Identification and Dissection of Four Major QTL Affecting Milk Fat Content in the German Holstein-Friesian Population

Xiaolong Wang[1]*[9], Christine Wurmser[1][9], Hubert Pausch[1], Simone Jung[1], Friedrich Reinhardt[2], Jens Tetens[3], Georg Thaller[3], Ruedi Fries[1]

1 Chair of Animal Breeding, Technische Universitaet Muenchen, Freising-Weihenstephan, Germany, 2 Vereinigte Informationssysteme Tierhaltung w.V., Verden/Aller, Germany, 3 Institute of Animal Breeding and Husbandry, Christian-Albrechts-Universitaet, Kiel, Germany

Abstract

Milk composition traits exhibit a complex genetic architecture with a small number of major quantitative trait loci (QTL) explaining a large fraction of the genetic variation and numerous QTL with minor effects. In order to identify QTL for milk fat percentage (FP) in the German Holstein-Friesian (HF) population, a genome-wide association study (GWAS) was performed. The study population consisted of 2327 progeny-tested bulls. Genotypes were available for 44,280 SNPs. Phenotypes in the form of estimated breeding values (EBVs) for FP were used as highly heritable traits. A variance components-based approach was used to account for population stratification. The GWAS identified four major QTL regions explaining 46.18% of the FP EBV variance. Besides two previously known FP QTL on BTA14 ($P = 8.91 \times 10^{-198}$) and BTA20 ($P = 7.03 \times 10^{-12}$) within *DGAT1* and *GHR*, respectively, we uncovered two additional QTL regions on BTA5 ($P = 2.00 \times 10^{-13}$) and BTA27 ($P = 9.83 \times 10^{-5}$) encompassing *EPS8* and *GPAT4*, respectively. *EPS8* and *GPAT4* are involved in lipid metabolism in mammals. Re-sequencing of *EPS8* and *GPAT4* revealed 50 polymorphisms. Genotypes for five of them were inferred for the entire study population. Two polymorphisms affecting potential transcription factor binding sites of *EPS8* ($P = 1.40 \times 10^{-12}$) and *GPAT4* ($P = 5.18 \times 10^{-5}$), respectively, were highly significantly associated with the FP EBV. Our results provide evidence that alteration of regulatory sites is an important aspect of genetic variation of complex traits in cattle.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: xiaolong.wang@tierzucht.tum.de

[9] These authors contributed equally to this work.

Introduction

Improvement of milk yield and composition is a major objective of dairy cattle breeding programs and highly reliable breeding values are estimated to this end. Milk composition traits such as protein and fat content are not only important production traits but also permit insights into the metabolic constitution of lactating cows [1].

Milk fat content, indicated as fat percentage (FP), is a prototypical complex quantitative trait determined by numerous loci with small effects and only few loci with major effects [2]. Family-based linkage studies and genome-wide association studies (GWAS) have already identified several genomic regions contributing to the genetic variation of FP in cattle (*e.g.* [2,3,4,5]). Among them, most prominently, a K232A-substitution within the acyl-CoA:diacylglycerol acyltransferase encoding gene *DGAT1* [6,7] and a F279Y-substitution within the growth hormone receptor encoding gene *GHR* [8,9] have been well characterized. These two polymorphisms account for a major fraction of the genetic variation of FP in various cattle breeds (*e.g.* [10,11,12]).

The aim of the present study was to identify major quantitative trait loci (QTL) for FP in the German HF population by genome-wide association analysis and to pinpoint the causal variants. The study revealed four QTL, two of them novel. We report candidate genes and putative causal variants for the newly identified QTL.

Results

Genotypes and Quality Control

The HF bulls were genotyped with the Illumina Bovine SNP 50 K BeadChip® comprising 54,001 single nucleotide polymorphisms (SNPs). Of 2401 genotyped animals, 62 were removed from the data set because genotyping failed for more than 10% of the SNPs. Twelve animals showed major differences between pedigree and genomic relationship and were excluded from further analysis (Figure S1). 549 SNPs with unknown chromosomal position and 7951 SNPs with minor allele frequencies <0.01 were omitted from subsequent analyses. 732 SNPs that were missing in more than 10% of the animals and 966 SNPs that deviated significantly from the Hardy-Weinberg equilibrium ($P < 0.001$) were excluded. The final data set comprised 2327 animals and 44,280 SNPs.
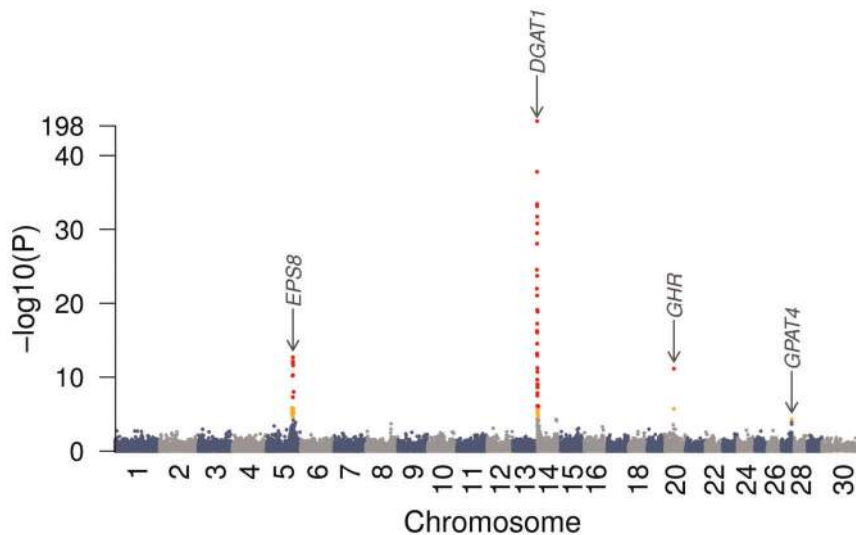
Figure 1. Association of 44,280 SNPs with the estimated breeding values for milk fat percentage in 2327 animals of the German Holstein-Friesian population. Red symbols represent SNPs with $P<1.3\times10^{-6}$ (Bonferroni-corrected significance level), orange symbols indicate chromosome-wide significance.
doi:10.1371/journal.pone.0040711.g001

Association study

The genome-wide association study based on 44,280 SNPs and 2327 progeny-tested bulls identified four QTL for FP on BTA5, BTA14, BTA20 and BTA27 in the German HF population (**Figure 1**).

