A post-GWAS confirming the genetic effects and functional polymorphisms of AGPAT3 gene on milk fatty acids in dairy cattle

Lijun Shi
Institute of Animal Science, Chinese Academy of Agricultural Sciences

Xin Wu
China Agricultural University

Zhu Ma
Beijing Dairy Cattle Center

Xiaoqing Lv
Beijing Dairy Cattle Center

Lin Liu
Beijing Dairy Cattle Center

Yanhua Li
Beijing Dairy Cattle Center

Feng Zhao
Beijing Dairy Cattle Center

Bo Han
China Agriculture University

Yuze Yang
Beijing General Station of Animal Husbandry

Dongxiao Sun (sundx@cau.edu.cn)
China Agricultural University https://orcid.org/0000-0002-7512-5946

Research

**Keywords:** AGPAT3, Chinese Holstein, Genetic effects, Milk fatty acids, Potential causal mutation

**DOI:** [https://doi.org/10.21203/rs.3.rs-38950/v2](https://doi.org/10.21203/rs.3.rs-38950/v2)

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](https://creativecommons.org/licenses/by/4.0/)
Abstract

**Background:** People are paying more attention to the healthy and balanced diet with the improvement of their living standards. Milk fatty acids (FAs) have been reported to be related to some atherosclerosis and coronary heart diseases in human. In our previous genome-wide association study (GWAS) on milk FAs in dairy cattle, 83 genome-wide significant single nucleotide polymorphisms (SNPs) were detected. Among them, two SNPs, ARS-BFGL-NGS-109493 and BTA-56389-no-rs associated with C18index ($P = 0.0459$), were located in the upstream of 1-acylglycerol-3-phosphate O-acetyltransferase 3 ($AGPAT3$) gene. $AGPAT3$ is involved in glycerol-lipid, glycerol-phospholipid metabolism and phospholipase D signaling pathways. Hence, it was inferred as a candidate gene for milk FAs. The aim of this study was to further confirm the genetic effects of the $AGPAT3$ gene on milk FA traits in dairy cattle.

**Results:** Through re-sequencing the complete coding region, and 3,000bp of 5' and 3' regulatory regions of the $AGPAT3$ gene, a total of 17 SNPs were identified, including four in 5' regulatory region, one in 5' untranslated region (UTR), three in introns, one in 3' UTR, and eight in 3' regulatory region. By the linkage disequilibrium (LD) analysis with Haploview4.1 software, two haplotype blocks were observed that were formed by four and 12 identified SNPs, respectively. Using SASH2, we performed single locus-based and haplotype-based association analysis on 24 milk FAs in 1,065 Chinese Holstein cows, and discovered that all the SNPs and the haplotype blocks were significantly associated with C6:0, C8:0 and C10:0 ($P < 0.0001 - 0.0384$). Further, with Genomatix, we predicted that four SNPs in 5' regulatory region (g.146702957G>A, g.146704373A>G, g.146704618A>G and g.146704699G>A) changed the transcription factor binding sites (TFBSs) for transcription factors SMARCA3, REX1, VIMY, BRACH, NIKX2, ZBED4, SP1, USF1, ARNT and FOX1. Out of them, two SNPs were validated to impact transcriptional activity by performing luciferase assay that the alleles A of both SNPs, g.146704373A>G and g.146704618A>G, increased the transcriptional activities of $AGPAT3$ promoter compared with alleles G ($P = 0.0004$).

**Conclusions:** In conclusion, our findings first demonstrated the significant genetic associations of the $AGPAT3$ gene with milk FAs in dairy cattle, and two potential causal mutations were detected.

Introduction

Milk fat is one of critical breeding objectives in dairy cattle. It is comprised of triglyceride (>95%), diglyceride (2%), phospholipids (1%), cholesterol (0.05%) and small amount of free fatty acids (FAs) (~0.1%) [1]. The main components of triglyceride are glycerin and FAs, in which, the FAs act as precursors for the formation of other aroma components, such as esters and alcohols [2]. For the various milk fatty acid traits in Holstein cows, the estimated heritability values have been reported to be 0.14-0.33 for saturated fatty acids (SFAs) and 0.08-0.29 for unsaturated fatty acids (UFAs) [3-7].

