


Article

GWAS-Based Identification of New Loci for Milk Yield, Fat, and Protein in Holstein Cattle

Liyuan Liu ^{1,2,†}, Jinghang Zhou ^{1,2,†}, Chunpeng James Chen ^{2,†} , Juan Zhang ¹, Wan Wen ³, Jia Tian ³, Zhiwu Zhang ^{2,*}  and Yaling Gu ^{1,*}

¹ School of Agriculture, Ningxia University, Yinchuan 750021, Ningxia, China; lly000917@gmail.com (L.L.); zjhang027@gmail.com (J.Z.); zhangjkathy@126.com (J.Z.)

² Department of Crop and Soil Sciences, Washington State University, Pullman, Washington, DC 99164, USA; chun-peng.chen@wsu.edu

³ Animal Husbandry Workstation, Yinchuan 750001, Ningxia, China; 13709503296@163.com (W.W.); tianjia214@126.com (J.T.)

* Correspondence: zhiwu.zhang@wsu.edu (Z.Z.); guyl@nxu.edu.cn (Y.G.)

† These authors contributed equally to this work.

Received: 23 September 2020; Accepted: 3 November 2020; Published: 5 November 2020



Simple Summary: Understanding the genetic architecture underlying milk production traits in cattle is beneficial so that genetic variants can be targeted toward the genetic improvement. In this study, we performed a genome-wide association study for milk production and quality traits in Holstein cattle. In the total of ten significant single-nucleotide polymorphisms (SNPs) associated with milk fat and protein, six are located in previously reported quantitative traits locus (QTL) regions. The study not only identified the effect of *DGAT1* gene on milk fat and protein but also found several novel candidate genes. In addition, some pleiotropic SNPs and QTLs were identified that associated with more than two traits, these results could provide some basis for molecular breeding in dairy cattle.

Abstract: High-yield and high-quality of milk are the primary goals of dairy production. Understanding the genetic architecture underlying these milk-related traits is beneficial so that genetic variants can be targeted toward the genetic improvement. In this study, we measured five milk production and quality traits in Holstein cattle population from China. These traits included milk yield, fat, and protein. We used the estimated breeding values as dependent variables to conduct the genome-wide association studies (GWAS). Breeding values were estimated through pedigree relationships by using a linear mixed model. Genotyping was carried out on the individuals with phenotypes by using the Illumina BovineSNP150 BeadChip. The association analyses were conducted by using the fixed and random model Circulating Probability Unification (FarmCPU) method. A total of ten single-nucleotide polymorphisms (SNPs) were detected above the genome-wide significant threshold ($p < 4.0 \times 10^{-7}$), including six located in previously reported quantitative traits locus (QTL) regions. We found eight candidate genes within distances of 120 kb upstream or downstream to the associated SNPs. The study not only identified the effect of *DGAT1* gene on milk fat and protein, but also discovered novel genetic loci and candidate genes related to milk traits. These novel genetic loci would be an important basis for molecular breeding in dairy cattle.

Keywords: dairy; milk production; quality traits; *DGAT1*; FarmCPU

1. Introduction

Milk production and quality are the most important economic traits in the dairy industry. Most milk phenotypes are quantitative traits that often controlled by both environmental factors and

multiple genes. A large number of studies have revealed numerous quantitative traits locus (QTL) regions for milk-related traits in dairy cattle population around the world over the past 20 years (CattleQTLdb: <https://www.animalgenome.org/cgi-bin/QTLdb/BT/index>), and many researchers conducted meta-analysis to identify genetic variants based on GWAS results for milk-related traits in different cattle breeds [1–3]. Whereas, in Chinese Holstein population, a previous study used 50,000 single-nucleotide polymorphisms (SNPs) and revealed some SNPs associated with milk production traits [4]. The aim of present study was to find new genetic loci in this population by using a higher density marker information.

Research has shown that high-density genotype could provide markers close to the QTL and help in fine mapping of causative mutations [5]. Vanraden et al. reported that high-density marker increased the precision of QTL detection in cattle population [6]. In addition, a study reported that the genomic prediction accuracy increased when the marker density was increased in cattle [7]; therefore, it is necessary to use dense genotype to identify important genetic variation, and provide some useful information for molecular breeding of dairy cattle and understanding the genetic architecture of milk traits.

Genome-wide association studies (GWAS) are very helpful for further genomic selection (GS) as they have proven to be a powerful method for identifying potential genetic variants, especially single-nucleotide polymorphisms (SNPs) associated with complex traits in humans and animals [8–10]. Resende et al. suggested that the prediction accuracy reached maximum when the genomic relationship matrix was constructed using causative quantities trait nucleotides (QTNs) [11]. Incorporating significant markers from GWAS results can improve the prediction accuracy in dairy cattle [1,12–14].

In this study, we conducted GWAS using Illumina BovineSNP150 BeadChip which contains about 150,000 SNPs. The population was from Holstein cows raised in the Ningxia area of northwest China. Our objectives were to identify new genetic variants associated with five milk production and quality traits, including milk yield, fat and protein. We expect the newly identified genetic variants and potential candidate genes would become valuable resources for genetic evaluation.

2. Material and Methods

2.1. Population and Phenotypic Data

The studied population was Holstein cows that were raised on 22 dairy farms in the Ningxia area of China. In total, about 452,920 test-day records estimated breeding values from 61,600 cows spanning a 9-year period (2011–2019) at their first lactation. The estimated breeding values (EBVs) as phenotypes to implement association analysis, milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP), and protein percentage (PP) measurements were recorded once a month for each cow after calving. The milk yield was automatically recorded by the milking system on each farm, the milk components are tested by the Dairy Herd Improvement lab at Animal Husbandry of Extension Station in Ningxia, using spectrometers. FY was calculated as $(FP \times MY) / 100$; PY was calculated as $(PP \times MY) / 100$. The distribution of phenotypes and correlations between the different phenotypic traits are illustrated in Figure S1.

2.2. Estimated Breeding Values

The Derivative-free approach to MULTivariate analysis (DMU) package [15] was used to estimate breeding values using Random Regression Test-Day Model [16,17]. We considered herd-test day and calving year-season as fixed effects, calving month-age as fixed regression effect, and individual additive genetic effect and permanent environment effect as random regression effects. Both fixed and

random regressions were modeled using a 4th order Legendre polynomial [18]. The model equation is as follows:

$$y_{ijklm} = HTD_i + cays_j + \sum_{m=0}^4 b_{km}X_m(\omega) + \sum_{m=0}^4 a_{lm}X_m(\omega) + \sum_{m=0}^4 p_{lm}X_m(\omega) + e_{ijklm} \quad (1)$$

where y_{ijklm} is the test-day records; HTD_i is the fixed effect of the i th herd-test day ($i = 1, \dots, 1913$); $cays_j$ is the fixed effect of the j th calving year-season ($j = 1, \dots, 36$); b_{km} is fixed regression coefficient for the k th class of calving month-age ($k = 1, \dots, 8$); a_{lm} is random regression coefficient for additive genetic effects specific to cow l ; p_{lm} is random regression coefficient for permanent environment effects specific to cow l ; $X_m(\omega)$ is the m th covariate of Legendre polynomial; ω is the days of lactation after standardization; and e_{ijklm} is the random residual effects, hypothesizing the residuals are homogeneous. The variance–covariance matrix is as follows [19]:

$$Var \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} G \otimes A & 0 & 0 \\ 0 & I \otimes P & 0 \\ 0 & 0 & R \end{bmatrix} \quad (2)$$

where a is additive genetic random regression coefficient vector; p is permanent environment random regression coefficient vector; G is variance–covariance matrix of additive genetic random regression coefficient; A is numerator relationship matrix; P is variance–covariance matrix of permanent environment random regression coefficient; I is the identity matrix; R is diagonal matrix of residual variance ($I\sigma_e^2$), which hypothesizes the residuals are homogeneous. The homogeneous option dramatically reduces computing time without sacrifice as there is a minimal difference between the homogeneous model and the heterogenous model.