Forty-eight significantly associated SNPs encompass the DGAT1 encoding gene (1.46 Mb –7.31 Mb) and define the QTL region on BTA14. The most significantly associated SNP ($P=1.57\times10^{-198}$) resides 1149 bp upstream of the postulated DGAT1-QTN [6,7]. Genotyping of the K232A-substitution within DGAT1 showed close to complete linkage disequilibrium with the SNP ARS-BFGL-NGS-4939 ($r^2=0.998$). The P-value for the latter SNP is marginally lower than for the K232A marker due to the imputed genotypes of two animals, probably indicating inaccurate imputation rather than incomplete linkage disequilibrium (**Figure 2A**).

On BTA20, a single SNP (BTA-84181-no-rs) located 357 kb downstream of the postulated causal F279Y-substitution within the GHR encoding gene [9] was significantly associated ($P=1.95\times10^{-6}$). The association signal for the F279Y-substitution was more prominent than for any other SNP on BTA20 ($P=7.03\times10^{-12}$) (**Figure 2B**). BTA-84181-no-rs SNP is only moderately linked with the F279Y-substitution ($r^2=0.2$). We additionally obtained genotypes for the S18N-substitution within the PRLR encoding gene [8], another presumed causal variant. The S18N-substitution is 7.56 Mb distant from the significantly associated SNP BTA-84181-no-rs. No association of the PRLR-variant ($P=0.319$) with the FP EBV was observed.

Seven genome-wide significantly associated SNPs located between 91.2 Mb and 97.1 Mb delineate the QTL region on BTA5. The most significantly associated SNP Hapmap49734-BTA-74577 ($P=2.00\times10^{-13}$) is located in the second intron of the epidermal growth factor receptor pathway substrate 8 encoding gene (EPS8) (**Table 1**) (**Figure 3**).

On BTA27, SNP ARS-BFGL-NGS-57448 is associated with $P=9.83\times10^{-5}$ and thus with chromosome-wide significance only. It is located about 56 kb upstream of the glycerol-3-phosphate acyltransferase 4 encoding gene (GPAT4) (**Table1**) (**Figure 4**).

Assessing the Impact of the Four QTL

Alleles increasing the FP EBVs were identified for the most significantly associated SNP for each of the four QTL regions. The frequency distribution of animals with an increasing number of alleles is displayed in **Figure 5**. The EBVs of animals with one and seven FP increasing alleles differ by more than three standard deviations.

Chromosomal Partitioning of the EBV Variance

The proportion of the EBV variance attributed to a particular chromosome/QTL was estimated with the effects of all chromosomes/QTL fitted simultaneously. Totally, the 44,280 SNPs account for 85.97% of the EBV variance. The contribution of particular chromosomes varies strongly (**Figure 6**). A major fraction of the EBV variance is attributable to BTA14 (33.60%), BTA5 (12.08%) and BTA20 (7.01%). BTA27 accounts for a minor fraction (1.19%) of the EBV variance only. Totally, the four identified QTL explain 46.18% of the FP EBV variance. The estimates of the EBV variance attributable to the four QTL are 8.35% (BTA5), 31.04% (BTA14), 5.91% (BTA20) and 0.88% (BTA27).

Molecular-genetic Analysis of the BTA5 QTL

EPS8 was considered as a positional candidate gene for the FP QTL region on BTA5 as the two most significantly associated SNPs were located within its second intron. EPS8 plays a role in the fat metabolism of mammals as it is a substrate for the EGFR-kinase [13]. The EGFR-kinase activates transcriptional regulators of fatty acid synthesis and thereby increases the concentration of intracellular fatty acids [14]. Re-sequencing of EPS8 in 24 animals resulted in the identification of 20 polymorphisms (**Table S4**). We genotyped two promoter polymorphisms (ss319604831 and ss319604833) in high linkage disequilibrium ($r^2>0.8$) with SNP Hapmap49734-BTA-74577 and for one non-synonymous mutation (M599T, ss319604845) located in a highly conserved region of EPS8 (**Table 1, Figure 3**). However, only ss319604833, located 100 bp upstream of the transcription start of EPS8, was highly associated with the FP EBV ($P=1.40\times10^{-12}$) (**Table 1**) (**Figure 3**). The significance level of the putative causal variant
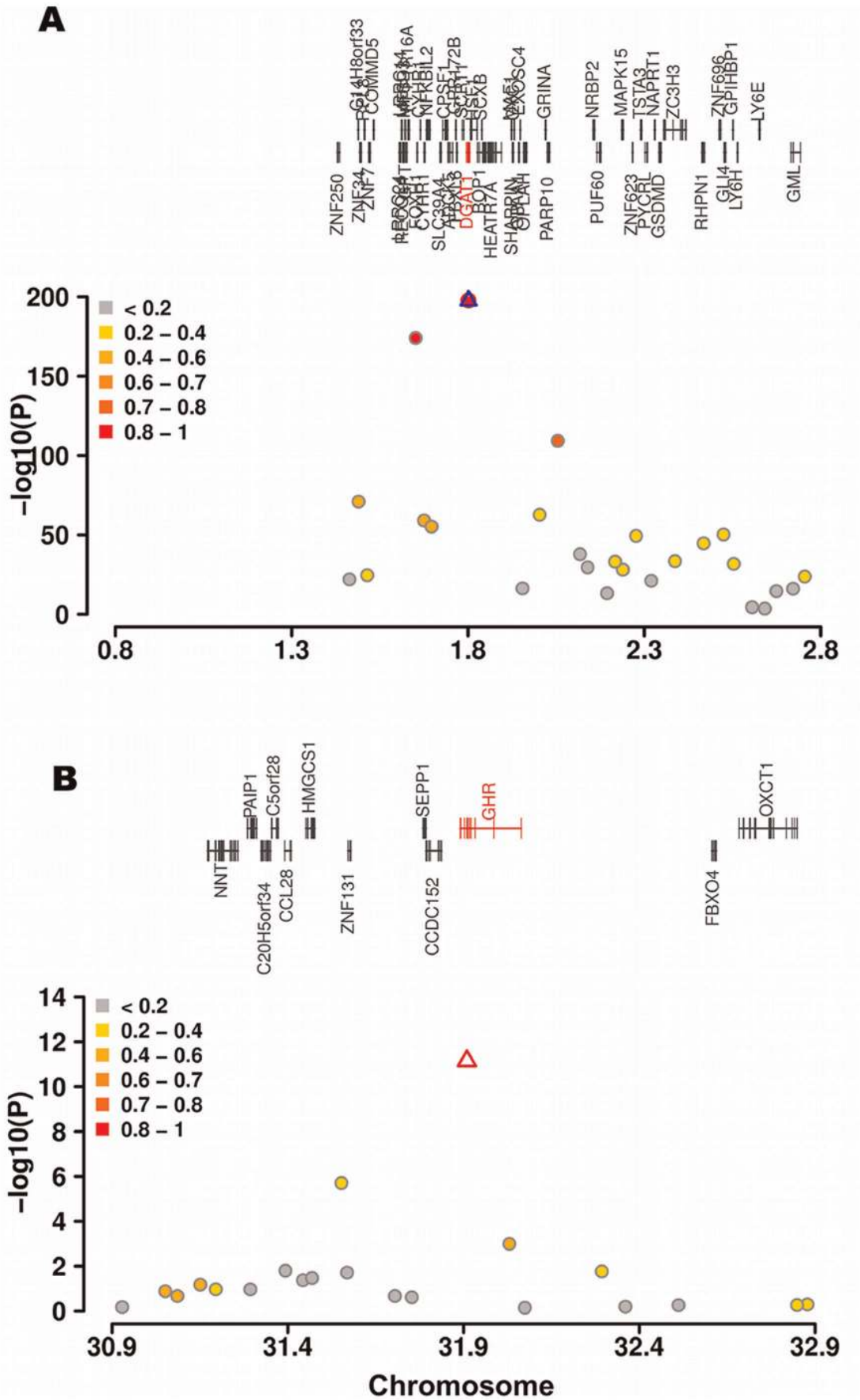
Figure 2. Detailed view of two genomic regions within known QTL for milk fat content in cattle. QTL regions on BTA14 (**A**) and BTA20 (**B**) encompassing *DGAT1* and *GHR*, respectively. Open symbols (blue and red triangles) represent the K232A- and the F279Y-substitution, respectively. Different colors indicate the extent of linkage disequilibrium (r^2) between the postulated QTN and all other SNPs. doi:10.1371/journal.pone.0040711.g002

