Characterization of the DGAT1 K232A and variable number of tandem repeat polymorphisms in French dairy cattle.

Mathieu M. Gautier, Aurelien A. Capitan, Sebastien S. Fritz, Andre A. Eggen, Didier Boichard, Tom T. Druet

HAL Id: hal-02654424

https://hal.inrae.fr/hal-02654424

Submitted on 29 May 2020

Characterization of the *DGAT1* K232A and Variable Number of Tandem Repeat Polymorphisms in French Dairy Cattle

M. Gautier,[*1] **A. Capitan,*** **S. Fritz,†** **A. Eggen,*** **D. Boichard,‡** **and T. Druet‡**
*INRA, UR339 Laboratoire de Génétique Biochimique et Cytogénétique, F-78352 Jouy-en-Josas, France
†Union Nationale des Coopératives agricoles d'Elevage et d'Insémination Animale, F-75595 Paris, France
‡INRA, UR337 Station de Génétique Quantitative et Appliquée, F-78350 Jouy-en-Josas, France

ABSTRACT

A quantitative trait locus (QTL) underlying different milk production traits has been identified with a high significance threshold value in the genomic region containing the acylCoA:diacylglycerol acyltransferase (*DGAT1*) gene, in the 3 main French dairy cattle breeds: French Holstein, Normande, and Montbéliarde. Previous studies have confirmed that the K232A polymorphism in *DGAT1* is responsible for a major QTL underlying several milk production traits in Holstein dairy cattle and several other bovine breeds. In this study, we estimate the frequency of the 2 alternative alleles, K and A, of the K232A polymorphism in French Holstein, Normande, and Montbéliarde breeds. Although the K allele segregates in French Holstein and Normande breeds with a similar effect on production traits, the existence of additional mutations contributing to the observed QTL effect is strongly suggested in both breeds by the existence of sires heterozygous at the QTL but homozygous at the K232A polymorphism. One allele at a variable number of tandem repeats (VNTR) locus in the 5′ noncoding region of *DGAT1* has been recently proposed as a putative causative variant. In our study, this marker was found to present a high mutation rate of 0.8% per gamete and per generation, making the allele diversity observed compatible with that expected under neutrality. Moreover, among the sires homozygous at the K232A polymorphism, no allele at the VNTR can fully explain their QTL status. Finally, no allele at the VNTR was found to be significantly associated with the fat percentage variation in the 3 breeds simultaneously after correction for the effect of the K232A polymorphism. Therefore, our results suggest the existence of at least one other causative polymorphism not yet described. Because the A allele is nearly fixed in the Montbéliarde breed, this breed represents an interesting model to identify and confirm other mutations that have a strong effect on milk production traits.

Key words: quantitative trait locus, dairy trait, acyl-CoA:diacylglycerol acyltransferase gene (*DGAT1*), variable number of tandem repeats

INTRODUCTION

A QTL with a major influence on several milk production traits, and particularly on fat content, has been identified at the centromeric end of bovine chromosome 14 (BTA14; Coppieters et al., 1998; Heyen et al., 1999; Looft et al., 2001; Boichard et al., 2003). Furthermore, fine mapping experiments have identified a nonconservative K232A polymorphism inside the acylCoA:diacylglycerol acyltransferase gene (**DGAT1**), which has been proposed as the causative mutation or quantitative trait nucleotide (**QTN**) explaining all the variation due to the QTL (Grisart et al., 2002). Additional studies have demonstrated the effect of this mutation on the function of the resulting enzyme, which plays a fundamental role in triacylglycerol metabolism (Grisart et al., 2004).

Other studies have shown that the 2 alleles at the K232A polymorphism segregate in several breeds from different countries (Kaupe et al., 2004). The K232A mutation probably occurred early in the history of cattle domestication or even before domestication (Winter et al., 2002; Kaupe et al., 2004). Nevertheless, the estimated frequency of the wild-type lysine variant (the K allele) varies greatly according to the population considered, which might reflect different breeding objectives regarding milk composition in different countries and breeds. The effect of the mutation on milk production traits has been extensively characterized and the K allele is always associated with an increase in fat content, fat yield, and protein content, and a decrease in protein and milk yields.

Evidence for the existence of additional mutations contributing to the observed QTL effect has been reported recently (Bennewitz et al., 2004). In particular, one allele at a variable number of tandem repeat (**VNTR**) locus located in the 5′ noncoding region of *DGAT1* has been

Table 1. Description of the 3 breed-specific granddaughter designs used in the study

Breed	Number of families (sires)	Average family size	Number of progeny-tested sons
Montbéliarde	6	65 (41 to 98)	387
Normande	9	60 (38 to 74)	541
Holstein	26	81 (34 to 231)	2,097
All	41[1]	74 (34 to 231)	3,025

[1]Seventeen sires were included in the design as sons; thus, a total of 3,049 (3,025 + 24) individuals were used.

proposed as a putative causative variant (Kuhn et al., 2004).

In this paper, we report the characterization of both the VNTR and the K232A polymorphisms in the *DGAT1* gene in the 3 main French dairy cattle breeds (Holstein, Normande, and Montbéliarde). We also investigate the existence of putative causative alleles at the VNTR locus.

MATERIALS AND METHODS

Experimental Data

Experimental Design. As detailed in Table 1, the animals considered in this study are organized according to 3 breed-specific granddaughter designs (**GDD**; Weller et al., 1990) including 6, 9, and 26 half-sib families for the Montbéliarde, Normande, and Holstein breeds, respectively. A total of 3,025 progeny-tested sons were used. In addition to the 14 initial families (Boichard et al., 2003), 27 new families (4, 6, and 17 in the Montbéliarde, Normande, and Holstein breeds, respectively) derived from the French marker-assisted selection (**MAS**) program (Boichard et al., 2002) were integrated in this design.

