QTL concordance between breeds was rather high, much higher than usually for many other traits. The most important QTL were shared across the three breeds (even DGAT1, likely with different mutations), and many others were shared between two breeds.

QTL Co-location between FA. Many QTL regions were found to be common for several FA. The best example is with BTA14 which affected nearly all traits defined in concentration in milk in the three breeds. The region on BTA 5 around 92 Mb was shared by up to 17 different FA in each breed. This highly significant region was observed when FA were expressed both in percentage of fat (Boichard et al, 2014) or milk. Moreover, in the three breeds, a second important region was also observed around 110 Mb with a strong co-location for many FA.
On average, co-location between FA was stronger when FA were expressed in percentage of milk. In percentages of fat, following the genetic correlations, the QTL were more diverse. This was especially clear for short and medium saturated FA and the different steps of the elongation chain are not influenced by the same QTL. Similarly, Omega3 and Omega6 have different and highly significant QTL, although one could have anticipated high prediction correlations through MIR data. These observations let us suppose that MIR predictions are of good quality.

Conclusion

Many highly significant QTL were detected for milk fatty acids in NO, MO and HO breeds. Confidence intervals of locations were narrow. Many QTL seem to be shared across breeds. These results open the way for new mutations discovery. A good accuracy of genomic selection is expected due to a low number of strong QTL and a rather high heritability of the traits.
See http://www.phenofinlait.fr/ for the *PhénoFinlait* project.

Acknowledgments. The PhénoFinlait project received financial support from ANR (ANR-08-GENO-005), Apis-Gène, CASDAR, CNIEL, FranceAgriMer, France Génétique Elevage and Ministry of Agriculture.

Acknowledgments to the farmers who participated in PhénoFinlait, INRA and Institut de l'Elevage colleagues who coordinated samples and data collection, the laboratories, manufacturers (Foss and Bentley) and DHI organizations which provided data.

Literature Cited

Boichard D., Govignon-Gion A., Larroque H., et al (2014) INRA Prod Anim (in press).

Bouwman A., Bovenhuis H., Visker M.H.P.W, et al (2011) BMC Genet 12,43

Ferrand M., Huquet B., Barbey S., et al (2010) Chemometr. Intell. Lab. Syst. 106, 183-189

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

Gion A., Larroque H., Brochard M., et al (2011) Interbull Bull 44, 185-189.

Meuwissen T.H.E., and Goddard M.E. (2001) Genet Sel Evol, 33, 605-634

Figure 1: Manhattan plot for 9 fatty acids expressed in percentage of fat in Normande breed.