A heatmap of estimated breeding values for milk production traits is illustrated in Figure S2.

2.3. Genotypic Data

Blood samples from the 1220 cows were collected by cattle farm staff in this study. DNA was extracted and genotyping was carried out by Compass Biotechnology (<http://www.kangpusen.com/>) using the Illumina BovineSNP150 BeadChip. *Bos_taurus_UMD_3.1* as the genome reference. In total, there were 124,743 variants for the association analysis after conducting quality control using Plink software [20]. Markers were removed if (1) the call rate of an individual genotype was less than 95%, (2) the call rate of a single SNP genotype was less than 90%, and (3) if the minor allele frequency (MAF) of an SNP was less than 0.05 and deviated from Hardy–Weinberg equilibrium ($p < 1.0 \times 10^{-6}$). We calculated marker intervals and linkage disequilibrium (LD) to estimate R square for all markers and plotted the marker distribution as show in Figure 1.

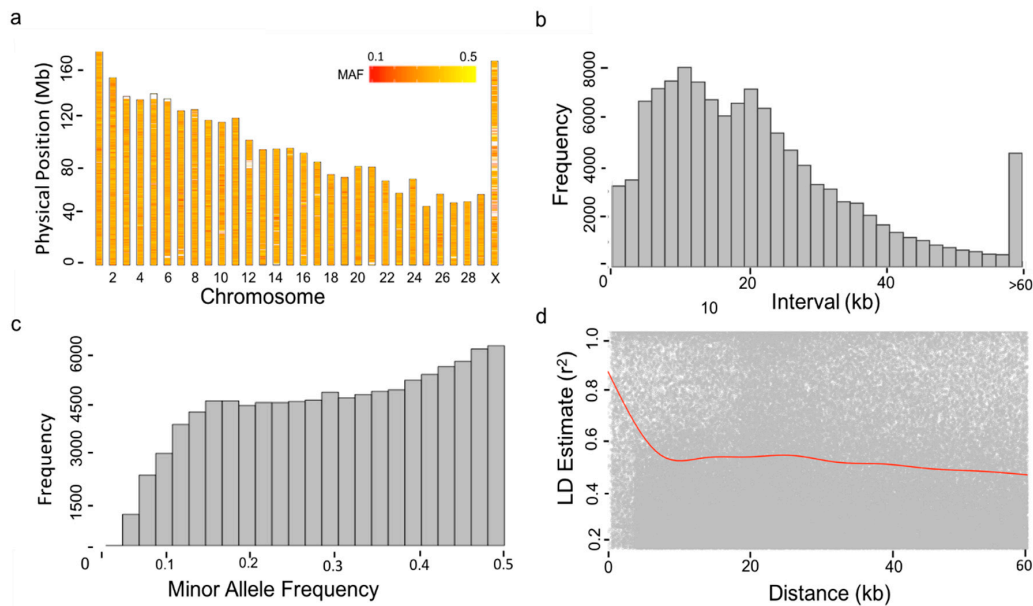


Figure 1. Properties of Single Nucleotide Polymorphisms (SNPs). A total of 1220 cows were genotyped by the Illumina Bovine 150 k BeadChip. After conducting quality control on both minor allele frequency (MAF, above 5%) and missing rate (<10%), 1220 individuals and 124,743 SNPs remained. The distribution of the filtered SNPs is displayed over the 30 bovine chromosomes except for Y chromosome (a). The MAFs of SNPs were re-calculated after the filtering and were displayed by a heat map. Consequently, the SNPs with MAF < 5% remained, as demonstrated by the histogram (c). The density of SNPs is displayed by the frequency of the distance between adjacent SNPs (b). The distances over 60 kb clustered into one group. The maximum distance was 100.06 kb. Pairwise Linkage Disequilibrium (LD) was calculated as the R square for SNPs within the 100 kb window. The decay of LD over distance (red line) is displayed by the pairwise LD and moving average (d).

2.4. Principal Component Analysis

Principal component analysis (PCA) was conducted using R function `Prcomp()` on 1220 cows genotyped with 124,743 markers covering the whole genome to study the population structure [21,22]. Most of these 1220 cows were the progeny of the frozen semen imported from multiple countries.

2.5. Association Analysis

We performed association analysis by a multi-locus linear mixed model using the FarmCPU (Fixed and random model Circuitous Probability Unification) [23]. FarmCPU method implements marker tests with associated markers as covariates in a fixed effect model and optimization on the associated covariate markers in a random effect model separately [23]. As we know, population stratification is an important factor that can cause false positive in association studies [22]. Therefore, the present study fitted the first three principal components (PCs) as covariate variables in the GWAS models [23], the fixed effect model is as follows:

$$y = Xb_X + M_t b_t + S_j d_j + e \quad (3)$$

where y is the EBVs of individual; X is a matrix of fixed effect for the first three PCs; M_t is the genotype matrix of t pseudo Quantitative Traits Nucleotides (QTNs), initiated as an empty set; b_X and b_t are the corresponding effects of X and M_t , respectively; S_j is the genotype of the j marker; d_j is the corresponding effect; e is the vector of residuals $e \sim N(0, I\sigma_e^2)$. The random effect model is as follows:

$$y = u + e \quad (4)$$

where y and e stay the same as in the fixed effect model; u is the genetic effect of the individual and $u \sim N(0, K\sigma_u^2)$, in which K is the kinship matrix derived from the pseudo QTNs.

The genome wise threshold corresponding type I error of 1% was 4.0×10^{-7} after Bonferroni multiple test correction (5%/124,743).

2.6. Annotation of Candidate Gene and Pathway Analysis

The genome reference Bos_taurus_UMD_3.1 was used to search candidate gene. The average pairwise LD was 0.46 corresponding to adjacent marker distance of 120 kb. This range was used to search candidate genes. The online websites "https://oct2018.archive.ensembl.org/Bos_taurus/Info/Index", "<https://www.ncbi.nlm.nih.gov/gene/>", <https://www.genome.jp/kegg/pathway.html>, <https://david.ncifcrf.gov/home.jsp> were used for functional analysis and pathway analysis of the candidate genes by GWAS.

3. Results

3.1. Phenotypic and Estimated Genetic Parameters

Phenotype distributions and correlations among phenotypic traits, estimated breeding values, and residuals are shown in the Supplementary Material (Figure S1). There were strong positive phenotypic correlations between "yield" type of traits, including MY, PY, and FY. Their phenotypic correlations were 0.90 (MY and PY), 0.70 (MY and FY), and 0.74 (PY and FY).

We also found strong positive genetic correlations between MY and PY ($r_g = 0.92$), MY and FY ($r_g = 0.84$), and PY and FY ($r_g = 0.88$). In contrast, there were weak negative phenotypic and genetic correlations between MY and FP ($r_p = -0.15$, $r_g = -0.32$), MY and PP ($r_p = -0.20$, $r_g = -0.44$).

In this study, we used EBVs as the dependent variables for GWAS, Figure S2 shows the heatmap of EBVs for five milk traits. A test-day model is used to estimate the heritability for each trait and breeding values for individuals. The heritability estimates for MY, FY, PY, FP, and PP are 0.12, 0.21, 0.23, 0.30, and 0.32, respectively (Table 1).

Table 1. Variance components and heritability of milk traits *.

| Variance Component | MY | FP | FY | PP | PY |
|-------------------------|------------|-------------|------------|------------|------------|
| Genetic | 1592.62 | 4.76 | 1.48 | 1.65 | 1.20 |
| Permanent environmental | 4267.93 | 10.58 | 5.46 | 3.45 | 3.96 |
| Residual | 7413.63 | 0.49 | 0.08 | 0.07 | 0.03 |
| h^2 (SE) | 0.12(0.01) | 0.30 (0.05) | 0.21(0.02) | 0.32(0.01) | 0.23(0.02) |

* MY, milk yield; FP, fat percentage; FY, fat yield; PP, protein percentage, PY, protein yield; SE, standard error.