ss319604833 was marginally lower compared to the SNP Hapmap49734-BTA-74577 ($P = 2.00 \times 10^{-13}$). This could be due to imperfect imputation, since 9.2% of the genotypes were missing after genotyping and thus were imputed.

ss319604833 is located in a region of *cis*-acting regulatory elements. Five transcription factors whose binding may be affected by the polymorphism were predicted. The promoter variant with the more frequent C allele, associated with a lower FP EBV, contains a potential MEF3-element. In contrast, TBF1-, Ptx1-, MafB- and TFAP2A-elements were predicted for the sequence with the rare T allele, associated with an increased FP EBV (**Figure 7A**). The expression of transcription factor TFAP2A is significantly correlated with the concentration of nonesterified fatty acids (NEFA) and liver triacylglycerol [15].

Molecular-genetic Analysis of the BTA27 QTL

The FP QTL region on BTA27 encompasses the *GPAT4* gene. *GPAT4* encodes the rate-limiting enzyme glycerol-3-phosphate acyltransferase in the triacylglycerol biosynthesis pathway and plays a key role in milk fat biosynthesis [16]. Re-sequencing of the exons and regulatory flanking regions in a panel of twelve animals revealed 30 polymorphisms (**Table S4**). Genotypes for two polymorphisms (ss410759404 and ss410758894) located within the 5′flanking region and within the 3′UTR, respectively, were obtained for the entire study population (**Table 1**) (**Figure 4**). ss410759404 (GA>T), was more strongly associated with the FP EBV ($P = 5.18 \times 10^{-5}$) than any other SNP on BTA27.

ss410759404 is located 1378 bp upstream of the translation start of *GPAT4*. Transcription factors SREB, CREB, RXR-α and RAR-β were predicted for the sequence with the frequent T allele, which is associated with a lower FP EBV. The PAX6-element was predicted for the sequence with the minor GA allele, that was associated with an increased FP EBV. (**Figure 7B**). RXR-α up-regulates genes involved in fatty acid and lipid metabolism during the process of adipogenesis [17], CREB takes part in the

regulation of gluconeogenesis [18] and SREB is considered to be one of the central regulation factors in milk fat synthesis [16].

The prediction of putative miRNA binding sites revealed that the binding of mmu-miR-712 might be affected by the ss410758894 polymorphism, with the minor A allele showing less similarity to the miRNA-binding site (**Figure 7C**). However, no conserved sites for miRNA families could be identified due to low conservation between species. Also, the *P*-value for ss410758894 ($P = 2.27 \times 10^{-4}$) did not meet the significance level.

Discussion

Our genome-wide association study was based on a medium-sized sample of the German HF population and on a dense SNP map. It revealed four major QTL for FP. The four identified QTL regions account for a large part of the EBV variance (46.18%). However, a large fraction of the EBV variance is attributable to chromosomes with no identified QTL, *e.g.* BTA11 and BTA19. The QTL region on BTA14 encompassing *DGAT1* accounts for 31.04% of the EBV variance which agrees with previous findings [10]. The QTL region on BTA20 encompassing the F279Y-substitution within the GHR encoding gene on BTA20 [9] accounts for 5.91% of the EBV variance. There was no evidence for a second FP QTL on BTA20 resulting from the S18N-substitution within the *PRLR* encoding gene [8]. We were able to identify two additional QTL regions for FP in the German HF population on BTA5 and BTA27 that together account for 9.23% of the EBV variance. Our findings support the proposed genetic architecture of FP with numerous loci with small effects and only few loci with larger effects [2].

The presence of a FP QTL on BTA5 agrees with findings in the Australian HF population [2]. We identified a highly significantly associated polymorphism in the promoter region of *EPS8* which is supposed to mediate the binding of TFAP2A and concomitantly the transcription rate of *EPS8*. EPS8 physically interacts with the

Table 1. Characteristics of the most significantly associated 50 K Illumina BeadChip SNPs and additional polymorphisms of four major QTL for milk fat percentage in the German Holstein-Friesian population.

