

## Article

# Genome-Wide Association Study for Udder Conformation Traits in Chinese Holstein Cattle

Mudasir Nazar<sup>1</sup>, Ismail Mohamed Abdalla<sup>1</sup>, Zhi Chen<sup>1</sup> , Numan Ullah<sup>1</sup>, Yan Liang<sup>1</sup> , Shuangfeng Chu<sup>1</sup>, Tianle Xu<sup>2</sup> , Yongjiang Mao<sup>1</sup>, Zhangping Yang<sup>1</sup>  and Xubin Lu<sup>1,\*</sup> 

<sup>1</sup> College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China

<sup>2</sup> Joint International Research Laboratory of Agriculture and Agri-Product Safety, Yangzhou University, Yangzhou 225009, China

\* Correspondence: dx120180094@yzu.edu.cn; Tel.: +86-182-5271-4348

**Simple Summary:** Udder conformation traits are economically important for dairy animals in the dairy industry. Milk production loss can be reduced via a better udder structure in dairy cattle. Udder traits are related to milk production and the somatic cell count, which is a sign of mammary infections (mastitis); as such, it is vital to understand the genetic architecture underlying udder traits in Holstein Friesian cattle for genetic development and long-term selection. Through a GWAS on udder structure traits in Chinese Holstein cattle, we identified several significant single-nucleotide polymorphisms (SNPs) and candidate genes associated with udder traits. The results could provide useful information regarding the genetics architecture of udder structure traits, thus improving the genetic information, health, longevity, and production of dairy cattle.

**Abstract:** Udder conformation traits are one of the most economic traits in dairy cows, greatly affecting animal health, milk production, and producer profitability in the dairy industry. Genetic analysis of udder structure and scores have been developed in Holstein cattle. In our research, we conducted a genome-wide association study for five udder traits, including anterior udder attachment (AUA), central suspensory ligament (CSL), posterior udder attachment height (PUAH), posterior udder attachment width (PUAW), and udder depth (UD), in which the fixed and random model circulating probability unification (FarmCPU) model was applied for the association analysis. The heritability and the standard errors of these five udder traits ranged from  $0.04 \pm 0.00$  to  $0.49 \pm 0.03$ . Phenotype data were measured from 1000 Holstein cows, and the GeneSeek Genomic Profiler (GGP) Bovine 100 K SNP chip was used to analyze genotypic data in Holstein cattle. For GWAS analysis, 984 individual cows and 84,407 single-nucleotide polymorphisms (SNPs) remained after quality control; a total of 18 SNPs were found at the GW significant threshold ( $p < 5.90 \times 10^{-7}$ ). Many candidate genes were identified within 200kb upstream or downstream of the significant SNPs, which include *MGST1*, *MGST2*, *MTUS1*, *PRKN*, *STXBP6*, *GRID2*, *E2F8*, *CDH11*, *FOXP1*, *SLF1*, *TMEM117*, *SBF2*, *GC*, *ADGRB3*, and *GCLC*. Pathway analysis revealed that 58 Gene Ontology (GO) terms and 18 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were enriched with adjusted  $p$  values, and these GO terms and the KEGG pathway analysis were associated with biological information, metabolism, hormonal growth, and development processes. These results could give valuable biological information for the genetic architecture of udder conformation traits in dairy Holstein cattle.

**Keywords:** udder structure traits; SNPs; GWAS; FarmCPU; gene network analysis; Holstein cattle



**Citation:** Nazar, M.; Abdalla, I.M.; Chen, Z.; Ullah, N.; Liang, Y.; Chu, S.; Xu, T.; Mao, Y.; Yang, Z.; Lu, X. Genome-Wide Association Study for Udder Conformation Traits in Chinese Holstein Cattle. *Animals* **2022**, *12*, 2542. <https://doi.org/10.3390/ani12192542>

Academic Editor: Martino Cassandro

Received: 11 July 2022

Accepted: 20 September 2022

Published: 22 September 2022

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



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Udder conformation traits play a vital economic role in dairy cattle [1]. The udder traits of the cattle is one of the most important factors that can be used to estimate production performance [2]. Udder composite values were included in official national genomic evaluation systems to account for the impact of cow conformation on health traits in 2009 [3].