Genome-wide association study (GWAS) is a commonly used strategy to identify potential genetic variants underlying important complex traits in human and domestic animals. So far, some candidate genes and QTL regions for milk production traits have been detected with GWA studies in dairy cattle, such as $DLGAP1$, $AP2B1$, $SCD$, BTA1 (1.59 ~ 3.37 Mbp), and BTA3 (70.34 ~ 73.69 Mbp) [8-13]. In our previous GWAS for milk FAs in Chinese Holstein cows, 83 genome-wide significant single nucleotide polymorphisms (SNPs) were detected in total [12], in which, two SNPs (ARS-BFGL-NGS-109493 and BTA-56389-no-rs) associated with C18index ($P = 0.0459$), were located in the upstream of 1-acylglycerol-3-phosphate O-acetyltransferase 3 ($AGPAT3$) gene. In addition, we performed a joint GWAS for milk FAs in combined Chinese and Danish Holstein populations and found that a chromosome-wide significant QTL region of 146.29 ~ 146.31 Mbp on BTA1 was associated with C18index ($P = 0.011$). The $AGPAT3$ gene was nearby this region with approximately 400 kb. 1-acylglycerol-sn-glycerol 3-phosphate acyltransferase (AGPAT), encoded by the $AGPAT3$ gene, is one of the isoforms of AGPATs [14] and is involved in the glycerolphospholipid metabolism (ko00564), and phospholipase D signaling pathway (ko04072). Mammalian AGPAT catalyzed the acylation of lysophosphatidic acid to form the phosphosphatidic acid that was the precursor of all glycerolipids. Therefore, it was implied that the $AGPAT3$ gene was a promising candidate gene for milk FA traits in dairy cattle. The purpose of the present study was to further determine whether the $AGPAT3$ gene had significant genetic effects on milk FAs in a Chinese Holstein cow population.

Materials And Methods

**Animals and phenotypic data**

In this study, a total of 1,065 Chinese Holstein cows were used as described in a previous research [15], which milk samples were collected in Beijing Dairy Cattle Center (www.bdcc.com.cn) to measure milk FA contents. With the gas chromatography method, a total of 16 milk FAs (C6:0, C8:0, C10:0, C11:0, C12:0, C13:0, C14:0, C15:0, C16:0, C16:1, C17:0, C18:0, C18:1 cis-9 and C20:0) were measured as the weight proportion of total fat weight [12]. With the phenotypes, we calculated five indices based on the formula:

\[
\text{C14index} = \frac{\text{C14:0} \times 100}{\sum \text{C14:0}}, \quad \text{C16index} = \frac{\text{C16:0} \times 100}{\sum \text{C16:0}}, \quad \text{C18index} = \frac{\text{C18:0} \times 100}{\sum \text{C18:0}}, \quad \text{C17index} = \frac{\text{C17:0} \times 100}{\sum \text{C17:0}}, \quad \text{C16:1index} = \frac{\text{C16:1} \times 100}{\sum \text{C16:1}}.
\]

In addition, the summarized SFA and UFA, and SFA/UFA were obtained as well.

**SNP identification and genotyping**

Based on the genomic sequence of bovine $AGPAT3$ gene (Gene ID: 506607), 14 pairs of primers (Table S1) were designed by the Primer 3 version 4.0 (http://bioinfo.ut.ee/primer3-0.4.0/) and were synthesized in the Beijing Genomics Institute (Beijing, China) to amplify all the exons with partial adjacent intron region, and 3,000 bp of 5' and 3' regulatory regions. As previously described [15], two DNA pools were constructed and the polymerase chain reaction (PCR) amplifications were performed with each DNA pool as template. To identify potential polymorphisms, the PCR amplification products were bi-directionally sequenced with an ABI3730XL DNA analyzer (Applied Biosystems, Foster, CA, USA). Then, the identified SNPs were genotyped for the 1,065 cows by the matrix-assisted laser-desorption/ionization time of flight mass spectrometry (MALDI-TOF MS, Sequenom MassARRAY, Biouyong Technologies Inc, HK).

**Linkage disequilibrium (LD) and association analyses**
We estimated the LD among the identified SNPs of AGPAT3 gene with Haploview 4.1 (Broad Institute, Cambridge, MA, USA).

For association analysis, the 1,065 cows were traced back to three-generation pedigrees to construct the kinship matrix (\( \mathbf{A} \)) matrix by SAS 9.2 (SAS institute, Cary, NC, USA), so that 3,335 individuals were totally included. Single-locus and haplotype-based associations with 24 kind of milk FAs were performed by the following mixed animal model with SAS 9.2:

\[
Y_{ijkm} = \mu + g_i + h_j + l_k + a_i + b_l + m_j + h_{ik} + \varepsilon_{ijkm}
\]

Here, \( Y_{ijkm} \) is the phenotypic value of each milk fatty acid trait; \( \mu \) is the overall mean; \( g_i \) is the fixed effect corresponding to the genotype or haplotype combination of individual \( i \); \( h_j \) and \( l_k \) were the fixed effect of farm \( j \) and stage of lactation \( k \), respectively; \( m_j \) is the random polygenic effect; \( h_{ik} \) (\( m_{j}=1\sim293 \)) is the fixed effect of age at calving \( m \); \( a_i \) is the regression coefficient of covariate \( q_i \) and \( b_l \) is the random residual. Further, we calculated the additive effect (a), dominant effect (d), and allele substitution effect (\( \alpha \)) according to \( a = \frac{A - B - 2AB}{2} \), \( d = AB - \frac{A + B}{2} \), and \( \alpha = a + d(q - P) \). Here, AA, AB and BB were the least square means of milk FAs corresponding to the genotypes, and \( p \) and \( q \) were the frequencies of allele A and B, respectively.