SNP	BTA	Physical position [a]	Minor allele (MAF)	P [b]	Neighboring gene
Hapmap49734-BTA-74577	5	94,570,828	A (0.09)	2.00×10^{-13}	EPS8
ARS-BFGL-NGS-4939	14	1,801,116	G (0.31)	1.57×10^{-198}	DGAT1
BTA-84181-no-rs	20	31,552,475	G (0.37)	1.95×10^{-6}	GHR
ARS-BFGL-NGS-57448	27	36,155,097	A (0.36)	9.83×10^{-5}	GPAT4
ss319604831	5	94,551,792	G (0.21)	4.92×10^{-6}	EPS8
ss319604833	5	94,553,580	T (0.09)	1.40×10^{-12}	EPS8
ss319604845	5	94,726,848	T (0.22)	2.40×10^{-5}	EPS8
K232A	14	1,802,265	A (0.31)	8.91×10^{-198}	DGAT1
F279Y	20	31,909,479	A (0.17)	7.03×10^{-12}	GHR
S18N	20	39,115,344	G (0.14)	3.19×10^{-1}	PRLR
ss410759404	27	36,211,257	GA (0.38)	5.18×10^{-5}	GPAT4
ss410758894	27	36,228,939	A (0.40)	2.27×10^{-4}	GPAT4

[a]The SNPs are ordered according to their position on the UMD3.1 assembly of the bovine genome sequence.
[b]The *P*-values are obtained after regression analysis and by using a variance components based approach to account for population stratification.
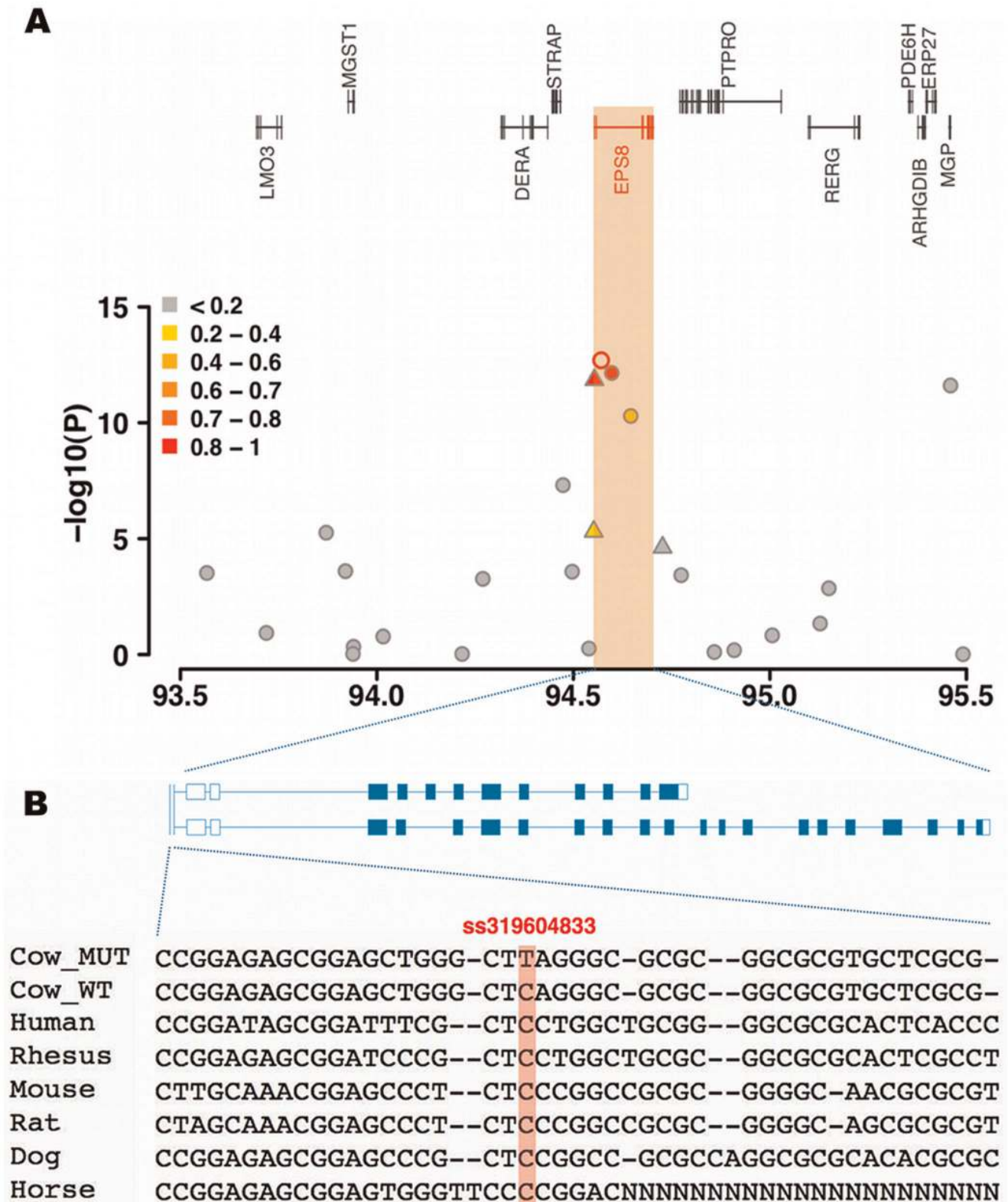doi:10.1371/journal.pone.0040711.t001