The udder structure traits are related to production losses [4] as a result of mastitis [5] and early culling of cattle [6]. A previous study showed that the selection for better udder condition reduced bovine mastitis and the somatic cell count in Danish Holstein cattle [7]. The structure of the udder is related to long-term milk production mastitis resistance [8]. Udder structure traits affect the mammary gland health and have been detected as phenotype traits for improving clinical mastitis resistance in dairy cattle [9]. The teat canal length measurements and the number of lactations are important parameters for udder health in dairy cattle [10]. Udder conformation traits are heritable and could be useful as phenotypes due to the decrease in mastitis in Holstein cattle [11]. The heritability of udder traits was reported at 0.23 and 0.32 for udder depth and teat length, respectively [12]. The heritability of udder traits ranged between 0.09 and 0.25 in Brown Swiss cattle [13]. In Czech Holstein cattle, genetic correlations between the current linear traits ranged from 0.75 between fore udder attachment and udder depth to 0.70 between rear udder height and rear udder width [14]. Bovine mammary infection is a serious risk, due to the fact of udder health problems that affect financial losses, wellbeing, and the production of dairy cattle [15]. Udder health can be improved using udder conformation traits [16]. Udder health is vital in dairy farming, as it is the foundation for cost-effective and clean milk production [10].

The identification of quantitative trait loci (QTL) is an essential step in identifying and understanding genetic variants associated with economically significant phenotypes. GWAS has become a widely used method for identifying QTL and genome regions associated with phenotypes in Holstein cattle [17]. GWAS has proven to be an effective tool for identifying genetic variant associated with economically important traits such as milk production [18], milk composition [19], fatty acid [20,21], milk protein composition [22], body conformation [23], reproduction [24], pigmentation [25], fertility [26] and mastitis [27,28]. GWAS has been performed on udder structure and teat size traits in Charolais cattle [29], and it has also been conducted for udder structure score and teat shape in Nellore-Angus crossbred cows [30]. A previous study revealed 12 QTL that control the different features of the mammary gland's morphology in the Fleckvieh cattle population [31]. Gene Ontology (GO) terms, KEGG pathway enrichment analysis, and gene networks refer to a better understanding of the biological mechanisms underlying the development traits and identify potential candidate genes [32,33]. Ten candidate genes were reported for udder structure traits, including *ESR1*, *FGF2*, *FGFR2*, *GLI2*, *IQGAP3*, *PGR*, *PRLR*, *RREB1*, *BTRC*, and *TGFBR2*, across three French dairy cattle breeds [33]. There are very few GWAS reports on mammary system performance in Chinese Holstein cows. Therefore, our research aimed to find SNPs and candidate genes associated with udder structure conformation traits, anterior udder attachment (AUA), central suspensory ligament (CSL), posterior udder attachment height (PUAH), posterior udder attachment width (PUAW), and udder depth (UD) in Chinese Holstein cattle. Our results enhance the knowledge of useful biological information regarding the genetic architecture of these udder traits., these identified SNPs and candidate genes would become helpful resources for the genetic selection in Holstein cattle.

## 2. Materials and Methods

### 2.1. Ethics Statement

Whole processes for hair sample collection and data collection of phenotype traits were measured according to the plan proposed by the Ministry of Agriculture of the People's Republic of China and the China Council on Animal Care. The Institutional Animal Health Care, and this study was also accepted by the Yangzhou University Animal Researchers Ethics Committee (Permit serial: SYXK (Su) IACUC 2012-0029). During the research, no animals were harmed.

## 2.2. Phenotype and Genotype Data Collection

The studied experimental population consisted of 1000 Holstein cows raised from 4 dairy farms in Jiangsu Province, China (Sihong Aideweigang Dairy Farm 1: 199; Xuyi Weigang Dairy Farm 2: 214; Xuzhou Yonghao Dairy Farm 3: 224; Huaxia Animal Husbandry Farm 4: 363 cattle). The hair samples for genotyping data were collected within two months (November–December, 2019). The five udder structure traits were measured individually by scores ranging from 1 to 9 points. The five udder structure traits consisted of the anterior udder attachment (AUA), central suspensory ligament (CSL), posterior udder attachment height (PUAH), posterior udder attachment width (PUAW) and udder depth (UD) of 1000 cows, were measured according to China National Standard (GB/T 35568-2017). Three expert technicians completed the phenotypic measurement of udder traits for each cow, and hair samples from the tail of 999 cattle were also collected for DNA extraction and genotyping.