3.2. Marker Information

We conducted GWAS analyses with 1220 Holstein dairy cows and 124,743 markers after quality control (QC). Markers covered all 29 autosomes plus the X sex chromosome (Figure 1a). After the QC filtering, we re-calculated the minor allele frequency (MAF) for all SNPs. The minimum MAF was 3.8%. There were only 0.1% of markers with MAF below 5% (Figure 1c). Marker density was high. Majority of markers (56%) are within 20 kb distances to their adjacent markers (Figure 1b). Within such distance, the LD was strong (average $R^2 = 0.46$) (Figure 1d).

3.3. Population Structure

To determine the level of population stratification, we plotted the population structure by principal component analysis (PCA). The population stratified into two unevenly sized groups (Figure 2). We also

produced a scatter plot of bull's country source (Figure S3). To adjust for the population stratification, the first three principal components (PCs) was fitted as covariate variables in the association analysis. The first three PCs explain the 1.6%, 1.3%, and 1.1% of variation, respectively, about 4% of the variation is explained by the first three PCs together. We also constructed a scatter plot between the first three PCs and the five milk traits. There are weak correlations observed between the PCs and these phenotypes (Figure S4).

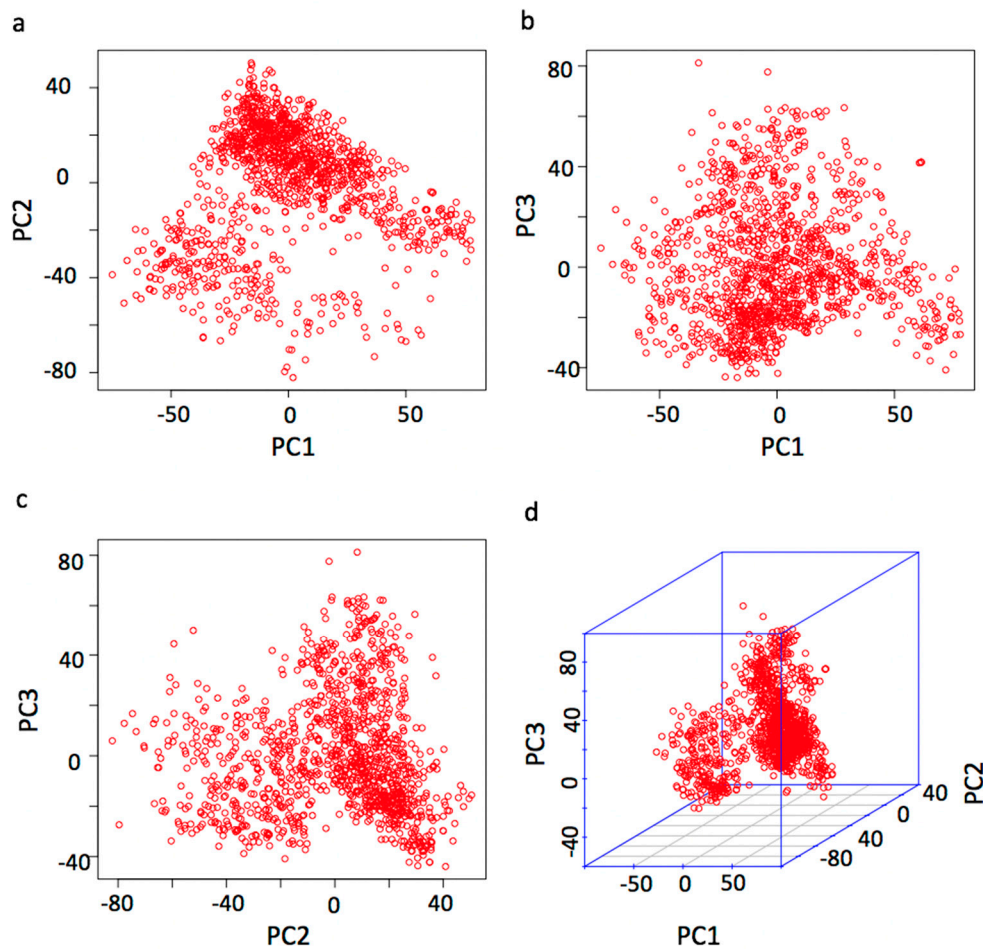


Figure 2. Population structure demonstrated by principal component analysis. Principal component analysis (PCA) was conducted with the 124,743 SNPs for the 1220 cows. The population structure is demonstrated by the pairwise scatter plots (a–c) and the 3D plot (d) of the first three principal components (PCs).

3.4. Results of the Genome-Wide Associations

By drawing the Quantile-Quantile (QQ) plots, we found that the model for GWAS analysis in this study was reasonable, and the point at the upper right corner also shown that some significant markers were found that associated with four milk quality traits (Figure 3). We used $p < 4.0 \times 10^{-7}$ as the threshold, which corresponds to 1% of type I error after Bonferroni multiple test correction. A total of ten highly significant SNPs are associated with fat and protein, but no threshold significant SNP is associated with MY (Table 2, Figure 3). Three SNPs (rs42295213, rs136949224, and rs109421300) associated with FP are located on BTA1, 8, and 14, four SNPs (rs43526055, rs137676276, rs109528658, and rs135780687) associated with FY are located on BTA7, 11, 17, and X, respectively. Three SNPs (rs109875012, rs109421300, and rs108996837) associated with PP are located on BTA5, 14, and 21, respectively. One SNP associated with PY is located on BTA5. Four of these ten significant SNPs are located inside genes EPH receptor A6 (*EPHA6*), solute carrier organic anion transporter family member

1A2 (*SLCO1A2*), diacylglycerol O-acyltransferase 1 (*DGAT1*), and E1A binding protein p400 (*EP400*). The SNP (rs109875012) on BTA5 is located close to the *ZNF384* (zinc finger protein 384). The SNP (rs10705865) on BTA8 is located close to *SCARA5* (scavenger receptor class A member 5) gene. The SNP (rs137676276) on BTA11 is located close to vitrin (*VIT*). The SNP (rs108996837) on BTA21 is located close to *EXOC3L4* (exocyst complex component 3 like 4), and the SNP (rs135780687) on X chromosome is located close to *GRPR* (gastrin releasing peptide receptor). The most significant SNP (rs109421300) associated with both FP and PP is located in the *DGAT1* gene. Two SNPs (rs137676276, rs108996837) exhibit notably smaller MAFs compared to other SNPs, 0.11 and 0.12, respectively (Table 2).