Figure 3. Schematic view of the BTA5 QTL region encompassing *EPS8*. (**A**) The open symbol represents the most significantly associated SNP. Different colors indicate the extent of linkage disequilibrium (r²) between the most significantly associated SNP and all other SNPs. Triangles and circles indicate SNPs resulting from re-sequencing and genotyping with the 50 K BeadChip, respectively. (**B**) Two alternative transcripts of *EPS8* are present in cattle. The multispecies sequence alignment of a segment in the promoter region of *EPS8* encompassing the highly significantly associated ss319604833 polymorphism (red background) illustrates the high conservation among species.
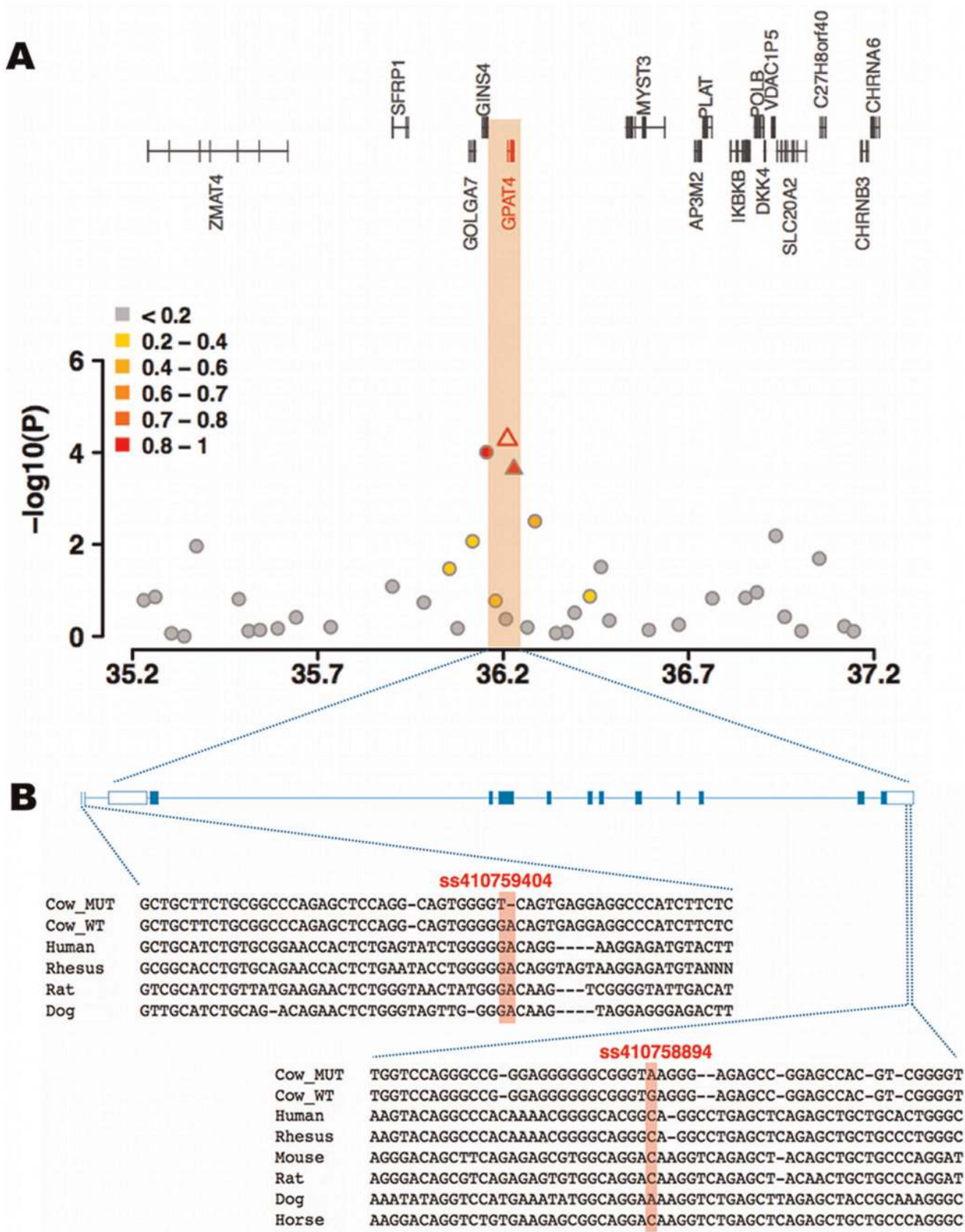doi:10.1371/journal.pone.0040711.g003

Figure 4. Schematic view of the BTA27 QTL region encompassing *GPAT4*. (**A**) The open symbol represents the most significantly associated SNP. Different colors indicate the extent of linkage disequilibrium (r^2) between the most significantly associated SNP and all other SNPs. The triangles indicate SNPs resulting from re-sequencing, circles indicate SNPs from genotyping with the 50 K BeadChip. (**B**) Gene structure of *GPAT4* and

epidermal growth factor receptor [13]. Recently it has been shown that sterol regulatory element-binding proteins (SREBPs) are regulated by the epidermal growth factor [19]. SREBPs control the expression of genes required for the uptake and synthesis of cholesterol, fatty acid and triglycerides. Thus, it seems likely that an enhanced transcription rate of *EPS8*, conferred by binding of TFAP2A, results in an increased milk fat biosynthesis in the lactating mammary gland. Therefore, a contribution of ss319604833 to the genetic variation of milk fat synthesis seems plausible.

Recently, Bouwman and colleagues [20] reported a QTL contributing to the genetic variation of milk fatty acid composition in the Dutch HF population nearby *GPAT4*, supporting our findings of a FP QTL in the German HF population on BTA27. *GPAT4* plays a crucial role in lipid biosynthesis in mammals [16]. The transcription rate of *GPAT4* is highly correlated with the concentration of milk diacylglycerols and triacylglycerols [21,22]. Prediction of transcription factor binding sites for the highly significantly associated SNP ss410759404 suggests that the binding of transcription factors involved in fat metabolism might be affected by this variant [16,17,18]. Hence, a contribution of ss410759404 to the milk fat biosynthesis capacity in lactating cows seems likely. Association analysis uncovered a second SNP (ss410758894) affecting a potential miRNA binding site of *GPAT4*. Although ss410758894 did not meet the criteria for significant association, it is possible that both variants contribute to the

genetic variation of the BTA27 FP QTL in the German HF population.

Materials and Methods

Animals and Phenotypes

The study population consisted of 2401 progeny-tested Holstein Friesian bulls. The animals descend from 376 different sires and 423 maternal grand-sires. The paternal half-sib families and maternal grandsire families encompass up to 83 members with an average of six members. Phenotypes in the form of estimated breeding values (EBVs) for milk FP were obtained from vit w.V. Verden (www.vit.de, April 2010 version). Breeding value estimation for FP was carried out using best linear unbiased prediction (BLUP).

Genome-wide Association Study

The HF bulls were genotyped with the Illumina Bovine SNP 50 K BeadChip® comprising 54,001 single nucleotide polymorphisms. The chromosomal positions of 53,452 SNPs were according to the University of Maryland UMD3.1 assembly of the bovine genome sequence [23]. SNPs with unknown position or if genotyping failed in more than 10% of the animals were excluded. SNPs with a minor allele frequency (MAF) <0.01 or significant (P<0.001) deviation from the Hardy-Weinberg equilibrium were omitted for subsequent analysis. Sporadically missing genotypes
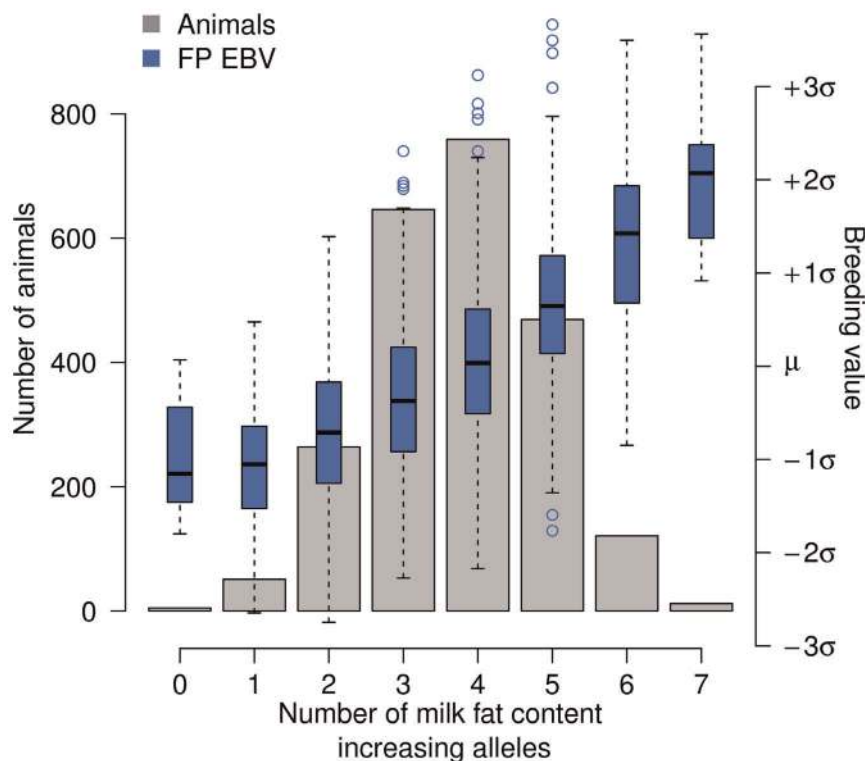