Daughter yield deviations (**DYD**) for 5 milk production traits (milk, fat, and protein yields and fat and protein percentages) and their corresponding reliabilities were obtained from the French national evaluation. DNA from all bulls had been extracted during the previously mentioned studies (Boichard et al., 2002, 2003) and was available.

Genetic Markers. Genotypes for 2 bovine microsatellite markers CSSM066 and ILSTS039 were available from the MAS and initial QTL programs (Boichard et al., 2002, 2003). However, in the 2003 QTL program, animals had been genotyped only for CSSM066; thus, 2,941 (97%) and 1,514 (50%) sons were genotyped for CSSM066 and ILSTS039, respectively.

For the VNTR, the genotyping procedure consisted of a fluorescent PCR amplification with the following primers VNTR_F: TCAGGATCCAGAGGTACCAG and VNTR_R: CAATGAGAAGGCACGGACTGTGAA. For a

5-repeat allele, the length of the resulting PCR product was 228 bp. The PCR reactions were performed using the Go-Taq Flexi (Promega, Paris, France) according to the manufacturer's recommendations on a PTC-100 thermocycler (MJ Research, Paris, France) in a final volume of 10 μL. Samples were preheated for 5 min at 95°C, subjected to 40 cycles (95°C for 30 s, 58°C for 30 s, and 72°C for 30 s), and a final extension step of 5 min at 72°C. The resulting 228-bp PCR products were purified on a Sephadex G50 column (Amersham, Paris, France) before running on a Megabace 96 capillary sequencer (Molecular Dynamics, Paris, France). Raw data were then analyzed using Genetic Profiler v1.5 (Molecular Dynamics). Thereafter, alleles were named according to the number of repetitions (**R**) of the 18-bp motif. Some genotyped individuals were sequenced to confirm the expected number of repeats. In total, 2,958 (98%) progeny-tested sons were genotyped for this marker.

The K232A polymorphism was analyzed using the TaqMan allelic discrimination technology according to the standard genotyping protocol. The TaqMan probes (CCGTTGGCCGCCTT and CCGTTGGCCTTCTTA specific for the A and K alleles, respectively) and primers (DGAT1-F: CGCTTGCTCGTAGCTTTGG and DGAT1-R: CCGCGGTAGGTCAGGTTGT) were provided by Applied Biosystems (Paris, France). This procedure genotyped 3,020 (100%) progeny-tested sons for this marker.

For some nongenotyped sires from the GDD, marker genotypes were reconstructed based on the genotypes of their sons. This provided 24 additional genotyped individuals for the VNTR, the K232A polymorphism, and CSSM066, and 14 more for ILSTS039.

Pedigrees. Extended pedigrees were used for mixed model analyses. In addition to the genotyped animals, their ancestors over 3 generations were included in the pedigree. Noninformative ancestors; that is, ancestors with no parents and at most one progeny, were deleted iteratively. The size of these pedigree files was 1,006, 1,427, and 4,833 animals for the Montbéliarde, Normande, and Holstein breeds, respectively. The data files contained 422, 578, and 2,259 animals with records, respectively.

Statistical Analyses

Estimation of Allelic Frequencies. Estimation of allelic frequencies was done on the founder alleles. In the 3 different breed-specific GDD, these corresponded to the 2 sire alleles and each son's maternally inherited alleles.

Linkage Map Construction. Marker order and map distances were estimated using CRIMAP 2.4 software (Green et al., 1990). The CHROMPIC option was subsequently used to identify unlikely double crossovers. Final

map distances were computed based on Haldane's mapping function.

QTL Mapping Analysis. Linkage analysis was performed using a regression interval analysis (Haley and Knott, 1992) with the web-based version of the QTL Express software (Seaton et al., 2002) fitting a 1-QTL model every centimorgan along the considered region. Records used were equal to twice the DYD of the bulls, weighted by their respective reliability. A 95% chromosome-wise significance threshold was computed based on 10,000 permutations. The *P*-value associated with the null hypothesis of a QTL homozygosity was calculated based on the *t*-test value at the maximum peak position. The rejection threshold was set to 5%.

VNTR and K232A Association Tests. Association tests were performed considering the 2 markers simultaneously with the following model:

$$y_{ij} = s_i + n_{kj}b_k + n_{vj}b_v + e_{ij}$$

where y_{ij} is the DYD of son j of sire i, s_i is the fixed effect of sire i, n_{kj} is the number of copies of the K allele carried by son j at the K232A polymorphism, n_{vj} is the number of copies of the VNTR allele considered, b_k and b_v are, respectively, half the allele substitution effect of the K allele at the K232A polymorphism and the VNTR allele, and e_{ij} is the residual effect associated to this record. The weight of each record was computed as already described (Kuhn et al., 2004). Because more than 2 alleles were segregating for the VNTR, all alleles with a frequency greater than 0.05 were successively tested using the above model.

Estimation of the Effect of the K232A Polymorphism. To estimate the effect of the K232A polymorphism, the traits were modeled separately using the following linear mixed model:

$$\mathbf{y} = \boldsymbol{\mu} + n_k\mathbf{b_k} + \mathbf{Zu} + \mathbf{e}$$

where **y** is a vector containing twice the DYD for bulls, **μ** is a vector with the fixed effect of the mean, n_k is a vector containing the number of copies of the K allele (0, 1, or 2) corresponding to the DYD, **b_k** is the allele substitution effect, **u** is a vector of random polygenic effects, and **e** is a vector of random residual effects; **Z** is a design matrix relating the DYD of the individuals to their corresponding polygenic effect.