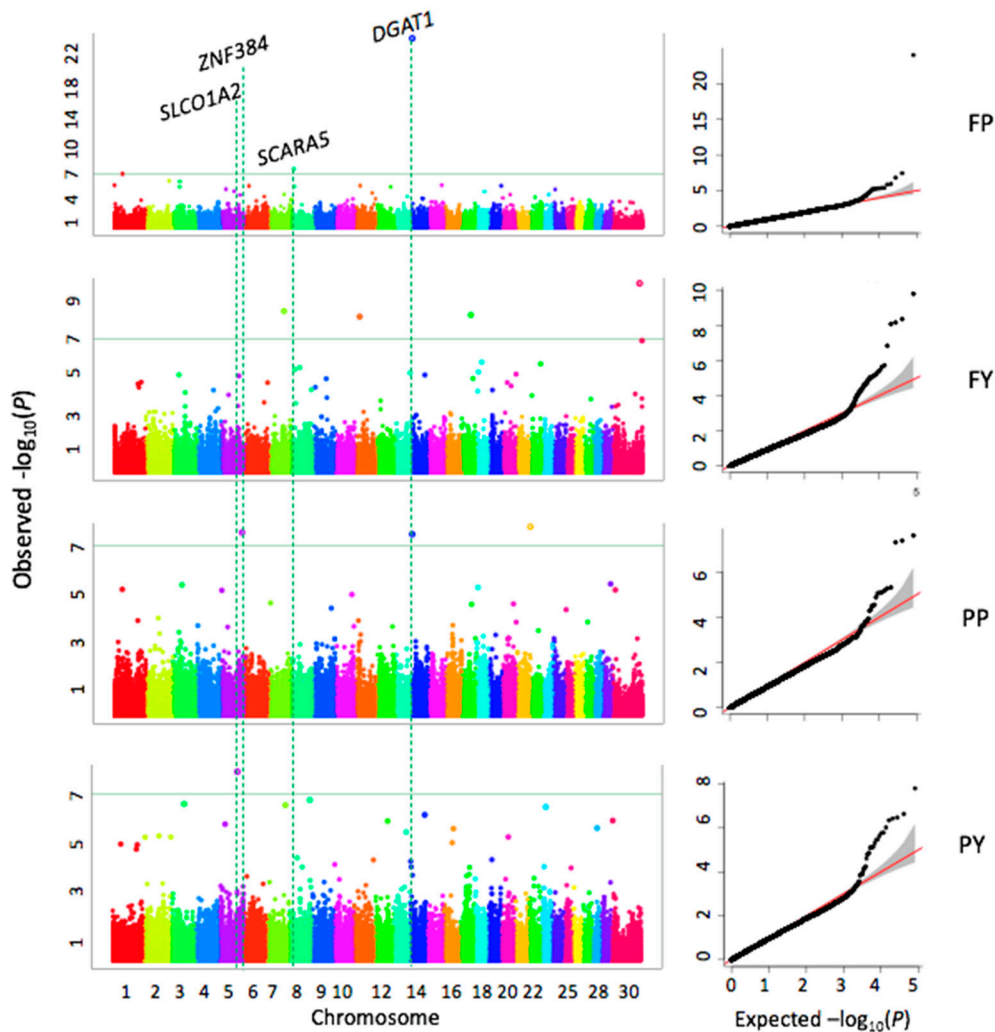


Figure 3. Associations between 124,743 SNPs and milk traits. Milk traits include fat yield (FY), protein yield (PY), fat percentage (FP), protein percentage (PP). The association analyses were conducted by the FarmCPU R package. Manhattan plots display the negative logarithms of the observed p values for SNPs across 30 bovine chromosomes (left panel). The green line indicates the Bonferroni multiple test threshold at $p = 4.0 \times 10^{-7}$. The Quantile-Quantile (QQ) plots represent the negative logarithms of the expected p values (X-axis) and observed p -values (Y-axis) (right panel).

Table 2. Genome-wide significant SNPs associated with milk traits *.

| Traits | SNP | CHR | Position (bp) | MAF | Nearest Gene | Distance (kb) | <i>p</i> -Value | Effect |
|--------|-------------|-----|---------------|------|----------------|---------------|------------------------|--------|
| FP | rs42295213 | 1 | 41,061,715 | 0.36 | <i>EPHA6</i> | within | 1.50×10^{-7} | 0.007 |
| PP | rs109875012 | 5 | 104,120,905 | 0.45 | <i>ZNF384</i> | 2.6 | 4.03×10^{-8} | −0.004 |
| PY | rs134480235 | 5 | 89,267,320 | 0.49 | <i>SLCO1A2</i> | within | 1.57×10^{-8} | 1.192 |
| FY | rs43526055 | 7 | 73,431,219 | 0.31 | <i>ADRA1B</i> | 179.8 | 4.48×10^{-9} | −1.523 |
| FP | rs136949224 | 8 | 10,705,865 | 0.14 | <i>SCARA5</i> | 43.5 | 3.57×10^{-8} | 0.012 |
| FY | rs137676276 | 11 | 19,277,448 | 0.11 | <i>VIT</i> | 25 | 8.58×10^{-9} | −2.281 |
| FP | rs109421300 | 14 | 1,801,116 | 0.23 | <i>DGAT1</i> | within | 9.92×10^{-25} | 0.018 |
| PP | | | | | | | 4.75×10^{-8} | 0.006 |
| FY | rs109528658 | 17 | 46,090,458 | 0.40 | <i>EP400</i> | within | 7.05×10^{-9} | 1.543 |
| PP | rs108996837 | 21 | 69,386,346 | 0.12 | <i>EXOC3L4</i> | 32.6 | 2.36×10^{-8} | −0.008 |
| FY | rs135780687 | X | 134,726,985 | 0.42 | <i>GRPR</i> | 83 | 1.63×10^{-10} | 1.629 |

* FP, fat percentage (%); PP, protein percentage (%); PY, protein yield (kg); FY, fat yield (kg).

3.5. Pleiotropic QTLs for Milk Production Traits

We used the SNPs with $p < 0.0005$ to make a heatmap to look for markers associated with two or more milk production traits because these milk traits are moderately or highly correlated (Figures S1, S2, and S5). There are some SNPs and QTLs associated with three milk traits (MY, PY, and FY) on BTA1, 2, 3, 11, 17, 20, and 22. QTLs associated with MY while PY on BTA1, 6, 14, and 17. A QTL associated with FY and FY is on BTA20.

4. Discussion

4.1. Population Structure

Population stratification is an important confounding factor due to systematic ancestry differences that can cause false positives in GWAS [24]. By the principal component analysis, the PCA scatter plot showed that there is population structure in this studied population (Figure 2 and Figure S3). Two probably reasons for this deviation, the first reason is that those Holstein semen from overseas is not used by all cows in dairy farms, and there are still some local Holstein semen used by cows, as we know most of the farms in this study participated in a dairy breeding project that introduced Holstein semen from overseas annually from 2013 to 2018. Another reason is that some cows are introduced from different countries and contain blood from other breeds. In general, not all the registered cattle are purebreds. Especially in cattle population, this point can be explained by looking at the cattle breed registration requirements in different countries. Here, we take the Holstein cattle as an example, one of the requirements to register a Holstein cattle in China is that the cattle at least has 87.5% blood of Holstein (Chinese Holstein, GB/T 3157 2008), all the animals with Holstein genetics can be registered in Canada (https://www.holstein.ca/Public/en/Services/Registration/Registration_Eligibilities) and there are similar clauses in the USA (http://www.holsteinusa.com/animal_id/register.html). According to the above standards, we can see the common ground to register a Holstein cattle between different countries is that not all the Holstein cattle are purebreds. Even most of the Holstein cattle are purebreds, some of registered cattle could still contain a little other blood in the long-term breeding progress. That is why the population structure analysis is necessary in this study.

As we observed the population structure, the principal components were fitted as covariance to association analysis to correct population stratification. After adjusted PC factors, there are four SNPs overlapped with the association model and not fitted PCs—results will be discussed in a later section.

4.2. GWAS for Milk Traits

Milk production and quality are important economic objectives in the dairy industry, good milk production performance can bring greater economic benefits because of the milk pricing system and

production efficiency. Most of milk phenotypes are quantitative traits and are regulated by polygenic, research on the relationship between genetics and milk traits began decades ago. As early as 1994, a research identified a QTL significantly associated with FY was linked to kappa-casein and a QTL for PY was linked to beta-lactoglobulin [25]. Subsequently, a growing number of studies detected tens and thousands of QTLs through the 30 chromosomes associated with 653 different traits in cattle (Cattle QTLdb) [26]. Even though there has been plenty of research in this field, the available evidence is still not enough to give a completed explanation of genetic mechanism for these traits, and more new research samples will still be valuable to help lay groundwork in this aspect. Therefore, 1220 Holstein cows with comprehensive herd-test data were genotyped for GWAS in this study. In the results, we found ten SNPs were significantly associated with four milk quality traits (FP, FY, PP, and PY), one of the most significant SNPs was located in the *DGAT1* gene and shown closely related to milk fat and protein percentage. No significant SNP passed the Bonferroni correction threshold for MY, the small number of markers detected may be due to the limited sample size in the present study, which is a critical factor limiting the statistical power.