Prediction of changes of transcription factor binding sites (TFBSs) caused by the SNPs in 5' regulatory region

We used the Genomatix software suite v3.9 (http://www.genomatix.de/cgi-bin/welcome/welcome.pl?s=d1b5c9a9015b02bb3b1a806f9c03293f) [18] to predict whether the four SNPs in 5' regulatory region of AGPAT3 (g.146702957G>A, g.146704373A>G, g.146704618A>G, and g.146704699G>A) changed the TFBSs.

Recombinant plasmid construction and luciferase assay

To detect the allele-specific effects of the SNPs g.146702957G>A, g.146704373A>G, g.146704618A>G, and g.146704699G>A on the transcriptional activity of AGPAT3 gene, eight luciferase reporter gene fragments (G and A of g.146702957G>A; A and G of g.146704373A>G; A and G of g.146704618A>G; and G and A of g.146704699G>A) corresponding to the eight alleles of the four SNPs (Fig. 1a) were designed and synthesized (Geneiz, Suzhou, China). The eight fragments with the KpnI and Nhel restriction sites at the 5' and 3' termini, respectively, were cloned into the pGL4.14 luciferase assay vector (Promega, Madison, USA). In addition, all the plasmids were purified by the Endo-free Plasmid DNA Mini Kit (OMEGA, omega bio-tek, Norcross, Georgia, USA), and were sequenced to confirm the integrity of each construct’s insertion.

The human embryonic kidney (HEK) 293T cells were cultured with Dulbecco’s modified Eagle’s medium (Gibco, Life Technologies) and 10% fetal bovine serum (Gibco) at 37°C in a humidified incubator containing 5% CO\(_2\). Before transfection, about 2×10\(^5\) cells were seeded in each 24-well plate. For eight luciferase reporter gene fragments of g.146702957G>A, g.146704373A>G, g.146704618A>G, and g.146704699G>A, 500 ng constructed plasmid was co-transfected with 10 ng pRL-TK Renilla luciferase reporter vector (Promega) into each well. All the experiments were performed in three replicates for each construct. Approximate 48h after transfection, the cells were harvested and the activity of both renilla and Renilla luciferases were measured with a Dual-Luciferase Reporter Assay System (Promega) on a Modulus microplate multimode reader (Turner Biosystems, CA, USA). The average statistic of three replicates were calculated as the normalized luciferase data (Firefly/Renilla).

Results

Identification of SNPs

A total of 17 SNPs of the AGPAT3 gene was detected in this study (Table 1), which consisted of four (g.146702957G>A, g.146704373A>G, g.146704618A>G, and g.146704699G>A) in 5' flanking region, one (g.146705692G>A) in 5' untranslated region (UTR), three (g.146725085T>C, g.146726096A>G, and g.146729107A>C) in introns, one (g.146735090G>T) in 3' UTR, and eight (g.146737188C>T, g.146737545G>A, g.146737748T>C, g.146737849C>T, g.146737879T>G, g.146737916T>C, g.146737946C>T, and g.146738055G>A) in 3' flanking region. The genotype and allele frequencies of the identified SNPs were presented in Table 1.

Estimation of LD among the identified SNPs of AGPAT3

We used the haplごview 4.1 to estimate the LD among these 17 SNPs, and observed two haplotype blocks (Fig. 2) that was formed by four and 12 SNPs, respectively. The haplotype block 1 included four haplotype combinations, namely, H1: GAAG (38%), H2: GAAA (32.2%), H3: AGGG (26.6%), and H4: GAGG (3%), and the haplotype block 2 had six haplotype combinations: H1 = GTAAGCGTCTTC, H2 = GCAGCGTCTTC, H3 = GCAATCGTCTTC, H4 = ACACCGGTCTTC, H5 = GTGATCGTCTTC, and H6 = GCAACCGTCTTC with their frequencies of 20%, 39.8%, 13.4%, 13.4%, 7.9% and 4.1%.