## 2.3. Phenotypic and Genotypic Parameters

Statistical analysis and Pearson correlation of the phenotype traits were investigated using the computer software, SPSS (v.19.0, Chicago, IL, USA), for data analysis (mean, standard error, minimum and maximum scores, standard deviation, coefficient variance %, and phenotypic correlation of traits). Genetic analysis was carried out with an animal model using the DMU software (v.5.6) [34] to determine the heritability, and the genetic correlation between pairs of traits, as shown in the following equation:

$$y_{ijklm} = u + Herd_i + Year_j + Season_k + Parity_l + a_m + e_{ijklm} \quad (1)$$

where  $y_{ijklm}$  is the phenotypes in the year  $j$ th;  $k$ th is the season;  $l$ th is the individual parity  $m$ th from the herd of animals  $i$ th;  $u$  is the mean of the population; animal  $herd_i$  is the herd effect according to a cows origin from one of the four herds;  $Year_j$  is the year levels of the effects in  $j$ th;  $Season_k$  is the season levels of the effects  $k$ th and parity is the effect of parity  $l$ th;  $a$  is the additive effect of the individual  $m$ th, which was conducted by the pedigree information; and  $e$  is the residual in the  $j$ th year,  $k$ th season, and  $l$ th parity of the individual  $m$ th from the  $i$ th herd. All the effects were evaluated as fixed effects except for the overall mean. The pedigree data of the cows age could be traced back at least 3 generations (2009–2020), the parities of the cattle were from 1 to 4, and the 4 seasons represented “September–November, December–February, March–May, and June–August”.

## 2.4. Genotyping Data and Quality Control

Genome DNA was extracted and genotyped using the GGP Bovine 100k SNP Chip, by Neogen Corporation (<http://www.neogenchina.com.cn/>; accessed on 25 June 2022) based on ARC-UCD1.2/bosTau9 as the genome reference. The GGP Bovine 100K SNP chip consisting of 100,000 SNPs was used for genotyping individuals. Then, quality control was conducted using Plink software (v.1.90, MA, USA, Cambridge) [35], to remove the markers that did not comply with the following standard: (1) the individual call rate lower than 95%; (2) the genotyping call rate of single SNP lower than 90%; (3) the minor allele frequency (MAF) of SNP > 0.05; (4) deviated from the Hardy–Weinberg equilibrium ( $p < 1.0 \times 10^{-6}$ ). After quality control, there remained 984 cattle and 84,407 markers (SNPs) variants for further analyses. The distribution of the SNPs information on 29 chromosomes within 1mb window size shown in Figure S1.

## 2.5. Principal Component Analysis

For the analysis of the population structure, Plink 1.90 software (v1.90) [35] was used to perform PCA (principal component analysis) on 984 cattle genotypes with 84406 SNPs and whole information genome. The ggplot2 package in R (v.4.0.4) was used to visually analyze the PCA plot.

## 2.6. Association Analysis

This study used the multi-locus linear mixed model to perform the GWAS analysis between the SNPs and traits using the fixed and random model circulating probability unification (FarmCPU) method [36]. The FarmCPU method performed marker tests with associated markers as covariates in a fixed-effects model, and in a random-effects model, iteratively [36]. In order to obtain a convenient illustration, the names of the associated markers detected in the fixed-effects model at each cycle were called pseudo-quantitative trait nucleotides (QTNs). Pseudo-QTNs were used to define the kinship of individuals to avoid a model over-fitting problem in the fixed-effect model [36]. Population stratification is a critical component that can lead to false-positive results in association studies [37]. Consequently, the present study fitted the highest two principal components Principal (PCs) as covariate variables in the GWAS models. The fixed-effects model was as follows [36].

$$y_i = M_{i1}b_1 + M_{i2}b_2 + \dots + M_{it}b_t + S_{ij}d_j + e_i \quad (2)$$

where  $y_i$  is the  $i$ th of an individual;  $M_{i1}, M_{i2}, \dots, M_{it}$  are the genotypic of the pseudo-QTNs, started with an empty set;  $b_1, b_2, \dots, b_t$  are the consistent effects of the pseudo-QTNs;  $S_{ij}$  is the genotype of the  $i$ th individual and  $j$ th genetic marker (SNPs);  $d_j$  is the corresponding effect of the  $j$ th genetic marker;  $e_i$  are the residuals having a distribution with a zero mean and a variance of  $\sigma^2_e$ .