A random regression test-day model was used to estimate breeding values in this study, which eliminated environmental factors (herd-test-day, calving year-season and calving month-age), and then EBV was used as a dependent variable to conduct association tests, whereas a study used the deregressed EBV as phenotypic records for the bulls to estimate SNP effects by using a single-marker regression model [27]. The deregressed EBV was proposed by Garrick et al., in which removing the parental average effects is more valuable in genomic analysis [28]. The phenotypes used by association analysis can be varied, raw phenotypes, adjusted phenotypes, EBVs, deregressed EBVs, and daughter yield deviation (DYD) [29–34], and certainly different phenotypes are suitable for different scenarios. In dairy breeds, EBVs or deregressed EBVs are preferred as the dependent variable of GWAS. Research has shown that the deregressed EBVs could be reliable in genomic analysis, but some studies showed that the accuracy of genomic analysis when using EBVs was only slightly lower than using deregressed EBVs [35,36]. In addition, a simulation study indicated that GWAS using EBVs or deregressed EBVs as the dependent variable with polygenic effect modeled had similar performance in controlling the false positive rates (FPR), even when using deregressed EBVs lowered the power in some degree [37]. Therefore, the present study prefers to use EBVs for analysis.

In this study, the three SNPs on BTA1, 8, 14 that were associated with FP were found, which are within reported QTL regions [38–40]. The study used *Bos_taurus_UMD_3.1* as a reference genome to search the candidate genes at a distance of 120 kb upstream or downstream of the associated SNPs. The SNP (0.007%; $p = 1.50 \times 10^{-7}$) associated with an increase of FP on BTA1 is located in the *EPHA6* gene, which has functions of ATP binding (GO:0005524), protein binding (GO: 0005515), protein tyrosine kinase activity (GO: 0004713), and has been proposed to participate in Axon guidance pathway (KEGG: bta04360). The SNP (0.012%; $p = 3.57 \times 10^{-8}$) on BTA8 located close to *SCARA5* gene resulted in an increase of FP. This gene is a member of the scavenger receptor (SR) family, which is broadly expressed in fat tissue in humans. Research showed that scavenger receptors involved in lipid accumulation and inflammation [41]. Some studies reported the gene plays a critical role in progression and metastasis of breast cancer [42] and is involved in breast carcinogenesis [43]. The SNP on BTA14 associated both with an increase of FP (0.018%, $p = 9.92 \times 10^{-25}$) and PP (0.006%; $p = 4.75 \times 10^{-8}$), which is in the *DGAT1* gene. As we know, *DGAT1* gene is widely reported associated with milk yield and composition, especially K232A polymorphism affecting on milk fat and protein [38,44,45]. In this study, the SNP (rs109421300) we identified is within an intron of *DGAT1* gene and is the most significantly affecting FP and also affecting PP, and a lot of studies have reported for milk yield and composition in Holstein, Jersey, and Holstein–Friesian cattle [46–48]. *DGAT1* gene as a key metabolic enzyme catalyzes the biosynthesis of triacylglycerols [49], and glycerolipid metabolism (KEGG: bta00561), retinol metabolism (KEGG: bta00830), metabolic pathways (KEGG: bta01100), and fat digestion and absorption (KEGG: bta04975) by KEGG pathway analysis (https://www.genome.jp/dbget-bin/www_bget?bta:282609). The SNP (−0.004%; $p = 4.03 \times 10^{-8}$) on BTA5 is associated with a reduction of PP and is within previously

reported milk fatty acid content QTL [39]. This SNP is close to the *ZNF384*, which encodes a C2H2-type zinc finger protein. It may function as a transcription factor and have DNA binding and metal ion binding functions. Another SNP (-0.008% ; $p = 2.36 \times 10^{-8}$) associated with a reduction of PP on BTA21 at 69Mb is located near the *EXOC3L4* gene. Four SNPs are associated with FY in the present study. The SNP (-1.523 kg; $p = 4.48 \times 10^{-9}$) on BTA7 near the adrenoceptor alpha 1B (*ADRA1B*) gene resulted in a reduction of FY, which is participated in the calcium signaling pathway (KEGG: bta04020), cGMP-PKG signaling pathway (KEGG: bta04022), and neuroactive ligand–receptor interaction (KEGG: bta04080). The SNP (-2.281 kg; $p = 8.58 \times 10^{-9}$) associated with a reduction of FY on BTA11 at 19Mb located near the *VIT* gene. This gene encodes an extracellular matrix (ECM) protein and has been suggested to contribute to normal brain asymmetry variation [50]. The SNP (1.543 kg; $p = 7.05 \times 10^{-9}$) on BTA17 is located in the *EP400* and resulted in an increase of FY, and also within previously reported QTL for milk protein composition [51], *EP400* gene participates in histone H2A acetylation, histone H4 acetylation, and has the function of ATP binding, DNA binding, protein binding, and helicase activity. The SNP (1.629 kg; $p = 1.63 \times 10^{-10}$) associated with an increase of FY on chromosome X is located nearby the *GRPR*. This gene is a gastrin releasing peptide receptor. Gastrin-releasing peptide regulates numerous functions of the gastrointestinal and central nervous system, which is involved in calcium signaling pathway (KEGG: bta04020) and neuroactive ligand–receptor interaction (KEGG: bta04080). These two pathways were enriched for FY in this study. A study revealed that the calcium signaling pathway was related to milk coagulation properties and curd nutrient recovery traits in dairy cattle [52], and interestingly, another study showed that these two pathways were significantly correlated with lactation performance in mice [53]. Song et al. reviewed a lot of studies on calcium signaling pathway that participated in the effect of the sympathetic nerve in regulating adipose metabolism [54]. We suspect that *ADRA1B* and *GRPR* genes may affect milk fat through the calcium signaling and neuroactive ligand–receptor interaction pathways. It is well known that the biological mechanisms of quantitative traits are very complicated, and the present study is only based on SNP data analysis. Some studies have shown that copy number variation and DNA methylation are also related to milk production and quality traits [55–57]. The SNP (1.192 kg; $p = 1.57 \times 10^{-8}$) on BTA5 is associated with an increase of PY and it is also within the previously reported QTL region, which is associated with milk fatty acid content. This SNP is in the *SLCO1A2* gene, which is encoding solute carrier anion transporter family, member 1A2, the gene participating in the digestive system, organic anion transport process, and bile secretion pathway (KEGG: 04976). Furthermore, when the association model used not fitted PCs as covariance, four SNPs are overlapped with the model with fitted PCs: rs109875012 (*ZNF384*), rs136949224 (*SCARA5*), rs109421300 (*DGAT1*), and rs109528658 (*EP400*), it is suggested that these significant SNPs could be more stable and reliable.