Associations between AGPAT3 and milk FAs

The associations of the 17 SNPs with 24 milk FAs were summarized in Table 2. Among these SNPs, 17 were strongly associated with C6:0 (\( P < 0.0001 \sim 0.0004 \)) and C8:0 (\( P < 0.0001 \sim 0.0384 \)); 14 were significantly associated with total index (\( P < 0.0001 \sim 0.0318 \)); ten were significantly associated with C10:0 (\( P = 0.0016 \sim 0.0151 \)); nine were strongly associated with C17:1 (\( P < 0.0001 \sim 0.0149 \)); seven were significantly associated with C20:0 (\( P < 0.0001 \sim 0.0072 \)); five had significant associations with C14:0 (\( P < 0.0001 \sim 0.0477 \)); five were strongly associated with C17index (\( P = 0.0006 \sim 0.0389 \)); five had strong associations with C18:1cis-9 (\( P < 0.0001 \sim 0.0258 \)); three had significant associations with C18:0 (\( P = 0.0020 \sim 0.0246 \)); three had strong associations with SFA (\( P < 0.0001 \sim 0.0434 \)); two were significantly associated with C17:0 (\( P = 0.0212 \sim 0.0413 \)); two were significantly associated with UFA (\( P < 0.0001 \) and \( P \sim 0.0001 \)).
effects, we deduced that the improved genes to positively regulate their transcriptional activity [28]. Hypermethylation status reduced the transcription factor activity of SP1 located in the weakly activated the transcription of the bindings of transcription factors BRACH and NKX26, respectively, and the allele G of g.146704618A>G altered the bindings of transcription factors ZBED4 under understanding the evolution of gene regulation [25]. In the present study, by luciferase assay, the alleles A of g.146704373A>G and g.146704618A>G in the 5' thousand base pairs in length, and can harbor many TFBSs [24]. It is essential to understand the evolution dynamics of transcription factor binding for Sequences-specifc binding of transcription factors to the regulatory regions on the DNA is a key regulatory mechanism that affects gene expression and GWASs and this study suggested that effects on C18index and C18:0. In addition, our results revealed that the independent Chinese Holstein population that was different from the precious GWA studies, we also observed that C16:1 [21]. In our previous GWA studies [12,13], Mammalian AGPAT catalyzed the acylation of lysophosphatidic acid to form the phosphatidic acid, which was the precursor of all glycerplipids [14]. For the FAs in dairy cattle.

lysophosphatic acid, a step in the phospholipid biosynthesis pathway [19]. Here, we detected that the metabolism (ko00561, ko00564 and ko04072). In human, docosapentaenoic acid as the substrate of AGPAT3 is involved in pathways related to lipid protein transfers a fatty acid in sn-2 position of lysophosphatic acid, a step in the phospholipid biosynthesis pathway [19]. Here, we detected that the AGPAT3 gene mainly impacted the medium-chain milk FAs in dairy cattle.

Mammalian AGPAT catalyzed the acylation of lysophosphatidic acid to form the phosphatidic acid, which was the precursor of all glycerlipids [14]. For the AGPAT families, AGPAT7 had significant association with milk FA CLA [20], and AGPAT6 was strongly associated with C14:0, C16:0, C10:1, C12:1, C14:1 and C16:1 [21]. In our previous GWA studies [12,13], AGPAT3 gene was identified as a candidate for two milk FAs, C18index and C18:0. In this study, using an independent Chinese Holstein population that was different from the precious GWAS studies, we also observed that AGPAT3 showed the significant genetic effects on C18index and C18:0. In addition, our results revealed that the AGPAT3 had strong associations with C6:0, C8:0 and C10:0. Overall, the previous GWASs and this study suggested that AGPAT3 gene had significantly genetic effects on milk FAs.

Discussion

This study was a follow-up investigation for our previous GWAS on milk FAs in Chinese Holstein [12]. AGPAT3 is involved in pathways related to lipid metabolism (ko00651, ko00564 and ko04072). In human, docosapentaenoic acid as the substrate of AGPAT3 protein transfers a fatty acid in sn-2 position of lysophosphatic acid, a step in the phospholipid biosynthesis pathway [19]. Here, we detected that the AGPAT3 gene mainly impacted the medium-chain milk FAs in dairy cattle.

Mammalian AGPAT catalyzed the acylation of lysophosphatidic acid to form the phosphatidic acid, which was the precursor of all glycerlipids [14]. For the AGPAT families, AGPAT7 had significant association with milk FA CLA [20], and AGPAT6 was strongly associated with C14:0, C16:0, C10:1, C12:1, C14:1 and C16:1 [21]. In our previous GWA studies [12,13], AGPAT3 gene was identified as a candidate for two milk FAs, C18index and C18:0. In this study, using an independent Chinese Holstein population that was different from the precious GWAS studies, we also observed that AGPAT3 showed the significant genetic effects on C18index and C18:0. In addition, our results revealed that the AGPAT3 had strong associations with C6:0, C8:0 and C10:0. Overall, the previous GWASs and this study suggested that AGPAT3 gene had significantly genetic effects on milk FAs.