The (co)variance structure was:

$$\text{var}\begin{bmatrix}\mathbf{u}\\\mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\sigma_a^2 & 0\\ 0 & \mathbf{R}\sigma_e^2\end{bmatrix}$$

where σ_a^2 is the variance associated to the random polygenic effect and σ_e^2 the residual variance; **A** is the additive

relationship matrix and **R** is a diagonal matrix containing the inverse of the weight of the corresponding DYD. These weights were daughter equivalents and were estimated as previously described (VanRaden and Wiggans, 1991) with a correction for the number of cows per herd.

These analyses were performed using the ASREML package (Gilmour et al., 2000).

Haplotype Analysis and Linkage Disequilibrium Estimation. The most likely haplotype phases of each individual from the different breed-specific pedigrees were computed using the deterministic algorithm already described (Boichard et al., 2003). Estimation of haplotype frequencies and linkage disequilibrium (**LD**) was performed on the founder haplotypes (see above).

Pairwise LD measures were derived from the standard measure of LD between 2 alleles at 2 different loci: $D_{ij} = p(A_iB_j) - p(A_i)p(B_j)$, where $p(A_i)$ is the frequency of allele A_i at locus A, $p(B_j)$ the frequency of allele B_j at locus B and $p(A_iB_j)$ the frequency of haplotype A_iB_j as estimated from the population. Three multiallelic LD measures derived from this standard measure were considered in the study.

First, multiallelic D′ was measured as (Hedrick, 1987):

$$D' = \sum_{i=1}^{k}\sum_{j=1}^{1} p(A_i)p(B_j)\left|\frac{D_{ij}}{D_{ij}^{max}}\right|,$$

where k and l are the number of alleles for marker A and B, respectively, and

$$D_{ij}^{max} = \min\{p(A_i)p(B_j),$$
$$[1 - p(A_i)][1 - p(B_j)]\} \text{ when } D_{ij} < 0,$$

$$D_{ij}^{max} = \min\{p(A_i)[1 - p(B_j)],$$
$$p(B_j)[1 - p(A_i)]\} \text{ when } D_{ij} \geq 0.$$

Second, multiallelic r^2 was measured as (Hill and Robertson, 1968)

$$r^2 = \sum_{i=1}^{k}\sum_{j=1}^{1} \frac{D_{ij}^2}{[1 - p(A_i)][1 - p(B_j)]}.$$

Finally, the multiallelic χ^2 statistic is defined as (Hill, 1975)

$$\chi^2 = N \sum_{i=1}^{k}\sum_{j=1}^{1} \frac{D_{ij}^2}{p(A_i)p(B_j)},$$

where N is the number of haplotypes considered to compute the corresponding pairwise measure. χ^2 is asymp-

Table 2. Frequency of the K allele and genotypes of animals with records in the 3 breeds

Breed	Frequency of allele K	Genotype		
		KK	AK	AA
Montbéliarde	0.040	0	14	370
Normande	0.130	11	122	402
Holstein	0.369	177	944	965

totically distributed according to a χ^2 distribution with $(k-1)(l-1)$ degrees of freedom. The threshold for significance of LD was set to 0.001.

To limit bias introduced by rare alleles in the estimation of the different measures and to preserve power to reject the null hypothesis of absence of LD, marker alleles with frequency <0.05 were pooled.

RESULTS

Characterization of the K232A and VNTR Polymorphisms

The lysine variant (K allele) at the K232A polymorphism was present in the 3 breeds and its frequency ranged from 0.04 in the Montbéliarde breed to 0.13 in the Normande breed and 0.37 in the Holstein breed (Table 2). Distributions of genotypes are presented in Table 2 for the 3 breeds. Only 14 out of the 384 Montbéliarde bulls were heterozygous for K232A and none was homozygous for the lysine variant (KK). Indeed, most animals were homozygous for the alanine variant (AA) in both the Montbéliarde and Normande breeds. In the Holstein breed, where the K allele is much more frequent, approximately 10% of the bulls were homozygous (KK) and 43% were heterozygous.

Seven different alleles (from 2 to 8 repetitions of the 18-bp motif) at the VNTR were identified in the present study with corresponding frequencies varying according to breed (Table 3). The 2R allele segregated both in the Montbéliarde and Normande breeds and was not reported in previous studies (Kuhn et al., 2004; Sanders et al., 2006). For Holsteins, the VNTR allele frequencies varied within a similar range in Germany and France (Table 3).

Map Construction and Estimation of the VNTR Mutation Rate

When considering all families simultaneously, 1,803, 472, 951, and 1,849 informative meioses were available for the VNTR, K232A, ILSTS039, and CSSM066 loci, respectively, to build a genetic map. This order was found to be the most likely and is in agreement with previous studies. Nevertheless, several inconsistencies appeared in the segregation of the VNTR when analyzing genotyping raw data. In particular, genotypes of 4 sons descending from 3 different homozygous sires (with 5R/5R genotypes) were not compatible for this marker. Filiations of these sons were unambiguously assessed based on genotyping data from other projects (data not shown). A total of 506 genotyped sons descended from a homozygous sire at the VNTR. Assuming the absence of genotyping errors, the estimated mutation rate for the VNTR could be set to 0.8% (4/506) per paternal gamete. This value was consistent with known features for a VNTR-type locus in other mammalian species (Bois and Jeffreys, 1999). Additionally, based on further haplotype analyses, 3 mutational events out of the 4 identified were characterized and corresponded to repetition gains (from 5R to 6R).