Each marker (SNP) had its own  $p$  value after substitution. The  $p$  values and the associated marker map were used to update the selection of the pseudo-QTNs using the SUPER algorithm [38] in a REM as follows:

$$y_i = u_i + e_i$$

where  $y_i$  and  $e_i$  are similar to in Equation (1), and  $u_i$  is the individual's  $i$ th overall genetic effect. The type I error false-positive rate was measured at the level of 5%, and the genome-wide significance of the threshold value was revealed according to this method (0.05/SNPs) after quality control, where  $N_{snps}$  were the remaining number of SNPs after quality control [39]. After Bonferroni correlation, the significant threshold value for the genome-wide association study was  $p < 5.90 \times 10^{-7}$  (0.05/84407). Then, the Manhattan plots and quantile–quantile plots (QQ plots) were drawn using the “CMplot” package in the R 4.1.0 software [40].

## 2.7. Annotation of Candidate Genes

All of the genes within a 200 kilobases (Kb) position of the significant SNPs were identified as candidate genes for udder traits. Two hundred kilobases is a common distance that has been used to annotate and find genes related to SNPs in previous association [41,42]. We identified genomic region and candidate gene through UCSC Genome Browser with the help of cow assembly in April 2018 (<http://genome.ucsc.edu/bosTau9>, ARS-CUCD1.2) accessed on 25 June 2022 [43].

## 2.8. Functional Pathway-Enrichment and Gene Network Analysis of Candidate Gene

In the study, to further understand the biological information among these candidate genes, we submitted all the candidate genes identified from the GWAS analysis into the DAVID database (Database for Visualization, Annotation and Integrated Discovery) software [44] for Gene Ontology terms (GO) [45] and KEGG pathway analysis [46].  $p$  values were adjusted for multiple tests using the Bonferroni correction. Gene network analysis was carried out among genes using online software for the identification of Interrelating Genes (STRING database software v.11.0) [47] and Cytoscape software v.3.8.1 (cytoscape.org) [48] to visualize the analysis of the results.

### 3. Results

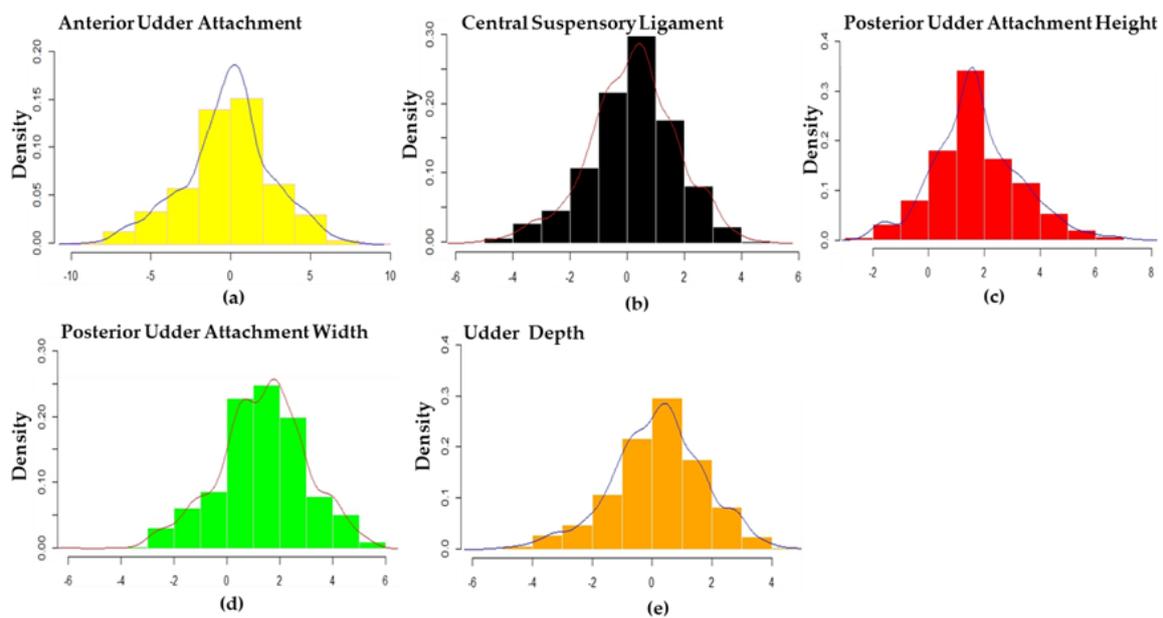
#### 3.1. Descriptive Statistical Data Analysis

The adjusted udder conformation traits of 984 cattle included the anterior udder attachment (AUA), central suspensory ligament (CSL), posterior udder attachment height (PUAH), posterior udder attachment width (PUAW), and udder depth (UD) are presented for normal distribution in this study. This descriptive statistical data analysis of the traits is shown in Table 1. The frequency distribution of the adjusted phenotype of udder traits is shown in Figure 1.