4.3. Correlations among Milk Traits and Pleiotropic QTLs

Estimates of heritability and genetic correlations are essential population genetic parameters in animal breeding research and application of animal breeding programs [58]. Genetic correlation can be useful for indirect selection, selection in different environment, and selecting multiple traits simultaneously. We found there was relatively high genetic correlation between milk yield and protein yield, and it is consistent with the results in Holsteins [59] and Jersey cattle [58]. We expect that there are overlapping regions of genetic variation between different traits that are relatively high correlation. A heatmap using p -values from GWAS results helped identify pleiotropic QTLs (Figure S5). The QTL on BTA2 at 134Mb is associated with MY, PY, and FY, and also reported associated with the milk composition in another Holstein study [60]. The MY- and PY-associated QTL is adjacent to the *ABCG2* gene on BTA6; previously, studies reported that a missense mutation in the *ABCG2* gene is associated with milk yield and composition in Holstein, Braunvieh, and Fleckvieh cattle [2,61], an intron variant affecting milk fatty acids in Chinese Holstein [62], and has also been reported to affect body weight, calving ease direct in US cattle breed [63,64]. The QTL on BTA17 is associated with MY, PY, and FY, with previous studies indicating that QTLs in this region are related to milk fatty acid

and body weight [60,65]. A previous study reported a QTL on BTA20 is associated with PP and body weight [63,66]. In this study, we found this QTL is associated with MY, PY, and FY. The QTL on BTA24 is pleiotropic and is associated with MY and FY, and previously reported for milk fatty acid [67] and conception rate in Holstein cows [68].

5. Conclusions

The study performed genome-wide association analysis using test-day records for five milk production and quality traits in Holstein cows. A total of ten significant SNPs associated with milk fat and protein percentage, fat and protein yield were found, six of them located within previously reported QTLs, including *DGAT1* gene on milk quality. We also found some new genetic loci and candidate genes related to milk fat and protein. In addition, some SNPs and QTLs associated with more than two milk traits were found. These results could provide some basis for molecular breeding and useful information to understanding the genetic architecture of milk production and quality traits in dairy cattle.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-2615/10/11/2048/s1>, Figure S1: Phenotype distributions and correlations among milk traits, Figure S2: Heatmap of estimated breeding values for milk traits, Figure S3: The relationship between origins and the first two principle components, Figure S4: Correlation between phenotypic values of milk traits and the first three principle components, Figure S5: Heatmap of significant markers associate with five milk traits.

Author Contributions: Conceived experiment: Z.Z. and Y.G.; data analyses: L.L., J.Z. (Jinghang Zhou), and J.C.; data collection: J.Z. (Juan Zhang), J.T., and W.W.; wrote manuscript: L.L., J.Z. (Jinghang Zhou), J.C., Z.Z., and Y.G.; Funding acquisition, C.J.C. and W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project of High-yield and High-quality Dairy Cattle Breeding (2019NYZ05), the High-Level Academic Papers in Ningxia University, the National Science Foundation (Award # DBI 1661348), and the USDA National Institute of Food and Agriculture Hatch project (1014919).

Acknowledgments: The authors would like to thank the DHI measurement center of Ningxia Animal Husbandry of Extension Station and 22 dairy farms for their phenotypic data. We also thank Linda R. Klein for valuable writing advice and editing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Ethics Statement: No applicable. This study only did statistical analysis based on an existing database on the Holstein breeding project in Ningxia province, China. All the authors in this study didn't participated in any sample collection process. All the farms involved were consent and agree to take part in this research.

References

1. Teissier, M.; Sanchez, M.P.; Boussaha, M.; Barbat, A.; Hoze, C.; Robert-Granie, C.; Croiseau, P. Use of meta-analyses and joint analyses to select variants in whole genome sequences for genomic evaluation: An application in milk production of French dairy cattle breeds. *J. Dairy Sci.* **2018**, *101*, 3126–3139. [[CrossRef](#)] [[PubMed](#)]
2. Pausch, H.; Emmerling, R.; Gredler-Grandl, B.; Fries, R.; Daetwyler, H.D.; Goddard, M.E. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. *BMC Genom.* **2017**, *18*, 853. [[CrossRef](#)]
3. Marete, A.G.; Guldbrandtsen, B.; Lund, M.S.; Fritz, S.; Sahana, G.; Boichard, D. A Meta-Analysis Including Pre-selected Sequence Variants Associated With Seven Traits in Three French Dairy Cattle Populations. *Front. Genet.* **2018**, *9*, 522. [[CrossRef](#)]
4. Jiang, L.; Liu, J.; Sun, D.; Ma, P.; Ding, X.; Yu, Y.; Zhang, Q. Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. *PLoS ONE* **2010**, *5*, e13661. [[CrossRef](#)]
5. Pryce, J.E.; Johnston, J.; Hayes, B.J.; Sahana, G.; Weigel, K.A.; McParland, S.; Spurlock, D.; Krattenmacher, N.; Spelman, R.J.; Wall, E.; et al. Imputation of genotypes from low density (50,000 markers) to high density (700,000 markers) of cows from research herds in Europe, North America, and Australasia using 2 reference populations. *J. Dairy Sci.* **2014**, *97*, 1799–1811. [[CrossRef](#)]

6. VanRaden, P.M.; Null, D.J.; Sargolzaei, M.; Wiggans, G.R.; Tooker, M.E.; Cole, J.B.; Sonstegard, T.S.; Connor, E.E.; Winters, M.; van Kaam, J.B.C.H.M.; et al. Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* **2013**, *96*, 668–678. [[CrossRef](#)] [[PubMed](#)]
7. Pryce, J.E.; Arias, J.; Bowman, P.J.; Davis, S.R.; Macdonald, K.A.; Waghorn, G.C.; Wales, W.J.; Williams, Y.J.; Spelman, R.J.; Hayes, B.J. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *J. Dairy Sci.* **2012**, *95*, 2108–2119. [[CrossRef](#)]
8. Kristensen, V.N.; Børresen-Dale, A.-L. SNPs associated with molecular subtypes of breast cancer: On the usefulness of stratified Genome-wide Association Studies (GWAS) in the identification of novel susceptibility loci. *Mol. Oncol.* **2008**, *2*, 12–15. [[CrossRef](#)]
9. Yang, W.; Gu, C.C. A whole-genome simulator capable of modeling high-order epistasis for complex disease. *Genet. Epidemiol.* **2013**, *37*, 686–694. [[CrossRef](#)]
10. Hayes, B.; Goddard, M. Genome-wide association and genomic selection in animal breeding This article is one of a selection of papers from the conference “Exploiting Genome-wide Association in Oilseed Brassicas: A model for genetic improvement of major OECD crops for sustainable farming”. *Genome* **2010**, *53*, 876–883. [[CrossRef](#)]
11. Fragomeni, B.O.; Lourenco, D.A.L.; Masuda, Y.; Legarra, A.; Misztal, I. Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genet. Sel. Evol.* **2017**, *49*, 59. [[CrossRef](#)] [[PubMed](#)]
12. Lopes, M.S.; Bovenhuis, H.; van Son, M.; Nordbø, Ø.; Grindflek, E.H.; Knol, E.F.; Bastiaansen, J.W.M. Using markers with large effect in genetic and genomic predictions. *J. Anim. Sci.* **2017**, *95*, 59. [[CrossRef](#)]
13. Zhang, Z.; Ober, U.; Erbe, M.; Zhang, H.; Gao, N.; He, J.; Li, J.; Simianer, H. Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *PLoS ONE* **2014**, *9*, e93017. [[CrossRef](#)] [[PubMed](#)]
14. Raymond, B.; Bouwman, A.C.; Schrooten, C.; Houwing-Duistermaat, J.; Veerkamp, R.F. Utility of whole-genome sequence data for across-breed genomic prediction. *Genet. Sel. Evol.* **2018**, *50*, 27. [[CrossRef](#)]
15. Madsen, P.; Jensen, J. A Package for Analysing Multivariate Mixed Models. Version 6, Release 5, 31. 2006. Available online: <http://www.wcgalp.org/system/files/proceedings/2010/dmu-package-analyzing-multivariate-mixed-models.pdf#:~:text=DMU%20-%20A%20Package%20For%20Analyzing%20Multivariate%20Mixed,%28BLUE%29%20and%20to%20predict%20ran-%20dom%20effects%20%28BLUP%29> (accessed on 24 September 2020).
16. Schaeffer, L.R. Application of random regression models in animal breeding. *Livest. Prod. Sci.* **2004**, *86*, 35–45. [[CrossRef](#)]
17. Naserkheil, M.; Miraie-Ashtiani, S.R.; Nejati-Javaremi, A.; Son, J.; Lee, D. Random Regression Models Using Legendre Polynomials to Estimate Genetic Parameters for Test-day Milk Protein Yields in Iranian Holstein Dairy Cattle. *Asian Australas. J. Anim. Sci.* **2016**, *29*, 1682–1687. [[CrossRef](#)]
18. Ren, X. Establishment of Genetic Evaluation System for Lactation Performance of Chinese Holstein Cattle in Ningxia. Master’s Thesis, China Agricultural University, Beijing, China, 2015. (In Chinese).
19. Padilha, A.H.; Cobuci, J.A.; Costa, C.N.; Neto, J.B. Random Regression Models Are Suitable to Substitute the Traditional 305-Day Lactation Model in Genetic Evaluations of Holstein Cattle in Brazil. *Asian Australas. J. Anim. Sci.* **2015**, *29*, 759–767. [[CrossRef](#)]
20. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)] [[PubMed](#)]
21. Yao, Z. Principal Component Methods in R: Practical Guide Principle Component Methods in R: Practical Guide. 2017. Available online: <http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/> (accessed on 24 September 2020).
22. Reich, D.; Price, A.L.; Patterson, N. Principal component analysis of genetic data. *Nat. Genet.* **2008**, *40*, 491–492. [[CrossRef](#)] [[PubMed](#)]
23. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet.* **2016**, *12*, e1005767. [[CrossRef](#)]
24. Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **2006**, *38*, 904–909. [[CrossRef](#)]