The estimated 2-point Haldane genetic distances for each pair of markers are given in Table 4. It is worth noting that the VNTR and K232A polymorphisms were less than 10 kb apart. Thus, distances appeared strongly inflated for markers linked to the VNTR. Respectively, 1 recombinant out of 289 coinformative meioses for the VNTR/K232A marker pair was observed, 16 out of 522 for the VNTR/ILSTS039 marker pair, and 114 out of 1,155 for the VNTR/CSSM066 marker pair (131 in total). A small fraction of these recombinant gametes might have been created by mutational events. Only paternal gametes are informative in our design. Consequently, a mutation can create a recombinant gamete only if the maternally inherited allele is identical in state to one of the paternal alleles, or if the mutated allele is identical in state to the alternative paternal allele. Therefore, the observed inflation in genetic distances was probably the result of a recombination hot spot in this centromeric region and an underestimation of genetic distances when considering the K232A polymorphism due to a lack of informative meioses.

QTL Analysis. The QTL was detected in the Holstein breed for the 5 traits ($P < 10^{-4}$) whereas in the Montbéliarde and Normande breeds, the QTL was significant only for fat percentage ($P < 10^{-4}$). The heterozygous status at the QTL was determined using fat percentage as a reference in the 3 breeds, because this trait shows the most pronounced QTL effect. As shown in Table 5, all the Montbéliarde sires were homozygous AA at the K232A polymorphism and 5 were heterozygous at the QTL. In the Normande breed, only sire #1611 (out of the 3 sires that were heterozygous at the QTL) was also heterozygous at the K232A polymorphism. In the Holstein breed, the 11 sires heterozygous at the K232A polymorphism were also heterozygous at the QTL and 2 sires homozygous at the mutation were heterozygous at the QTL.

As previously reported (Bennewitz et al., 2004), the existence of sires heterozygous at the QTL and homozy-

Table 3. Allele frequencies of the variable number of tandem repeat polymorphism (VNTR) in the 3 breeds and comparisons with previous reports

VNTR allele[1]	Frequency				
	Angeln[2]	German Holstein[3]	Holstein	Normande	Montbéliarde
2R	0	0	0	0.002	0.001
3R	0.010 (A)	0.017 (1)	0.057	0.279	0.003
4R	0.250 (B)	0.169 (2)	0.172	0.478	0.442
5R	0.300 (C)	0.477 (3)	0.459	0.082	0.501
6R	0.060 (D)	0.175 (4)	0.204	0.155	0.048
7R	0.380 (E)	0.163 (5)	0.107	0.004	0.005
8R	0.003 (F)	0	0.001	0	0

[1]The names of the alleles are given according to the number of repetitions (R) of the 18-bp motif.

[2]From Sanders et al. (2006); the names of the alleles in this publication are indicated in parentheses.

[3]From Kuhn et al. (2004); the names of the alleles in this publication are indicated in parentheses. Allele frequencies from this study correspond to those estimated on the basis of maternally inherited haplotypes (individuals belonging to a granddaughter design).

gous at the K232A polymorphism indicates that this latter mutation is not solely responsible for the QTL variation. The VNTR has been proposed as a putative candidate locus. In the Holstein and Normande breeds, the 4 sires (in Normande: #1406 and #1788; in Holstein: #3498 and #3501) homozygous AA at the K232A polymorphism and heterozygous at the QTL have the 4R allele at the VNTR in common. Nevertheless, 4 Holstein sires (#3442, #3480, #3481, and #6145) and 3 Normande sires (#1961, #2409, and #2618) homozygous (AA) at the K232A polymorphism carried 1 copy of this same allele at the VNTR but were homozygous at the QTL. In the Montbéliarde breed, among the 5 sires homozygous (AA) at the K232A polymorphism and heterozygous at the QTL, 4 (#1298, #1312, #1368, and #1531) carried 1 copy of the 5R and the 4R allele at the VNTR and the other sire (#1103) was homozygous for the 5R allele at the VNTR. The 7R allele, which was proposed as having the most pronounced effect in German Holsteins (Kuhn et al., 2004) and in the Angeln breed (Sanders et al., 2006), was not carried by any Normande and Montbéliarde sires. Out of the 6 Holstein sires (#3501, #3442, #3480, #3500, #3514, #4729, and #6865) homozygous at the K232A polymorphism and carrying 1 copy of this allele at the VNTR, only sire #3501 was heterozygous at the QTL.

VNTR and K232A Association Tests. Both the VNTR and K232A polymorphisms were previously suggested as causative. Thus, we performed association tests considering the 2 markers simultaneously to test the effect of the different VNTR alleles after correction for the K232A effect. As expected, the allele substitution effect for the K232A polymorphism was highly significant ($P < 10^{-4}$) for fat content in the 3 breeds. In the Montbéliarde breed it was also significant ($P < 0.05$) for milk yield and protein content. In the Normande and Holstein breeds, the effect was significant ($P < 10^{-4}$) for all traits except for protein yield in the Normande breed.

No allele at the VNTR was associated with trait variation across all 3 breeds. Nevertheless, in the Montbéliarde breed, the 4R allele was associated with milk and protein yield ($P < 0.01$), fat content ($P < 10^{-4}$), and protein content ($P < 0.05$) variations, whereas in the Normande breed, only the 3R allele was associated with protein yield variation ($P < 0.05$). Finally, as observed in the German Holstein population (Kuhn et al., 2004), the 7R allele in the French Holstein breed was significantly associated with an increase in fat yield ($P < 0.05$), protein content ($P < 0.05$), and fat content ($P < 10^{-4}$).