Sequences-specific binding of transcription factors to the regulatory regions on the DNA is a key regulatory mechanism that affects gene expression and hence heritable phenotypic variation [22,23]. Eukaryotic regulatory sequences, including enhancers and promoters, are typically between a hundred and several thousand base pairs in length, and can harbor many TFBSs [24]. It is essential to understand the evolution dynamics of transcription factor binding for understanding the evolution of gene regulation [25]. In the present study, by luciferase assay, the alleles A of g.146704373A>G and g.146704618A>G in the 5'-flanking region strongly increased the transcription activity of the AGPAT3 gene than the alleles G, in which, the alleles A and G of g.146704373A>G changed the bindings of transcription factors BRACH and NKX26, respectively, and the allele G of g.146704618A>G altered the bindings of transcription factors ZBED4 and SP1. In addition, for g.146704373A>G and g.146704618A>G, the individuals with the AA genotypes had significantly lower contents of C6:0 and C8:0 than that of alleles G. BRACH as a regulatory factor induced NANOG expression to participate in establishment of cancer stem cell characteristics [26], and NKX26 weakly activated the transcription of Cx40 to impact the heart development [27]. ZBED4 bound in vivo to specific elements in the promoter region of retinal genes to positively regulate their transcriptional activity [28]. Hypermethylation status reduced the transcription factor activity of SP1 located in the ZNF132 promoter region to make the ZNF132 promote the abilities of the tumorigenicity of cells in a nude mouse model [29]. Considering the significant association effects, we deduced that the improved AGPAT3 gene expression by changing the binding status of the transcription factor BRACH, to reduce the C6:0 and C8:0. Also, the transcription factors NKX26, ZBED4 and SP1 might have the contrary effects.
Nowadays, genomic selection is the main implication for dairy cattle breeding, where the genomic chips are used. Among the SNP markers in these chips, most of them were collected from the current SNP database and almost evenly distributed across the whole genome. Hence, g.146704373A>G and g.146704618A>G of AGPAT3 as the potentially causal mutations could be put into the SNP chip instead of used in marker selection to increase selection efficiency in some specific dairy cattle populations to improve the contents of milk FAs.

Conclusion
In conclusion, through a post-GWAS approach, our study firstly indicated there were significant genetic associations between the AGPAT3 gene and milk FAs in dairy cattle. Further, we found that two SNPs in S' regulatory region (g.146704373A>G and g.146704618A>G) changed the transcriptional activity of AGPAT3 implying their potential causal function. These findings provided important molecular information for dairy cattle breeding.

Abbreviations
a: Additive; AGPAT: 1-acylglycerol-sn-glycero 3-phosphate acyltransferase; AGPAT3: 1-acylglycerol-3-phosphate 0-acyltransferase 3; ARNT: AhR nuclear translocator homodimers; BRACH: Brachyury; d: dominant; FA: fatty acid; FOXA1: Forkhead box protein A1, hepatocyte nuclear factor 3-alpha (HNF-3-alpha); GWAS: genome-wide association study; HEK: Human Embryonic Kidney; LD: linkage disequilibrium; NKX26: NK2 homeobox 6, Csx2; PCR: polymerase chain reaction; REX1: REX1 transcription factor; zinc finger protein 42; SFA: saturated fatty acids; SMARCA3: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3; SNP: single nucleotide polymorphism; SP1: Stimulating protein 1, ubiquitous zinc finger transcription factor; TFBS: transcription factor binding site; UFA: unsaturated fatty acid; USF1: Upstream stimulating factor 1; UTR: untranslated region; VMYB: v-Myb, variant of AMV v-myb; ZBED4: Zinc finger, BED-type containing 4; GC-box binding sites; α: substitution.

Declarations
Acknowledgements
We appreciate Beijing Dairy Cattle Center and Beijing Sanyuanlvhe Dairy Farming Center for providing the milk, blood, and semen samples of Chinese Holstein.

Author contributions
DS and YY conceived and designed the experiments, LL prepared the DNA samples for SNP identification and genotyping with the help of XW, ZM, XL, YL, and FZ, XL measured the phenotypes of milk fatty acids, LS and XW analyzed the data, and LS, BH and DS prepared the manuscript. All authors read and approved the final manuscript.

Funding
This work was financially supported by the Beijing Science and Technology Program (202000105, D171100002417001), National Science and Technology Programs of China (2013AA102504), National Natural Science Foundation of China (31872330, 31802041), earmarked fund for Modern Agro-industry Technology Research System (CARS-36), and the Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62).

Availability of data and materials
All relevant data are available within the article and its supplementary information.

Ethics Approval and consent to participate
All protocols for collection of the samples of experimental individuals and phenotypic observations were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University (Permit Number: DK996). Milk, blood and semen samples were collected specifically for this study following standard procedures with the full agreement of the Beijing Sanyuanlvhe Dairy Farming Center who owned the Holstein cows and bulls, respectively.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Staniewski B, Kielczewska K, Smoczynski M, Baranowska M, Czerniewicz M, Brandt W. Effect of high pressures on the composition of milk fat triacylglycerols. Milchwissenschaft. 2012;67(1):18-21.
2. Delafuente MA, Fontecha J, Juarez M. Fatty-Acid Composition of the Triglyceride and Free Fatty-Acid Fractions in Different Cows-Milk, Ewes-Milk and Goats-Milk Cheeses. Z Lebensm Unters For. 1993;196(2):155-8.
3. Krag K, Poulsen NA, Larsen MK, Larsen LB, Janss LL, Buitenhuis B. Genetic parameters for milk fatty acids in Danish Holstein cattle based on SNP markers using a Bayesian approach. Bmc Genet. 2013;14:79.

4. Petrin J, Lung LH, Rodriguez MA, Salvian M, Pertille F, Rovadoscki GA, et al. Genetic parameters for milk fatty acids, milk yield and quality traits of a Holstein cattle population reared under tropical conditions. Journal of animal breeding and genetics. 2016;133(5):384-95.