Figure 5. The combined impact of the four identified QTL on BTA5, 14, 20 and 27 on the estimated breeding value for milk fat percentage in the German Holstein-Friesian population. 2327 Holstein-Friesian animals are grouped according to the number of alleles that increase the milk FP EBV. The grey bars indicate the number of animals with an increasing number of FP increasing alleles. The box plots represent the FP EBVs for each group. No animals carrying eight FP increasing alleles were observed in the study population.
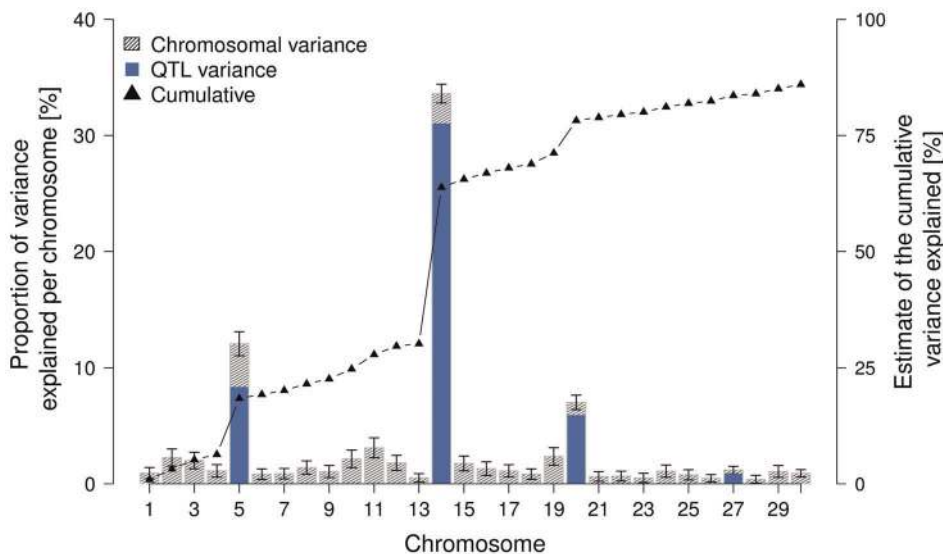doi:10.1371/journal.pone.0040711.g005

Figure 6. Partitioning of the genetic variance onto 30 chromosomes and four identified QTL regions on BTA5, 14, 20 and 27. The grey shaded bars indicate the fraction of EBV variance attributed to a particular chromosome and the corresponding standard error. The dark grey bars represent the fraction of EBV variance attributed to each of the four identified QTL regions. The black triangles represent the cumulative proportion of EBV variance explained.
doi:10.1371/journal.pone.0040711.g006

were imputed using default parameters of *Beagle* (version 3.2.1) [25]. Animals with more than 10% missing genotypes were not considered for further analyses. The genomic relationship of each pair of animals was obtained as proposed by VanRaden (2008) [24] and was compared with the corresponding pedigree relationship. Animals showing major inconsistencies between pedigree and genomic relationship were excluded from further analysis.

To account for population stratification, *EMMAX* [26] was used to fit the model $y = Xb + u + \varepsilon$, where y is a vector of EBVs for FP, b is the SNP effect, X is a design matrix of SNP genotypes, u is the random polygenic effect with $(0, \sigma_a^2 G)$, where σ_a^2 is the additive genetic variance, G is the genomic relationship matrix (GRM) among the 2327 animals (see above) and ε is the non heritable component of the random variation.

Chromosomal Partitioning of the Genetic Variance

In order to estimate the proportion of EBV variance attributable to a particular chromosome and QTL, a GRM was built (see above) for each of the 30 chromosomes and the four QTL regions separately. A QTL was defined by the SNPs within a 5 Mb interval centered on the most significantly associated SNP. Chromosome-specific variance was estimated based on GRMs that included all SNPs on a chromosome except those within the 5 Mb QTL interval. We used *GCTA* [27] to fit the model $y = \sum_{i=1}^{34} g_i + e$, where y is a vector of EBVs for FP, g is a vector of genetic effects attributed to the ith chromosome/QTL, and e is a vector of random residual deviates. g_i is assumed to be normally distributed with $N(0, G_i \sigma_{g_i}^2)$, where G_i is the GRM built based on SNPs of the ith chromosome/QTL. The proportion of variance attributable to the ith chromosome/QTL was calculated as $\sigma_i^2 / (\sum_{j=1}^{34} \sigma_j^2 + \sigma_e^2)$.

Candidate Gene Annotation

The *GenomeThreader* software tool [28] was used to predict the genomic structure and localization of the genes based on the University of Maryland UMD3.1 assembly of the bovine genome sequence [23] and the Dana-Farber Cancer Institute bovine gene index release 12.0 [29] together with the annotated RNA sequences of the UMD3.1 assembly. The *GenomeThreader* output was viewed and edited using the *Apollo Genome Annotation and Curation Tool* [30]. The exon-intron organization of the annotated genes is summarized in **Table S1**. Transcription factor binding sites were predicted with *MatInspector* [31], *JASPAR* [32] and *TESS* [33]. Prediction of microRNAs was carried out using *MicroCosm Targets* (http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/) and *TargetScanHuman 6.1* (http://www.targetscan.org/).

Re-sequencing of Candidate Genes

Candidate genes for the BTA5 and BTA27 FP QTL were re-sequenced in 24 and 12 HF animals, respectively. The animals were selected based on their genotypes for the most significantly associated SNP of the particular QTL. PCR primers were designed for the promoter regions (3000 bp and 1500 bp upstream of the transcription start site for *EPS8* and *GPAT4*, respectively), for all exons and intron-exon boundaries as well as for the 5′ and 3′ untranslated regions (UTR) (**Table S2**). The PCR products were sequenced using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) on the ABI 3130×l Genetic Analyzer (Applied Biosystems). The *Phred/Phrap/Polyphred* software suite [34] was used for base calling, sequence alignment and polymorphism identification, and *consed* [35] was used for viewing.