25. Bovenhuis, H.; Weller, J.I. Mapping and Analysis of Dairy Cattle Quantitative Trait Loci by Maximum Likelihood Methodology Using Milk Protein Genes as Genetic Markers. *Genetics* **1994**, *137*, 267–280.
26. Hu, Z.-L.; Park, C.A.; Reecy, J.M. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* **2019**, *47*, D701–D710. [[CrossRef](#)]
27. Viale, E.; Tiezzi, F.; Maretto, F.; De Marchi, M.; Penasa, M.; Cassandro, M. Association of candidate gene polymorphisms with milk technological traits, yield, composition, and somatic cell score in Italian Holstein-Friesian sires. *J. Dairy Sci.* **2017**, *100*, 7271–7281. [[CrossRef](#)] [[PubMed](#)]
28. Garrick, D.J.; Taylor, J.F.; Fernando, R.L. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* **2009**, *41*, 55. [[CrossRef](#)] [[PubMed](#)]
29. Jiang, J.; Ma, L.; Prakapenka, D.; VanRaden, P.M.; Cole, J.B.; Da, Y. A large-scale genome-wide association study in US Holstein cattle. *Front. Genet.* **2019**, *10*, 412. [[CrossRef](#)]
30. Yin, T.; König, S. Genome-wide associations and detection of potential candidate genes for direct genetic and maternal genetic effects influencing dairy cattle body weight at different ages. *Genet. Sel. Evol.* **2019**, *51*, 4. [[CrossRef](#)] [[PubMed](#)]
31. Zhang, F.; Wang, Y.; Mukiibi, R.; Chen, L.; Vinsky, M.; Plastow, G.; Basarab, J.; Stothard, P.; Li, C. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: Feed efficiency and component traits. *BMC Genom.* **2020**, *21*, 36. [[CrossRef](#)] [[PubMed](#)]
32. Zhou, C.; Li, C.; Cai, W.; Liu, S.; Yin, H.; Shi, S.; Zhang, Q.; Zhang, S. Genome-wide association study for milk protein composition traits in a Chinese Holstein population using a single-step approach. *Front. Genet.* **2019**, *10*, 72. [[CrossRef](#)] [[PubMed](#)]
33. Lu, H.; Wang, Y.; Bovenhuis, H. Genome-wide association study for genotype by lactation stage interaction of milk production traits in dairy cattle. *J. Dairy Sci.* **2020**. [[CrossRef](#)] [[PubMed](#)]
34. Sermyagin, A.A.; Gladyr, E.A.; Plemashov, K.V.; Kudinov, A.A.; Dotsev, A.V.; Deniskova, T.E.; Zinovieva, N.A. Genome-wide association studies for milk production traits in Russian population of Holstein and black-and-white cattle. In *Proceedings of the Scientific-Practical Conference “Research and Development-2016”, Moscow, Russia, 14–15 December 2016*; Springer: Cham, Switzerland, 2018; pp. 591–599. [[CrossRef](#)]
35. Ramírez-Flores, F.; López-Ordaz, R.; Domínguez-Viveros, J.; García-Muñiz, J.G.; Ruíz-Flores, A. Accuracy of genomic values predicted using deregressed predicted breeding values as response variables. *Rev. Mex. Cienc. Pecu.* **2017**, *8*, 445–451. [[CrossRef](#)]
36. Guo, G.; Lund, M.S.; Zhang, Y.; Su, G. Comparison between genomic predictions using daughter yield deviation and conventional estimated breeding value as response variables. *J. Anim. Breed. Genet.* **2010**, *127*, 423–432. [[CrossRef](#)] [[PubMed](#)]
37. Ning, C.; Kang, H.; Zhou, L.; Wang, D.; Wang, H.; Wang, A.; Fu, J.; Zhang, S.; Liu, J. Performance gains in genome-wide association studies for longitudinal traits via modeling time-varied effects. *Sci. Rep.* **2017**, *7*, 590. [[CrossRef](#)] [[PubMed](#)]
38. Kühn, C.; Thaller, G.; Winter, A.; Bininda-Emonds, O.R.P.; Kaupe, B.; Erhardt, G.; Bennewitz, J.; Schwerin, M.; Fries, R. Evidence for Multiple Alleles at the *DGAT1* Locus Better Explains a Quantitative Trait Locus With Major Effect on Milk Fat Content in Cattle. *Genetics* **2004**, *167*, 1873–1881. [[CrossRef](#)]
39. Bouwman, A.C.; Bovenhuis, H.; Visker, M.H.; van Arendonk, J.A. Genome-wide association of milk fatty acids in Dutch dairy cattle. *BMC Genet.* **2011**, *12*, 43. [[CrossRef](#)]
40. Sanchez, M.P.; Govignon-Gion, A.; Ferrand, M.; Gelé, M.; Pourchet, D.; Amigues, Y.; Fritz, S.; Boussaha, M.; Capitan, A.; Rocha, D.; et al. Whole-genome scan to detect quantitative trait loci associated with milk protein composition in 3 French dairy cattle breeds. *J. Dairy Sci.* **2016**, *99*, 8203–8215. [[CrossRef](#)] [[PubMed](#)]
41. Nayeri, S.; Sargolzaei, M.; Abo-Ismael, M.K.; May, N.; Miller, S.P.; Schenkel, F.; Moore, S.S.; Stothard, P. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet.* **2016**, *17*, 75. [[CrossRef](#)]
42. You, K.; Su, F.; Liu, L.; Lv, X.; Zhang, J.; Zhang, Y.; Liu, B. SCARA5 plays a critical role in the progression and metastasis of breast cancer by inactivating the ERK1/2, STAT3, and AKT signaling pathways. *Mol. Cell. Biochem.* **2017**, *435*, 47–58. [[CrossRef](#)]
43. Ulker, D.; Ersoy, Y.E.; Gucin, Z.; Muslumanoglu, M.; Buyru, N. Downregulation of SCARA5 may contribute to breast cancer via promoter hypermethylation. *Gene* **2018**, *673*, 102–106. [[CrossRef](#)] [[PubMed](#)]