Characterization of the Effect of the K232A Polymorphism. The average allelic substitution effect of the K232A polymorphism for dairy traits is presented in Table 6. The allelic substitution effect was estimated using twice the DYD. Therefore, it indicates the difference in terms of EBV and not of PTA. In the Holstein breed, the K allele increased fat yield (+14.90 kg) and fat (+0.34 %) and protein (+0.08%) contents, whereas it decreased milk (–351.2 kg) and protein (–4.6 kg) yields. These results are consistent for the 3 breeds but the magnitude of substitution effects varied. These estimated average allelic substitution effects were in agreement with previous studies (Grisart et al., 2002; Spelman et al., 2002; Thaller et al., 2003; Weller et al., 2003).

The variance due to the QTL can be estimated as $2p(1 - p)a^2$, where a is the additive effect and p is the

Table 4. Two-point and multipoint (in parentheses) Haldane genetic distances (cM) for the different marker pairs

Marker name	K232A	ILSTS039	CSSM066
VNTR	0.035 (0.035)	3.16 (2.61)	11.0 (10.2)
K232A		1.61 (2.57)	8.82 (10.1)
ILSTS039			6.96 (7.54)

Table 5. Results for the within-family QTL analysis for milk fat percentage trait

Breed	Family	K232A	VNTR_DGAT haplotype	t-value for fat yield
Montbéliarde	1103	A/A	5R_A/5R_A	2.15*
	1298	A/A	5R_A/4R_A	2.63**
	1312	A/A	5R_A/4R_A	3.25**
	1368	A/A	5R_A/4R_A	2.01*
	1531	A/A	5R_A/4R_A	3.23**
	1736	A/A	4R_A/4R_A	0.33
Normande	1406	A/A	4R_A/3R_A	2.05*
	1788	A/A	4R_A/6R_A	3.48**
	1611	K/A	4R_A/4R_K	4.37**
	1528	K/K	5R_K/4R_K	0.23
	1537	A/A	3R_A/6R_A	0.43
	1855	A/A	4R_A/4R_A	0.2
	1961	A/A	4R_A/3R_A	0.88
	2409	A/A	4R_A/3R_A	1.59
	2618	A/A	4R_A/3R_A	0.39
Holstein	3498	A/A	4R_A/6R_A	2.08*
	3501	A/A	4R_A/7R_A	4.7**
	2108	K/A	5R_K/7R_A	4.53**
	6146	K/A	5R_K/7R_A	6.69**
	3465	K/A	5R_K/4R_A	5.11**
	3468	K/A	5R_K/3R_A	8.07**
	3487	K/A	5R_K/6R_A	8.08**
	3492	K/A	5R_K/4R_A	3.98**
	3502	K/A	5R_K/6R_A	9.87**
	4442	K/A	5R_K/4R_A	6.42**
	5245	K/A	5R_K/5R_A	5.84**
	5970	K/A	5R_K/6R_A	6.85**
	5974	K/A	5R_K/5R_A	7.32**
	3442	A/A	7R_A/4R_A	1.86
	3480	A/A	7R_A/4R_A	0.93
	3500	A/A	7R_A/6R_A	0.56
	3514	A/A	7R_A/6R_A	0.51
	4729	A/A	7R_A/5R_A	0.19
	3481	A/A	4R_A/6R_A	0.46
	5972	A/A	5R_A/6R_A	1.79
	6109	A/A	4R_A/4R_A	1.49
	6143	A/A	5R_A/3R_A	1.85
	6145	A/A	4R_A/3R_A	0.27
	6151	A/A	5R_A/6R_A	0.74
	6865	A/A	6R_A/6R_A	—[1]
	7210	K/K	5R_K/5R_K	—

[1]Sires are homozygous for the 4 markers and thus could not be analyzed.

*$P < 0.05$; **$P < 0.01$.

frequency of one allele. The variances obtained with this approximation are presented for the Holstein breed, expressed as a percentage of the total genetic variance, in Table 6. The variances range from 2.4 to 42.5% of the total genetic variance for protein yield and fat content, respectively.

Haplotype Analysis and D′ Among Markers. As shown in Table 7, all markers were found to be in significant LD with each other among the 3 different populations considered. As expected, irrespective of the multiallelic measure considered (r^2 or D′), LD decreases when the genetic distance separating the markers increases. Interestingly, r^2 values for the 3 marker pairs involving the K232A polymorphism are all far from 1, indicating

that no allele at the 3 other markers (even the VNTR located only 10 kb away) behaved as a perfect surrogate for the K232A alleles. For instance, as shown in Table 8, in the Montbéliarde breed the K allele at the K232A polymorphism was always associated with the 5R allele at the VNTR (the D′ value is thus equal to 1), but an almost equal fraction of this latter allele at the VNTR was associated with the A allele at the K232A polymorphism (the r^2 value is thus far from 1). Similarly, in French Holsteins, as previously observed in German Holsteins (Kuhn et al., 2004), the K allele was almost always associated with the 5R allele at the VNTR (D′ = 0.98). In the Normande breed, the K allele was associated mainly with the 4R allele at the VNTR, whereas a much smaller fraction was also associated with the 5R allele (D′ = 0.68).