**Table 1.** Descriptive statistics for the adjustment of udder traits of Holstein cows (984).

Traits	Mean	SE	Min	Max	SD	CV%	Skewness	Kurtosis	p Value
AUA	0.17	0.09	−10.15	8.04	2.81	7.94	0.22	0.32	$2.53 \times 10^{-6}$
CSL	0.45	0.04	−1.88	8.85	2.67	3.50	0.33	0.50	$2.20 \times 10^{-7}$
PUAH	1.73	0.05	−2.2	7.32	1.55	2.41	0.34	0.53	$5.70 \times 10^{-8}$
PUAW	1.34	0.05	−5.9	5.59	1.68	2.83	0.26	0.23	$1.00 \times 10^{-4}$
UD	0.18	0.05	−5.29	4.75	1.56	2.43	0.24	0.34	$3.206 \times 10^{-5}$

Mean; SE, standard error; Min, minimum; Max, maximum; SD, standard deviation; CV, coefficient of variation. AUA, anterior udder attachment; CSL, central suspensory ligament; PUAH, posterior udder attachment height; PUAW, posterior udder attachment width; UD, udder depth.



**Figure 1.** Distribution of the adjusted phenotype of AUA (a), CSL (b), PUAH (c), PUAW (d), and UD (e) in the population of Holstein cows. The adjusted phenotypes of the five traits shown about normal distribution.

#### 3.2. Phenotypic and Genetic Correlations, and Heritability Estimation of Udder Traits

The phenotypic and genetic correlations between these traits are shown in Table 2. In the phenotypic correlation, the CSL trait was found to be positively correlated with PUAW and UD, and the other traits were negatively correlated with CSL. The CSL trait was genetically positively associated with PUAW, and other traits were genetically negative correlated. In contrast, the AUA trait was phenotypically positively correlated with CSL, PUAW and UD except PUAH. The distribution and phenotypic correlation among the udder traits are shown in Figure S2. The AUA trait was genetically positively correlated with PUAH and UD, and the other traits were negatively correlated. The PUAH trait was genetically positively correlated with PUAW and AUA, and it was phenotypically negative correlated. The PUAW trait was genetically highly negatively correlated with AUA and UD, and other traits were found to have low to moderate genetically positive

correlation, while the PUAW trait was found to be positively phenotypic correlated with all traits. The genotypic correlation among the udder traits is shown (Figure S3). Moreover, the heritability and the standard errors were estimated for the udder traits at  $0.24 \pm 0.02$ ,  $0.34 \pm 0.03$ ,  $0.04 \pm 0.00$ ,  $0.13 \pm 0.03$ , and  $0.49 \pm 0.03$ , respectively (Table 2).

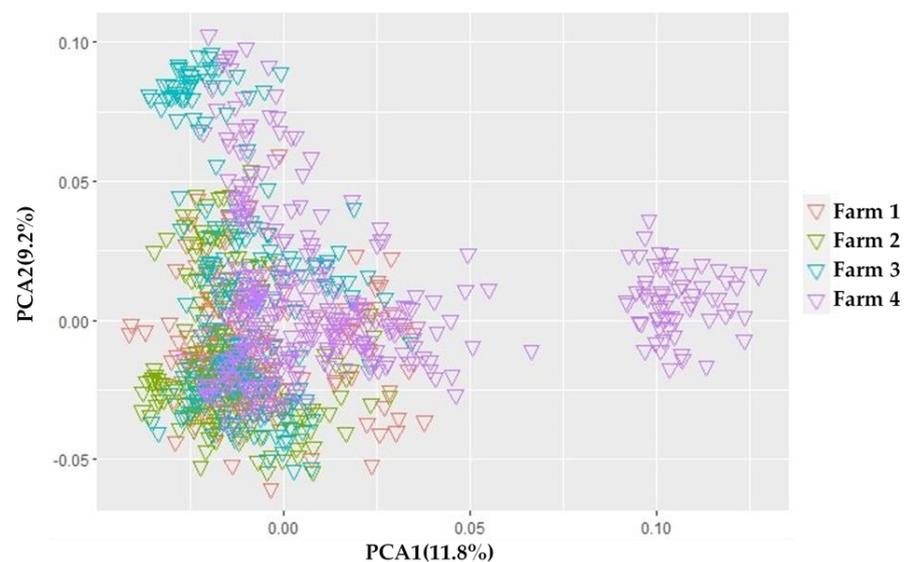
**Table 2.** Genotyping (upper diagonal) and phenotyping (lower diagonal) correlations, and the heritability (grey, diagonal) for udder traits in Holstein cattle.