5. Narayana SG, Schenk FK, Fleming A, Koeck A, Malchiodi F, Jamrozek J, et al. Genetic analysis of groups of mid-infrared predicted fatty acids in milk. Journal of dairy science. 2017;100(6):4731-44.

6. Stoop WM, Schenck K, Visser MH, Muller et al. Genome-wide scan for bovine milk-fat composition. I. Quantitative trait loci for short- and medium-chain fatty acids. Journal of dairy science. 2009;92(9):4664-75.

7. Schenck K, Stoop WM, Visser MH, van der Poel JJ, Bovenhuis H, van Arendonk JA. Short communication: Genome-wide scan for bovine milk-fat composition. II. Quantitative trait loci for long-chain fatty acids. Journal of dairy science. 2009;92(9):4676-82.

8. Chen Z, Yao Y, Ma P, Wang Q, Pan Y. Haplotype-based genome-wide association study identifies loci and candidate genes for milk yield in Holsteins. Plos One. 2018;13(2):e0192695.

9. Palombo V, Milanesi M, Sgorlon S, Capomaccio S, Mele M, Nicolazzi E, et al. Genome-wide association study of milk fatty acid composition in Italian Simmental and Italian Holstein cows using single nucleotide polymorphism arrays. Journal of dairy science. 2018.

10. Knutsen TM, Olsen HG, Tafintseva V, Svendsen M, Kohler A, Kent MP, et al. Unravelling genetic variation underlying de novo-synthesis of bovine milk fatty acids. Sci Rep. 2018;8(1):2179.

11. Yue SJ, Zhao YQ, Gu XR, Yin B, Jiang YL, Wang ZH, et al. A genome-wide association study suggests new candidate genes for milk production traits in Chinese Holstein cattle. Anim Genet. 2017;48(6):677-81.

12. Li C, Sun D, Zhang S, Wang S, Wu X, Zhang Q, et al. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. Plos One. 2014;9(5):e96186.

13. Li X, Buitenhuis AJ, Lund MS, Li C, Sun D, Zhang Q, et al. Joint genome-wide association study for milk fatty acid traits in Chinese and Danish Holstein populations. Journal of dairy science. 2015;98(11):8152-63.

14. Lu B, Jiang YJ, Zhou Y, Xu FY, Hatch GM, Choy PC. Cloning and characterization of murine 1-acyl-sn-glycerol 3-phosphate acyltransferases and their regulation by PPARalpha in murine heart. Biochem J. 2005;385(Pt 2):469-77.

15. Shi L, Liu L, Ma Z, Lv X, Li C, Xu L, et al. Identification of genetic associations of ECHS1 gene with milk fatty acid traits in dairy cattle. Anim Genet. 2019;50(5):430-8.

16. Kelsey JA, Cori BA, Collier RJ, Bauman DE. The effect of breed, parity, and stage of lactation on conjugated linoleic acid (CLA) in milk fat from dairy cows. J Dairy Sci. 2003;86(8):2588-97.

17. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 1996: Ed. 4 xv + 464 pp.

18. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics. 2005;21(13):2933-42.

19. Lemaitre RN, Tanaka T, Tanabe K, Manichaikul A, Foy M, Kabagambe EK, et al. Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. PLoS genetics. 2011;7(7):e1002193.

20. Gebreyesus G, Buitenhuiss AJ, Poulsen NA, Visser M, Zhang Q, van Valenberg HJF, et al. Multi-population GWAS and enrichment analyses reveal novel genomic regions and promising candidate genes underlying bovine milk fatty acid composition. Bmc Genomics. 2019;20(1):178.

21. Bouwman AC, Bovenhuis H, Visser MH, van Arendonk JA. Genome-wide association of milk fatty acids in Dutch dairy cattle. Bmc Genet. 2011;12:43.

22. Tugnul M, Paixao T, Barton NH, Tkacik G. Dynamics of Transcription Factor Binding Site Evolution. PLoS genetics. 2015;11(11):e1005639.

23. Talebzadeh M, Zare-Mirakabad F. Transcription factor binding sites prediction based on modified nucleosomes. Plos One. 2014;9(2):e89226.

24. Yao P, Lin P, Gokoolparsad A, Assareh A, Thang MW, Voineagu I. Coexpression networks identify brain region-specific enhancer RNAs in the human brain. Nature neuroscience. 2015;18(8):1168-74.

25. Schmidt D, Wilson MD, Ballester B, Schwalm PC, Brown GD, Marshall A, et al. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. Science. 2010;328(5981):1036-40.

26. Sarkar D, Shields B, Davies ML, Muller J, Wakeman JA. BRACHYURY confers cancer stem cell characteristics on colorectal cancer cells. International journal of cancer. 2012;130(2):328-37.

27. Heathcote K, Braybrook C, Abushaban L, Guy M, Khetyar ME, Patton MA, et al. Common arterial trunk associated with a homeodomain mutation of NKX2.6. Human molecular genetics. 2005;14(5):585-93.