44. Grisart, B.; Farnir, F.; Karim, L.; Cambisano, N.; Kim, J.J.; Kvasz, A.; Mni, M.; Simon, P.; Frere, J.-M.; Coppieters, W.; et al. 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* **2004**, *101*, 2398–2403. [[CrossRef](#)]
45. Bovenhuis, H.; Visker, M.H.P.W.; Poulsen, N.A.; Sehested, J.; van Valenberg, H.J.F.; van Arendonk, J.A.M.; Larsen, L.B.; Buitenhuis, A.J. Effects of the diacylglycerol o-acyltransferase 1 (DGAT1) K232A polymorphism on fatty acid, protein, and mineral composition of dairy cattle milk. *J. Dairy Sci.* **2016**, *99*, 3113–3123. [[CrossRef](#)]
46. Meredith, B.K.; Kearney, F.J.; Finlay, E.K.; Bradley, D.G.; Fahey, A.G.; Berry, D.P.; Lynn, D.J. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genet.* **2012**, *13*, 21. [[CrossRef](#)]
47. Aliloo, H.; Pryce, J.E.; González-Recio, O.; Cocks, B.G.; Hayes, B.J. Validation of markers with non-additive effects on milk yield and fertility in Holstein and Jersey cows. *BMC Genet.* **2015**, *16*, 89. [[CrossRef](#)]
48. Frańczak, M.; Szyda, J. Comparison of significant single nucleotide polymorphisms selections in GWAS for complex traits. *J. Appl. Genet.* **2016**, *57*, 207–213. [[CrossRef](#)]
49. Yen, C.-L.E.; Stone, S.J.; Koliwad, S.; Harris, C.; Farese, R.V. Thematic Review Series: Glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *J. Lipid Res.* **2008**, *49*, 2283–2301. [[CrossRef](#)]
50. Tadayon, S.H.; Vaziri-Pashkam, M.; Kahali, P.; Ansari Dezfouli, M.; Abbassian, A. Common Genetic Variant in VIT Is Associated with Human Brain Asymmetry. *Front. Hum. Neurosci.* **2016**, *10*. [[CrossRef](#)]
51. Schopen, G.C.B.; Koks, P.D.; van Arendonk, J.A.M.; Bovenhuis, H.; Visker, M.H.P.W. Whole genome scan to detect quantitative trait loci for bovine milk protein composition. *Anim. Genet.* **2009**, *40*, 524–537. [[CrossRef](#)]
52. Dadousis, C.; Pegolo, S.; Rosa, G.J.M.; Gianola, D.; Bittante, G.; Cecchinato, A. Pathway-based genome-wide association analysis of milk coagulation properties, curd firmness, cheese yield, and curd nutrient recovery in dairy cattle. *J. Dairy Sci.* **2017**, *100*, 1223–1231. [[CrossRef](#)]
53. Wei, J.; Ramanathan, P.; Martin, I.C.; Moran, C.; Taylor, R.M.; Williamson, P. Identification of gene sets and pathways associated with lactation performance in mice. *Physiol. Genom.* **2013**, *45*, 171–181. [[CrossRef](#)]
54. Song, Z.; Wang, Y.; Zhang, F.; Yao, F.; Sun, C. Calcium Signaling Pathways: Key Pathways in the Regulation of Obesity. *IJMS* **2019**, *20*, 2768. [[CrossRef](#)]
55. Wen, Y.; He, H.; Liu, H.; An, Q.; Wang, D.; Ding, X.; Shi, Q.; Feng, Y.; Wang, E.; Lei, C.; et al. Copy number variation of the USP16 gene and its association with milk traits in Chinese Holstein cattle. *Anim. Biotechnol.* **2020**. [[CrossRef](#)]
56. Di Gerlando, R.; Sutera, A.M.; Mastrangelo, S.; Tolone, M.; Portolano, B.; Sottile, G.; Bagnato, A.; Strillacci, M.G.; Sardina, M.T. Genome-wide association study between CNVs and milk production traits in Valle del Belice sheep. *PLoS ONE* **2019**, *14*, e0215204. [[CrossRef](#)]
57. Zhao, H.; Zhang, S.; Wu, X.; Pan, C.; Li, X.; Lei, C.; Chen, H.; Lan, X. DNA methylation pattern of the goat PITX1 gene and its effects on milk performance. *Arch. Anim. Breed.* **2019**, *62*, 59–68. [[CrossRef](#)]
58. Missanjo, E.; Imbayarwo-Chikosi, V.; Halimani, T. Estimation of Genetic and Phenotypic Parameters for Production Traits and Somatic Cell Count for Jersey Dairy Cattle in Zimbabwe. *ISRN Vet. Sci.* **2013**, *2013*, 470585. [[CrossRef](#)]
59. Zaabza, H.B.; Gara, A.B.; Rekik, B. Genetic analysis of milk production traits of Tunisian Holsteins using random regression test-day model with Legendre polynomials. *AJAS* **2018**, *31*, 636–642. [[CrossRef](#)]
60. Gebreyesus, G.; Buitenhuis, A.J.; Poulsen, N.A.; Visker, M.H.P.W.; Zhang, Q.; van Valenberg, H.J.F.; Sun, D.; Bovenhuis, H. Multi-population GWAS and enrichment analyses reveal novel genomic regions and promising candidate genes underlying bovine milk fatty acid composition. *BMC Genom.* **2019**, *20*, 178. [[CrossRef](#)]
61. Cohen-Zinder, M. 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.* **2005**, *15*, 936–944. [[CrossRef](#)]
62. Li, C.; Sun, D.; Zhang, S.; Yang, S.; Alim, M.A.; Zhang, Q.; Li, Y.; Liu, L. Genetic effects of FASN, PPARGC1A, ABCG2 and IGF1 revealing the association with milk fatty acids in a Chinese Holstein cattle population based on a post genome-wide association study. *BMC Genet.* **2016**, *17*, 110. [[CrossRef](#)]
63. Saatchi, M.; Schnabel, R.D.; Taylor, J.F.; Garrick, D.J. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genom.* **2014**, *15*, 442. [[CrossRef](#)]

64. Smith, J.L.; Wilson, M.L.; Nilson, S.M.; Rowan, T.N.; Oldeschulte, D.L.; Schnabel, R.D.; Decker, J.E.; Seabury, C.M. Genome-wide association and genotype by environment interactions for growth traits in U.S. Gelbvieh cattle. *BMC Genom.* **2019**, *20*, 926. [[CrossRef](#)]
65. McClure, M.C.; Morsci, N.S.; Schnabel, R.D.; Kim, J.W.; Yao, P.; Rolf, M.M.; McKay, S.D.; Gregg, S.J.; Chapple, R.H.; Northcutt, S.L.; et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim. Genet.* **2010**, *41*, 597–607. [[CrossRef](#)]
66. Snelling, W.M.; Allan, M.F.; Keele, J.W.; Kuehn, L.A.; McDanel, T.; Smith, T.P.L.; Sonstegard, T.S.; Thallman, R.M.; Bennett, G.L. Genome-wide association study of growth in crossbred beef cattle¹². *J. Anim. Sci.* **2010**, *88*, 837–848. [[CrossRef](#)] [[PubMed](#)]
67. Ibeagha-Awemu, E.M.; Peters, S.O.; Akwanji, K.A.; Imumorin, I.G.; Zhao, X. High density genome wide genotyping-by-sequencing and association identifies common and low frequency SNPs, and novel candidate genes influencing cow milk traits. *Sci. Rep.* **2016**, *6*, 31109. [[CrossRef](#)]
68. Kiser, J.N.; Clancey, E.; Moraes, J.G.N.; Dalton, J.; Burns, G.W.; Spencer, T.E.; Neibergs, H.L. Identification of loci associated with conception rate in primiparous Holstein cows. *BMC Genom.* **2019**, *20*, 840. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).