Traits	AUA	CSL	PUAH	PUAW	UD
AUA	0.24 (0.02)	−0.38	0.22	−0.62	0.10
CSL	0.02	0.34 (0.03)	−0.44	0.52	−0.45
PUAH	−0.11 **	−0.37 **	0.04 (0.00)	0.14	0.09
PUAW	0.17 **	0.14 **	0.002	0.13 (0.01)	−0.49
UD	0.26 **	0.14 **	−0.07 *	0.07 *	0.49 (0.03)

AUA, anterior udder attachment; CSL, central suspensory ligament; PUAH, posterior udder attachment height; PUAW, posterior udder attachment width; UD, udder depth. The upper subscript \* and \*\* represent significant correlation at 0.05 and 0.01, respectively.

### 3.3. Population Structure

This research used Principal Component Analysis to visualize the family structure. The result showed that all the populations were divided into two sized clusters group. The PCA results in Figure 2 identified that the first two principal components PCA1 and PCA2 participated in 11.8% and 9.2%, respectively; at the same time, the total variation was approximately 21%; therefore, the FarmCPU model was used to fit the first two PCs in the PCA association analysis as covariate variables. The mixed linear model considered several population stratifications based on the results of the PCA analysis. Two groups were represented in the population structure (Figure 2).

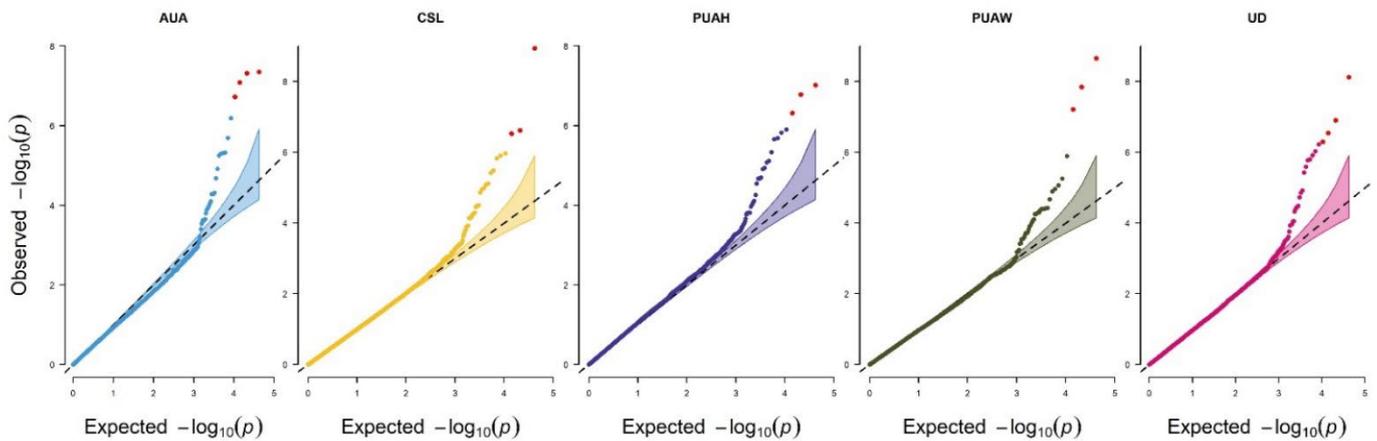


**Figure 2.** Population structure demonstrated by the 984 cattle raised at four animal farms. Principal component analysis (PCA).