28. Saghizadeh M, Gribanova Y, Akhmedov NB, Farber DB. ZBED4, a cone and Muller cell protein in human retina, has a different cellular expression in mouse. Mol Vis. 2011;17(219-20):2011-8.

29. Jiang D, He ZL, Wang CJ, Zhou YH, Li F, Pu WL, et al. Epigenetic silencing of ZNF132 mediated by methylation-sensitive Sp1 binding promotes cancer progression in esophageal squamous cell carcinoma. Cell Death Dis. 2018;10(1):1.

Tables

Table 1: Information of 17 SNPs of AGPAT3 gene with genotypic and allele frequencies
| SNP name | Location | Position (GRCh 31.1) | Gen Bank no. | Origin | Genotypes | NO. | Frequency | Allele | Frequency |
|----------|-----------|----------------------|--------------|--------|------------|-----|-----------|--------|-----------|
| g.146702957C>T | 5' flanking | Chr1:146702957 | rs210638640 | NCBI | AA | 78 | 0.0741 | A | 0.2662 |
|            |            |                      |              |        | GG | 564 | 0.5356 | G | 0.7338 |
|            |            |                      |              |        | GA | 411 | 0.3844 | G | 0.7338 |
| g.146743738G>A | 5' flanking | Chr1:146743738 | rs209442450 | NCBI | AA | 560 | 0.5333 | A | 0.4667 |
|            |            |                      |              |        | GG | 78 | 0.0748 | G | 0.9252 |
|            |            |                      |              |        | GA | 405 | 0.3815 | G | 0.9252 |
| g.146746613G>A | 5' flanking | Chr1:146746613 | rs110551271 | NCBI | AA | 531 | 0.4910 | A | 0.5090 |
|            |            |                      |              |        | GG | 90 | 0.0850 | G | 0.9150 |
|            |            |                      |              |        | GA | 406 | 0.4240 | G | 0.9150 |
| g.146746899G>A | 5' flanking | Chr1:146746899 | rs110278717 | NCBI | AA | 113 | 0.1076 | A | 0.3219 |
|            |            |                      |              |        | GG | 487 | 0.4638 | G | 0.5362 |
|            |            |                      |              |        | GA | 450 | 0.4262 | G | 0.5362 |
| g.146750865G>A | 5' UTR    | Chr1:146750865 | rs63281404 | NCBI | AA | 10 | 0.0095 | A | 0.9805 |
|            |            |                      |              |        | GG | 778 | 0.7273 | G | 0.2727 |
|            |            |                      |              |        | AG | 247 | 0.2327 | G | 0.2727 |
| g.146750865G>A | Inters-5   | Chr1:146750865 | rs110867007 | NCBI | GG | 634 | 0.6021 | G | 0.3979 |
|            |            |                      |              |        | TT | 37 | 0.0379 | T | 0.9621 |
|            |            |                      |              |        | CT | 352 | 0.3269 | C | 0.6731 |
| g.146769948G>A | Inters-6   | Chr1:146769948 | rs578203374 | NCBI | TT | 144 | 0.1503 | T | 0.3997 |
|            |            |                      |              |        | CT | 492 | 0.4760 | C | 0.5240 |
| g.146729107T>A | Inters-7   | Chr1:146729107 | rs43276015 | NCBI | AA | 166 | 0.1582 | A | 0.4418 |
|            |            |                      |              |        | GG | 370 | 0.3618 | G | 0.6382 |
|            |            |                      |              |        | GA | 155 | 0.4814 | G | 0.5186 |
| g.146730099G>T | 3' UTR     | Chr1:146730099 | rs579405887 | NCBI | CC | 169 | 0.1608 | C | 0.4392 |
|            |            |                      |              |        | TT | 376 | 0.3718 | T | 0.6282 |
|            |            |                      |              |        | CT | 506 | 0.4914 | T | 0.5086 |
| g.146731738T>C | 3' flanking | Chr1:146731738 | rs63532320 | NCBI | CC | 376 | 0.3580 | C | 0.6420 |
|            |            |                      |              |        | TT | 145 | 0.1574 | T | 0.8426 |
|            |            |                      |              |        | CT | 507 | 0.4926 | C | 0.5074 |
| g.146737545G>A | 3' flanking | Chr1:146737545 | rs43766238 | NCBI | GG | 146 | 0.1587 | G | 0.8413 |
|            |            |                      |              |        | TT | 377 | 0.3604 | T | 0.6396 |
|            |            |                      |              |        | CT | 502 | 0.4974 | C | 0.5026 |
| g.146737774C>T | 3' flanking | Chr1:146737774 | rs43760756 | NCBI | CC | 149 | 0.1608 | C | 0.4392 |
|            |            |                      |              |        | TT | 376 | 0.3718 | T | 0.6282 |
|            |            |                      |              |        | CT | 506 | 0.4914 | T | 0.5086 |
| g.146738068C>T | 3' flanking | Chr1:146738068 | rs43760757 | NCBI | CC | 371 | 0.3540 | C | 0.6460 |
|            |            |                      |              |        | TT | 147 | 0.1594 | T | 0.8406 |
|            |            |                      |              |        | CT | 510 | 0.4866 | C | 0.5134 |
| g.146738797G>C | 3' flanking | Chr1:146738797 | rs43760758 | NCBI | AA | 12 | 0.0114 | A | 0.9886 |
|            |            |                      |              |        | GG | 780 | 0.7293 | G | 0.2707 |
|            |            |                      |              |        | AG | 263 | 0.2403 | G | 0.7597 |
| g.146739042C>T | 3' flanking | Chr1:146739042 | rs43760759 | NCBI | CC | 536 | 0.5174 | C | 0.4826 |
|            |            |                      |              |        | TT | 78 | 0.0726 | T | 0.9274 |
|            |            |                      |              |        | TC | 422 | 0.4774 | T | 0.5226 |
| g.146739786C>T | 3' flanking | Chr1:146739786 | rs43760760 | NCBI | AA | 894 | 0.8442 | A | 0.1558 |
|            |            |                      |              |        | GG | 6 | 0.0057 | G | 0.9943 |
|            |            |                      |              |        | AG | 158 | 0.1513 | G | 0.8487 |
| g.146739853G>A | 3' flanking | Chr1:146739853 | rs432206340 | NCBI | AA | 206 | 0.2086 | A | 0.7914 |
|            |            |                      |              |        | CC | 287 | 0.2744 | C | 0.7256 |
|            |            |                      |              |        | CA | 542 | 0.5274 | A | 0.4726 |