DISCUSSION

In our study, a QTL underlying fat percentage was detected in the 3 main French dairy cattle breeds in the region containing *DGAT1*. Thus, this region harbors one or several genes affecting this trait. In the Holstein breed, the K232A polymorphism appeared to contribute to a large part of the genetic variation due to the QTL. All sires heterozygous at the K232A polymorphism were also heterozygous at the QTL. The frequency of the K allele estimated in our Holstein population is in the range reported in other studies: 0.40 in New Zealand (Spelman et al., 2002), 0.35 (Winter et al., 2002) or 0.55 (Thaller et al., 2003) in Germany, 0.16 in Israel (Weller et al., 2003), and 0.27 in Brazil (Lacorte et al., 2006).

The effects of the K232A polymorphism were estimated on an EBV scale but with twice the DYD as records (which were not regressed as EBV). For the Holstein breed, the results were close to those previously published (Grisart et al., 2002; Spelman et al., 2002; Thaller et al., 2003), which were computed on a PTA scale with DYD as records. The K allele improved fat yield and fat and protein contents and had a negative effect on the other dairy traits. The effects in the 3 breeds were consistent with respect to their direction but the magnitude of the estimated effects changed. Similar tendencies have been described in Fleckvieh and Braunvieh cattle and Ayrshire and Jersey breeds, respectively (Spelman et al., 2002; Thaller et al., 2003). However, because the K allele is rare in the Montbéliarde and Normande breeds, the standard error of the estimators of these effects is large in these breeds.

The combination of allele frequency and allele substitution effects gives an indication on genetic variance due to this polymorphism. As expected, this variance was very large for fat content in the Holstein breed (more than 40% of the genetic variation). The proportion of

Table 6. Estimates of the average substitution effects on 5 production traits in the 3 different breeds from the regression of individual daughter yield deviations on the number of K alleles inherited in the 3 breeds

Trait	Montbéliarde	Normande	Holstein
Milk yield			
Average substitution effect, kg (SE)	−611 (170)	−142 (61)	−351 (22)
Proportion of genetic variance explained,[1] %	NA[2]	NA	13.1
Fat yield			
Average substitution effect, kg (SE)	6.24 (6.5)	9.30 (2.5)	14.9 (0.88)
Proportion of genetic variance explained,[1] %	NA	NA	14.8
Protein yield			
Average substitution effect, kg (SE)	−14.3 (5.3)	−0.610 (2.0)	−4.63 (0.66)
Proportion of genetic variance explained,[1] %	NA	NA	2.40
Fat content			
Average substitution effect, ‰ (SE)	4.25 (0.51)	2.54 (0.26)	3.43 (0.090)
Proportion of genetic variance explained,[1] %	NA	NA	42.5
Protein content			
Average substitution effect, ‰ (SE)	0.960 (0.32)	0.690 (0.14)	0.760 (0.050)
Proportion of genetic variance explained,[1] %	NA	NA	13.0

[1]Estimates of the proportion of genetic variance due to the K232A polymorphism have only been computed in the Holstein breed (see text).

[2]NA = not analyzed.

explained genetic variance estimated in this study for the different traits was close to that originally reported (Grisart et al., 2002). In addition, using a variance component approach on a larger sample of the French Holstein population, the proportion of genetic variance explained by the QTL traced by closely linked markers has recently been found to be very similar (Druet et al., 2006). Nevertheless, because of a lower frequency of the K allele in the Israeli Holstein population, lower values have been observed (Weller et al., 2003). Interestingly, the estimated variance for protein yield was relatively small, indicating that the use of *DGAT1* to improve this trait (the most important in the French breeding goal) is less efficient.

Although there is no doubt that the K232A polymorphism had a strong effect, at least in the Holstein breed, it seemed not solely responsible for the QTL as previously postulated (Bennewitz et al., 2004; Kuhn et al., 2004). Indeed, 2 Holstein sires appeared to be heterozygous at the QTL but were not heterozygous at the K232A polymorphism. A similar trend was observed in the Normande breed. Moreover, in the Montbéliarde breed, the K232A polymorphism was clearly not the underlying QTN. These results confirm that another mutation must

Table 7. Linkage disequilibrium of marker pairs within the different breeds

Marker pair	Breed	df	Sample size[1]	Haplotype number[2]	D′	r[2]	χ[2]	P
VNTR/K232A	Holstein	4	1,767	8	0.98	0.269	752	2×10^{-161}
	Montbéliarde	3	379	5	1.0	0.042	23.1	4×10^{-05}
	Normande	4	420	9	0.68	0.049	49.3	5×10^{-10}
VNTR/ILSTS039	Holstein	16	838	23	0.49	0.065	570	5×10^{-111}
	Montbéliarde	15	245	17	0.38	0.039	66.9	2×10^{-08}
	Normande	16	262	18	0.61	0.122	170	7×10^{-28}
VNTR/CSSM066	Holstein	20	1,783	30	0.42	0.035	1,057	4×10^{-211}
	Montbéliarde	15	373	20	0.30	0.016	75.0	6×10^{-10}
	Normande	20	417	24	0.36	0.026	164	2×10^{-24}
K232A/ILSTS039	Holstein	4	821	10	0.70	0.174	269	6×10^{-57}
	Montbéliarde	5	251	7	1.0	0.020	29.5	2×10^{-05}
	Normande	4	253	9	0.74	0.067	48.7	7×10^{-10}
K232A/CSSM066	Holstein	5	1,814	12	0.45	0.075	337	1×10^{-70}
	Montbéliarde	5	379	8	0.93	0.041	70.2	9×10^{-14}
	Normande	5	408	12	0.24	0.008	25.5	1×10^{-04}
ILSTS039/CSSM066	Holstein	20	846	29	0.55	0.096	585	4×10^{-111}
	Montbéliarde	25	245	26	0.40	0.027	218	9×10^{-33}
	Normande	20	250	22	0.31	0.016	119	5×10^{-16}

[1]The number of reconstructed haplotypes containing information for both markers in the considered population.