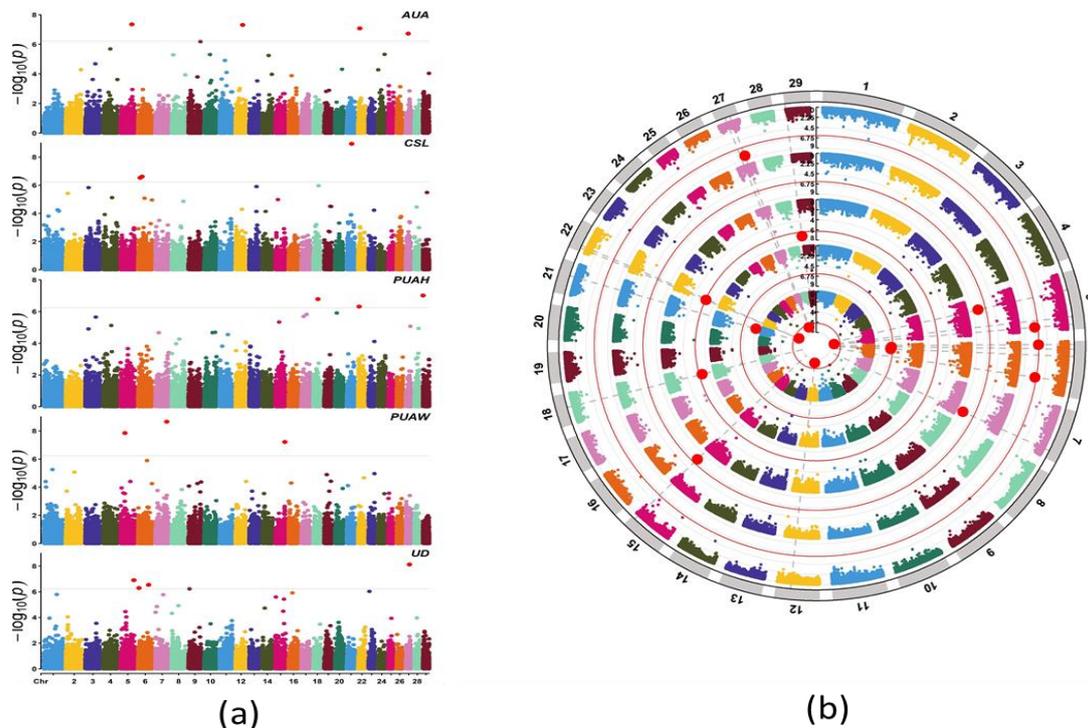
### 3.4. Genome-Wide Association Study

The FarmCPU model was used in the current study to conduct the genome-wide association analysis. The quantile–quantile (QQ) plots demonstrated that the genome-wide association study analysis model was reasonable in this research (Figure 3). The lambda values ( $\lambda$ ) for the AUA, CSL, PUAH, PUAW and UD were 0.91, 1, 1.03, 0.94, and 1.05, respectively, and they were all close to 1.06. The red dot at the top right corner of the QQ plots also displayed the significant SNPs that were associated with the udder conformation traits in this study (Figure 3), and the population stratification was sufficiently controlled.

The Manhattan plots reveal the results of the GWAS significance levels ( $-\log_{10}$  of the  $p$ -value of each SNP) by chromosomes position; each chromosome showed a different color. Significant SNPs in the Manhattan plot were strongly associated with traits (Figure 4a). A Circular Manhattan plot for significance  $\log_{-10}$  ( $p$  values) of the associated SNPs were shown relations to five udder traits. The five circles from outside to inside represent the udder conformation traits. (Figure 4b).



**Figure 3.** The quantile–quantile plots of the five udder traits from the GWAS in Holstein cow. The separation between the observed and expected values was analyzed using QQ plots. The null hypothesis indicated no relationship with the red lines. Deviation of the expected  $p$ -value showed that the population stratification was adequately controlled in the tails for every trait. The red dot significant SNPs indicate the threshold value.



**Figure 4.** (a) Manhattan plots established from the GWAS results of the udder conformation traits in Holstein cattle. The significant threshold was  $p < 5.9 \times 10^{-7}$ . The Manhattan plot displays genomic SNPs on the horizontal axis (x-axis) along with chromosomes and the negative logarithm of each SNP' association  $p$  value on the vertical axis (y-axis). After the Bonferroni correction, the green line shows a significant threshold level. (b) In the Circular Manhattan plot, the udder conformation traits were plotted from outside to inside, respectively.