Note: NO.: Number of cows with corresponding genotypes. UTR: untranslated region.

Table 2: Association between 17 SNPs and milk fatty acid traits in Chinese Holstein cows (LSM±SE)
| Compound | Mean ± SEM | p-value | Mean ± SEM | p-value | Mean ± SEM | p-value | Mean ± SEM | p-value | Mean ± SEM | p-value | Mean ± SEM | p-value |
|----------|------------|---------|------------|---------|------------|---------|------------|---------|------------|---------|------------|---------|
| C20:0 (%) | 0.1894 ± 0.0042 | 0.1931 ± 0.0026 | 0.1956 ± 0.0025 | 0.1922 ± 0.0024 | 0.1945 ± 0.0025 | 0.1968 ± 0.0042 | 0.1978 ± 0.0032 | 0.1986 ± 0.0032 | 0.1994 ± 0.0032 | 0.2011 ± 0.0032 | 0.2029 ± 0.0032 | 0.2047 ± 0.0032 |
| C18:0 (%) | 0.1894 ± 0.0042 | 0.1931 ± 0.0026 | 0.1956 ± 0.0025 | 0.1922 ± 0.0024 | 0.1945 ± 0.0025 | 0.1968 ± 0.0042 | 0.1978 ± 0.0032 | 0.1986 ± 0.0032 | 0.1994 ± 0.0032 | 0.2011 ± 0.0032 | 0.2029 ± 0.0032 | 0.2047 ± 0.0032 |
Note: LSM means least square mean. SE means standard error. *P* indicates the significances of the association analysis between the haplotype block and milk fatty acid traits. *P* is the raw value. **P** < 0.05, ***P** < 0.01. Different letter (small letters: *P* < 0.05; capital letters: *P* < 0.01) superscripts indicate significant differences among the haplotype combinations. The number in the brackets represents the number of cows for the corresponding haplotype combination.

Table 4 Changes of transcription factor binding site (TFBS) caused by the SNP in the 5′untranslated (UTR) and flanking regions of *AGPAT3*

| SNP        | Sequence | Transcription factor Name                  | Note                                      |
|------------|----------|--------------------------------------------|-------------------------------------------|
| g.146702657G>A | TCCCTGACOCCATTCCACCTGA | VMYB | v-Myb, variant of AMV v-myb |
|            | TCCCTGACCATTTCCACCTGA | SMARCA3 | SWI/SNF related, matrix associated, actin-dependent regulator of chromatin, subfamily A, member 3 |
|            |          | REX1 | REX1 transcription factor, zinc finger protein 42 |
| g.146704373A>G | CACGGGAAGTGGGGAGAAGT | BRACH | Brachyury |
|            | CACGGGAAGTGGGGAGAAGT | NKX2B | NK2 homeobox 2, Cdx2 |
| g.146704618A>G | CTCTTCCACC | ZBED4 | Zinc finger, BESD type containing 4, GC-box binding sites |
|            | CTCTTCCACC | SP1 | Stimulating protein 1, ubiquitous zinc finger transcription factor |
| g.146704699G>A | AATGGGAAAC | USF1 | Upstream stimulating factor 1 |
|            | AATGGGAAAC | ARNT | AhR nuclear translocator homodimers |
| g.146704699G>A | AATGGGAAAC | FOXA1 | Forkhead box protein A1, hepatocyte nuclear factor 3-alpha (HNF-3-alpha) |

Note: The SNPs in sequences are highlighted in red.