[2]The number of different haplotypes exiting in the considered population.

Table 8. Population allele frequencies of the variable number of tandem repeat (VNTR)/K232A haplotypes

VNTR/K232A haplotype	Population frequency			
	German Holstein[1]	Holstein	Normande	Montbéliarde
3R_A	0.021	0.041	0.249	0.000
4R_A	0.176	0.160	0.307	0.456
4R_K	0.006	0.002	0.101	0.000
5R_A	0.167	0.203	0.118	0.393
5R_K	0.207	0.350	0.025	0.045
6R_A	0.218	0.128	0.198	0.095
6R_K	0.000	0.001	0.002	0.000
7R_A	0.207	0.115	0.000	0.011

[1]From Kuhn et al. (2004).

contribute to the observed QTL effect. Recently, a VNTR in the 5′ noncoding region of *DGAT1* has been proposed to explain the additional part of the QTL variance, and one allele (7R in our study) was suggested as a putative causal allele (Kuhn et al., 2004; Sanders et al., 2006). Our study showed little evidence for this hypothesis. Indeed, among the 6 Holstein sires homozygous at the K232A polymorphism and carrying 1 copy of the 7R allele, only one was heterozygous at the QTL. In the Normande and Montbéliarde breeds, this allele was only present at a low frequency (<0.005) and its effect cannot be estimated. Recently, a functional analysis (Furbass et al., 2006) established that the Sp1 transcription factor could bind to the VNTR and demonstrated experimentally its functional relevance regarding induction of the transduction of a reporter gene. Nevertheless, no significant difference in the stimulating effect of the 3 different alleles compared (3R, 6R, and 7R) was observed. These different elements do not support a direct effect of the VNTR. For each breed considered, a different allele (3R in Normande breed, 4R in Montbéliarde breed, and 7R in Holstein breed) was found to be associated with a significant effect after correction for the K232A effect. This association of some VNTR alleles may be the result of residual LD existing with at least one additional causal polymorphism other than the K232A polymorphism underlying the QTL. Yet, a comprehensive screen for mutations in the *DGAT1* gene (in particular in the 5′ noncoding region) did not identify any significant association with the trait for any of the polymorphisms found except the VNTR (Winter et al., 2002; Kuhn et al., 2004); therefore, the additional causal variants might have to be searched in the surrounding intergenic regions.

The mutational event creating the K232A polymorphism from the K allele (the ancestral state) probably occurred early in the history of domestication of European cattle (Winter et al., 2002) more than 10,000 yr ago. In our study, the most important allelic diversity at the ILSTS039 microsatellite and the VNTR markers

was associated with the A allele at the K232A polymorphism. The high mutation rate at the VNTR (0.8% per generation and paternal gamete from our study) and microsatellite loci (known to range from 10^{-3} to 10^{-4}) can be considered as the primary factor generating allelic diversity in the haplotypes carrying the A allele, given that the expected numbers of alleles observed at these 2 markers was in almost perfect agreement with a stepwise mutation model (Ohta and Kimura, 1973).

Additionally, the frequency of the ancestral K allele at the K232A polymorphism has increased only recently, particularly in the Holstein breed, possibly due to more pronounced selection favoring fat content. This might have contributed to markedly reduced allelic diversity at surrounding markers by a genetic hitchhiking phenomenon (Grisart et al., 2004). Alternatively, local and recent introgression of the K allele in some European breeds might explain the low level of within-breed diversity associated with the K-carrying haplotypes (Kaupe et al., 2004). Indeed, some ancestral K alleles, probably from different origins, still segregate in the different cattle populations. In our study for instance, within the Normande and Holstein populations, the haplotypes carrying the K allele were mainly associated with different alleles at both the ILSTS039 and VNTR markers. On the other hand, the few K alleles segregating in the Montbéliarde breed might have been introduced recently during the known limited crossbreeding events with Red Holstein, which occurred in the 1970s. Indeed, in the Montbéliarde breed, the K alleles were all associated with the same ILSTS039 and VNTR alleles (218 and 5R, respectively) whereas in the Holstein breed, 85 and 99% of the K alleles were associated with the same 2 alleles.

CONCLUSIONS

The ancestral K allele of the *DGAT1* K232A polymorphism was present in the 3 breeds studied, whereas the alternate A allele is nearly fixed in the Montbéliarde breed. In the Normande and French Holstein breeds, the effects of the allele were consistent with a large effect on milk production and composition traits. The VNTR polymorphism explained only a small fraction of the variance of the QTL (after correction for the K232A effect). Other uncharacterized polymorphism(s) certainly exist at the centromeric end of bovine chromosome 14 because it has been shown that the K232A polymorphism is not solely responsible for the QTL variation. In the Montbéliarde breed, all the heterozygous sires at the QTL were homozygous at the K232A polymorphism. This breed represents an excellent model to identify additional mutations contributing to the observed QTL effect.

REFERENCES

Bennewitz, J., N. Reinsch, S. Paul, C. Looft, B. Kaupe, C. Weimann, G. Erhardt, G. Thaller, C. Kuhn, M. Schwerin, H. Thomsen, F. Reinhardt, R. Reents, and E. Kalm. 2004. The DGAT1 K232A mutation is not solely responsible for the milk production quantitative trait locus on the bovine chromosome 14. J. Dairy Sci. 87:431–442.