Genotyping of Selected Polymorphisms

Three previously proposed quantitative trait nucleotides (QTN) for FP, namely the ones responsible for the K232A-substitution within *DGAT1* [6,7], the F279Y-substitution within *GHR* [9] and the S18N-substitution within *PRLR* [8] were obtained by
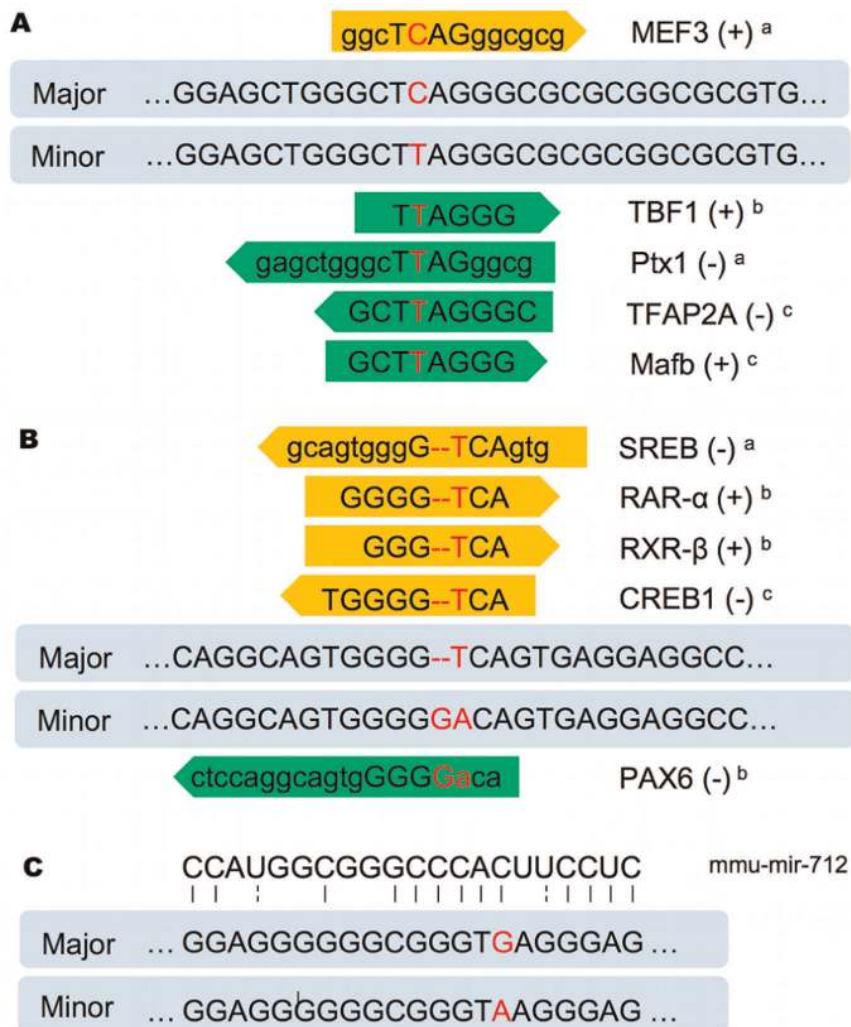
Figure 7. Prediction of regulatory sites for the polymorphisms in *EPS8* **and** *GPAT4***.** Prediction of transcription factor binding sites within the promoter sequences of *EPS8* and *GPAT4* encompassing the SNP ss319604833 (**A**) and the SNP ss410759404 (**B**). (+) and (−) indicate forward and reverse direction. *MatInspector*[a], *TESS*[b] and *JASPAR*[c] were used for the prediction of transcription factor binding sites. (**C**) Prediction of a miRNA binding site for mmu-mir-712 (free energy: −27.93 kcal/mol) within the 3′UTR of *GPAT4* encompassing the polymorphism ss410758894.
doi:10.1371/journal.pone.0040711.g007

TaqMan® genotyping analysis (Applied Biosystems Applera, Darmstadt, Germany). Potentially functional polymorphisms for the BTA5 and BTA27 QTL were genotyped in 2327 animals of our study population as well (**Table S3**). Sporadically missing genotypes were imputed using default parameters of *Beagle* (version 3.2.1) [25]. The missing genotype rates were 3.1% for both the *DGAT1* and the *GHR* variant, 9.2% for the *EPS8* polymorphism ss319604833, 8.3% for ss410759404 and 11.3% for ss410758894 within *GPAT4*.

Supporting Information

Figure S1 Comparison of the pairwise pedigree *vs.* genomic relationship. Pairwise pedigree *vs.* genomic relationship for the studied Holstein-Friesian population before (A) and after (B) the exclusion of 12 animals with inconsistencies.
(PDF)

Table S1 Exon/intron boundaries of bovine genes for re-sequencing.
(PDF)

Table S2 Primers used for re-sequencing bovine genes.
(PDF)

Table S3 Primers and probes used for Taqman Genotyping Assay.
(PDF)

Table S4 Polymorphisms identified in bovine genes.
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: XW CW HP RF. Performed the experiments: XW CW HP. Analyzed the data: HP XW CW. Contributed reagents/materials/analysis tools: SJ FR JT GT. Wrote the paper: XW CW HP RF.

References

1. Buttchereit N, Stamer E, Junge W, Thaller G (2010) Evaluation of five lactation curve models fitted for fat:protein ratio of milk and daily energy balance. Journal of Dairy Science 93: 1702–1712. doi:10.3168/jds.2009–2198

2. Hayes BJ, Pryce J, Chamberlain AJ, Bowman PJ, Goddard ME (2010) Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits. PLoS Genet 6: e1001139. doi:10.1371/journal.pgen.1001139

3. Schennink A, Bovenhuis H, Léon-Kloosterziel KM, van Arendonk JAM, Visker MHPW (2009) Effect of polymorphisms in the FASN, OLR1, PPARGC1A, PRL and STAT5A genes on bovine milk-fat composition. Anim. Genet. 40: 909–916. doi:10.1111/j.1365–2052.2009.01940.x

4. Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Everts-van der Wind A, et al. (2005) Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. 15: 936–944. doi:10.1101/gr.3806705

5. Jiang L, Liu J, Sun D, Ma P, Ding X, et al. (2010) Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. PLoS ONE 5: e13661. doi:10.1371/journal.pone.0013661

6. Grisart B, Coppieters W, Farnir F, Karim L, Ford C, et al. (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. doi:10.1101/gr.224202

7. Winter A, Krämer W, Werner FAO, Kollers S, Kata S, et al. (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. PNAS 99: 9300–9305. doi:10.1073/pnas.142293799

8. Viitala S, Szyda J, Blott S, Schulman N, Lidauer M, et al. (2006) The Role of the Bovine Growth Hormone Receptor and Prolactin Receptor Genes in Milk, Fat and Protein Production in Finnish Ayrshire Dairy Cattle. Genetics 173: 2151–2164. doi:10.1534/genetics.105.046730

9. Blott S, Kim J-J, Moisio S, Schmidt-Küntzel A, Cornet A, et al. (2003) Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics 163: 253–266.