Boichard, D., S. Fritz, M. N. Rossignol, M. Y. Boscher, A. Malafosse, and J. J. Colleau. 2002. Implementation of marker-assisted selection in French dairy cattle. Commun. no. 22-03 in Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France.

Boichard, D., C. Grohs, F. Bourgeois, F. Cerqueira, R. Faugeras, A. Neau, R. Rupp, Y. Amigues, M. Y. Boscher, and H. Leveziel. 2003. Detection of genes influencing economic traits in three French dairy cattle breeds. Genet. Sel. Evol. 35:77–101.

Bois, P., and A. J. Jeffreys. 1999. Minisatellite instability and germline mutation. Cell. Mol. Life Sci. 55:1636–1648.

Coppieters, W., J. Riquet, J. J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar, and M. Georges. 1998. A QTL with major effect on milk yield and composition maps to bovine chromosome 14. Mamm. Genome 9:540–544.

Druet, T., S. Fritz, D. Boichard, and J. J. Colleau. 2006. Estimation of genetic parameters for quantitative trait loci for dairy traits in the French Holstein population. J. Dairy Sci. 89:4070–4076.

Furbass, R., A. Winter, R. Fries, and C. Kuhn. 2006. Alleles of the bovine DGAT1 variable number of tandem repeat associated with a milk fat QTL at chromosome 14 can stimulate gene expression. Physiol. Genomics 25:116–120.

Gilmour, A. R., B. R. Cullis, S. J. Welham, and R. Thompson. 2000. ASREML reference manual. http://www.ciser.cornell.edu/FAQ/ASReml_user_guide.pdf

Green, E., K. Falls, and S. Crooks. 1990. Documentation for CRI-MAP, version 2.4. http://linkage.rockefeller.edu/soft/crimap

Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford, P. Berzi, N. Cambisano, M. Mni, S. Reid, P. Simon, R. Spelman, M. Georges, and R. Snell. 2002. Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Res. 12:222–231.

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J. J. Kim, A. Kvasz, M. Mni, P. Simon, J. M. Frere, W. Coppieters, and M. Georges. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Natl. Acad. Sci. USA 101:2398–2403.

Haley, C. S., and S. A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324.

Hedrick, P. W. 1987. Gametic disequilibrium measures: Proceed with caution. Genetics 117:331–341.

Heyen, D. W., J. I. Weller, M. Ron, M. Band, J. E. Beever, E. Feldmesser, Y. Da, G. R. Wiggans, P. M. VanRaden, and H. A. Lewin. 1999. A genome scan for QTL influencing milk production and health traits in dairy cattle. Physiol. Genomics 1:165–175.

Hill, W. G. 1975. Linkage disequilibrium among multiple neutral alleles produced by mutation in finite population. Theor. Popul. Biol. 8:117–126.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231.

Kaupe, B., A. Winter, R. Fries, and G. Erhardt. 2004. DGAT1 polymorphism in *Bos indicus* and *Bos taurus* cattle breeds. J. Dairy Res. 71:182–187.

Kuhn, C., G. Thaller, A. Winter, O. R. Bininda-Emonds, B. Kaupe, G. Erhardt, J. Bennewitz, M. Schwerin, and R. Fries. 2004. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics 167:1873–1881.

Lacorte, G. A., M. A. Machado, M. L. Martinez, A. L. Campos, R. P. Maciel, R. S. Verneque, R. L. Teodoro, M. G. C. D. Peixoto, M. R. S. Carvalho, and C. G. Fonseca. 2006. DGAT1 K232A polymorphism in Brazilian cattle breeds. Genet. Mol. Res. 5:475–482.

Looft, C., N. Reinsch, C. Karall-Albrecht, S. Paul, M. Brink, H. Thomsen, G. Brockmann, C. Kuhn, M. Schwerin, and E. Kalm. 2001. A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine Chromosome 14. Mamm. Genome 12:646–650.

Ohta, T., and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet. Res. 22:201–204.

Sanders, K., J. Bennewitz, N. Reinsch, G. Thaller, E. M. Prinzenberg, C. Kuhn, and E. Kalm. 2006. Characterization of the DGAT1 mutations and the CSN1S1 promoter in the German Angeln dairy cattle population. J. Dairy Sci. 89:3164–3174.

Seaton, G., C. S. Haley, S. A. Knott, M. Kearsey, and P. M. Visscher. 2002. QTL Express: Mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18:339–340.

Spelman, R. J., C. A. Ford, P. McElhinney, G. C. Gregory, and R. G. Snell. 2002. Characterization of the DGAT1 gene in the New Zealand dairy population. J. Dairy Sci. 85:3514–3517.

Thaller, G., W. Kramer, A. Winter, B. Kaupe, G. Erhardt, and R. Fries. 2003. Effects of DGAT1 variants on milk production traits in German cattle breeds. J. Anim. Sci. 81:1911–1918.

VanRaden, P. M., and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737–2746.

Weller, J. I., M. Golik, E. Seroussi, E. Ezra, and M. Ron. 2003. Population-wide analysis of a QTL affecting milk-fat production in the Israeli Holstein population. J. Dairy Sci. 86:2219–2227.

Weller, J. I., Y. Kashi, and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73:2525–2537.

Winter, A., W. Kramer, F. A. Werner, S. Kollers, S. Kata, G. Durstewitz, J. Buitkamp, J. E. Womack, G. Thaller, and R. Fries. 2002. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proc. Natl. Acad. Sci. USA 99:9300–9305.