10. Thaller G, Krämer W, Winter A, Kaupe B, Erhardt G, et al. (2003) Effects of DGAT1 variants on milk production traits in German cattle breeds. J. Anim. Sci. 81: 1911–1918.

11. Schennink A, Stoop WM, Visker MHPW, Heck JML, Bovenhuis H, et al. (2007) DGAT1 underlies large genetic variation in milk-fat composition of dairy cows. Anim. Genet. 38: 467–473. doi:10.1111/j.1365–2052.2007.01635.x

12. Signorelli F, Orrù L, Napolitano F, De Matteis G, Scatà MC, et al. (2009) Exploring polymorphisms and effects on milk traits of the DGAT1, SCD1 and GHR genes in four cattle breeds. Livestock Science 125: 74–79. doi:10.1016/j.livsci.2009.02.014

13. Fazioli F, Minichiello L, Matoska V, Castagnino P, Miki T, et al. (1993) Eps8, a substrate for the epidermal growth factor receptor kinase, enhances EGF-dependent mitogenic signals. EMBO J 12: 3799–3808.

14. Guo D, Prins RM, Dang J, Kuga D, Iwanami A, et al. (2009) EGFR Signaling Through an Akt-SREBP-1-Dependent, Rapamycin-Resistant Pathway Sensitizes Glioblastomas to Antilipogenic Therapy. Sci. Signal. 2: ra82. doi:10.1126/scisignal.2000446

15. Loor JJ, Dann HM, Everts RE, Oliveira R, Green CA, et al. (2005) Temporal gene expression profiling of liver from periparturient dairy cows reveals complex adaptive mechanisms in hepatic function. Physiol. Genomics 23: 217–226. doi:10.1152/physiolgenomics.00132.2005

16. Bionaz M, Loor JJ (2008) ACSL1, AGPAT6, FABP3, LPIN1, and SLC27A6 are the most abundant isoforms in bovine mammary tissue and their expression is affected by stage of lactation. J. Nutr. 138: 1019–1024.

17. Hamza MS, Pott S, Vega VB, Thomsen JS, Kandhadayar GS, et al. (2009) De-novo identification of PPARgamma/RXR binding sites and direct targets during adipogenesis. PLoS ONE 4: e4907. doi:10.1371/journal.pone.0004907

18. Chakravarty K, Cassuto H, Reshef L, Hanson RW (2005) Factors that control the tissue-specific transcription of the gene for phosphoenolpyruvate carboxykinase-C. Crit. Rev. Biochem. Mol. Biol. 40: 129–154. doi:10.1080/10409230590935479

19. Chatterjee S, Szustakowski JD, Nanguneri NR, Mickanin C, Labow MA, et al. (2009) Identification of Novel Genes and Pathways Regulating SREBP Transcriptional Activity. PLoS ONE 4: e5197. doi:10.1371/journal.pone.0005197

20. Bouwman AC, Bovenhuis H, Visker MH, van Arendonk JA (2011) Genome-wide association of milk fatty acids in Dutch dairy cattle. BMC Genetics 12: 43. doi:10.1186/1471-2156-12-43

21. Beigneux AP, Vergnes L, Qiao X, Quatela S, Davis R, et al. (2006) Agpat6–a novel lipid biosynthetic gene required for triacylglycerol production in mammary epithelium. J. Lipid Res. 47: 734–744. doi:10.1194/jlr.M500556-JLR200

22. Vergnes L, Beigneux AP, Davis R, Watkins SM, Young SG, et al. (2006) Agpat6 deficiency causes subdermal lipodystrophy and resistance to obesity. J. Lipid Res. 47: 745–754. doi:10.1194/jlr.M500553-JLR200

23. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, et al. (2009) A whole-genome assembly of the domestic cow, Bos taurus. Genome Biology 10: R42. doi:10.1186/gb-2009-10-4-r42

24. VanRaden PM (2008) Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414–4423. doi:10.3168/jds.2007-0980

25. Browning BL, Browning SR (2009) A Unified Approach to Genotype Imputation and Haplotype-Phase Inference for Large Data Sets of Trios and Unrelated Individuals. The American Journal of Human Genetics 84: 210–223. doi:10.1016/j.ajhg.2009.01.005

26. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, et al. (2010) Variance component model to account for sample structure in genome-wide association studies. Nature Genetics 42: 348–354. doi:10.1038/ng.548

27. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics 88: 76–82. doi:10.1016/j.ajhg.2010.11.011

28. Gremme G, Brendel V, Sparks ME, Kurtz S (2005) Engineering a software tool for gene structure prediction in higher organisms. Information and Software Technology 47: 965–978. doi:10.1016/j.infsof.2005.09.005

29. Quackenbush J, Cho J, Lee D, Liang F, Holt I, et al. (2001) The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic Acids Res. 29: 159–164.

30. Lee E, Harris N, Gibson M, Chetty R, Lewis S (2009) Apollo: A Community Resource for Genome Annotation Editing. Bioinformatics 25: 1836–1837. doi:10.1093/bioinformatics/btp314

31. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, et al. (2005) MatInspector and Beyond: Promoter Analysis Based on Transcription Factor Binding Sites. Bioinformatics 21: 2933–2942. doi:10.1093/bioinformatics/bti473

32. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, et al. (2009) JASPAR 2010: The Greatly Expanded Open-Access Database of Transcription Factor Binding Profiles. Nucl. Acids Res.

33. Schug J (2008) Using TESS to predict transcription factor binding sites in DNA sequence. Curr Protoc Bioinformatics Chapter 2: Unit 2.6. doi:10.1002/0471250953.bi0206s21

34. Nickerson DA, Tobe VO, Taylor SL (1997) PolyPhred: Automating the Detection and Genotyping of Single Nucleotide Substitutions Using Fluores-cence-Based Resequencing. Nucl. Acids Res. 25: 2745–2751. doi:10.1093/nar/25.14.2745

35. Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res. 8: 195–202.