As stated before, the threshold value was  $5.90 \times 10^{-7}$  for significant SNPs in the GWAS. The 18 SNPs passed the threshold and were significantly associated with five udder traits (AUA, CSL, PUAH, PUAW, and UD), four SNPs (DB-340-seq-rs208014256, Hapmap58214-rs29015775, BovineHD2700005329, and BovineHD0900028603) positioned on chromosomes 5, 22, 27 and 9 were identified to be associated with the trait AUA. Three SNPs (ARS-BFGL-BAC-29174, Hapmap32447-BTC-033214, and BovineHD0600005127, located on chromosomes 21, 6, and 6, respectively) were identified to be associated with the trait CSL. For PUAH traits, three SNPs were presented on chromosomes 29, 18, and 22 (BovineHD2900000083, BovineHD1800011193, and BovineHD2200002408, respectively) Moreover, three significantly associated SNPs (BovineHD0700028083, BovineHD0500010522, and BovineHD1500023322) located on chromosomes 7, 5, and 15, respectively, were indicated to be associated with the trait PUAW. While five SNPs (BTA-75047-no-rs, BovineHD0600024277, BovineHD0600001885, BovineHD0900001933, and BovineHD2300001734) detected on chromosomes 5, 6, 6, 9, and 23, respectively, were found to be associated with the trait UD (Table 3).

**Table 3.** GWAS significant SNPs associated with udder structure conformation traits in Holstein cattle.

Traits	SNPs	CHR	Position (kb)	MAF	Nearest Gene	Distance (kb)	p-Value	Effect
AUA	DB-340-seq-rs208014256	5	93520616	0.46	<i>MGST1</i>	Within	$4.48 \times 10^{-8}$	0.330783
	Hapmap58214-rs29015775	22	13159539	0.49	<i>LOC101903734</i>	Within	$8.34 \times 10^{-8}$	-0.35043
	BovineHD2700005329	27	19594311	0.16	<i>MTUS1</i>	Within	$1.90 \times 10^{-7}$	-0.47118
	BovineHD0900028603	9	97665052	0.25	<i>PRKN</i>	Within	$6.48 \times 10^{-7}$	0.371653
CSL	ARS-BFGL-BAC-29174	21	40773446	0.43	<i>STXBP6</i>	100 kb	$1.16 \times 10^{-9}$	0.36344
	Hapmap32447-BTC-033214	6	32254947	0.42	<i>GRID2</i>	Within	$2.45 \times 10^{-7}$	-0.31119
	BovineHD0600005127	6	17417238	0.39	<i>LOC112447148</i>	Within	$3.02 \times 10^{-7}$	-0.31736
PUAH	BovineHD2900000083	29	702083	0.44	<i>E2F8</i>	100 kb	$9.70 \times 10^{-8}$	0.298361
	BovineHD1800011193	18	37485453	0.47	<i>CDH11</i>	100 kb	$1.66 \times 10^{-7}$	0.26997
	BovineHD2200002408	22	8008314	0.12	<i>FOXP1</i>	Within	$4.89 \times 10^{-7}$	-0.42465
PUAW	BovineHD0700028083	7	93970405	0.38	<i>SLF1</i>	Within	$2.26 \times 10^{-9}$	-0.33939
	BovineHD0500010522	5	36446050	0.50	<i>TMEM117</i>	Within	$1.45 \times 10^{-8}$	-0.33019
	BovineHD1500023322	15	78715609	0.40	<i>SBF2</i>	Within	$6.19 \times 10^{-8}$	-0.30795
UD	BTA-75047-no-rs	5	109433376	0.05	<i>LGALS2</i>	Within	$1.26 \times 10^{-7}$	-0.66283
	BovineHD0600024277	6	86964714	0.21	<i>GC</i>	Within	$2.92 \times 10^{-7}$	0.321971
	BovineHD0600001885	6	7087395	0.35	<i>UBE2K</i>	50 kb	$5.16 \times 10^{-7}$	-0.27781
	BovineHD0900001933	9	8043006	0.28	<i>ADGRB3</i>	Within	$5.98 \times 10^{-7}$	-0.28799
	BovineHD2300001734	23	6986268	0.26	<i>GCLC</i>	Within	$9.36 \times 10^{-7}$	-0.32781

SNP: single nucleotide polymorphism; CHR: chromosomes; AUA, anterior udder attachment; CSL, central suspensory ligament; PUAH, posterior udder attachment height; PUAW, posterior udder attachment width; UD, udder depth; MAF: minor allele frequency; Effect: each variation regression coefficient.

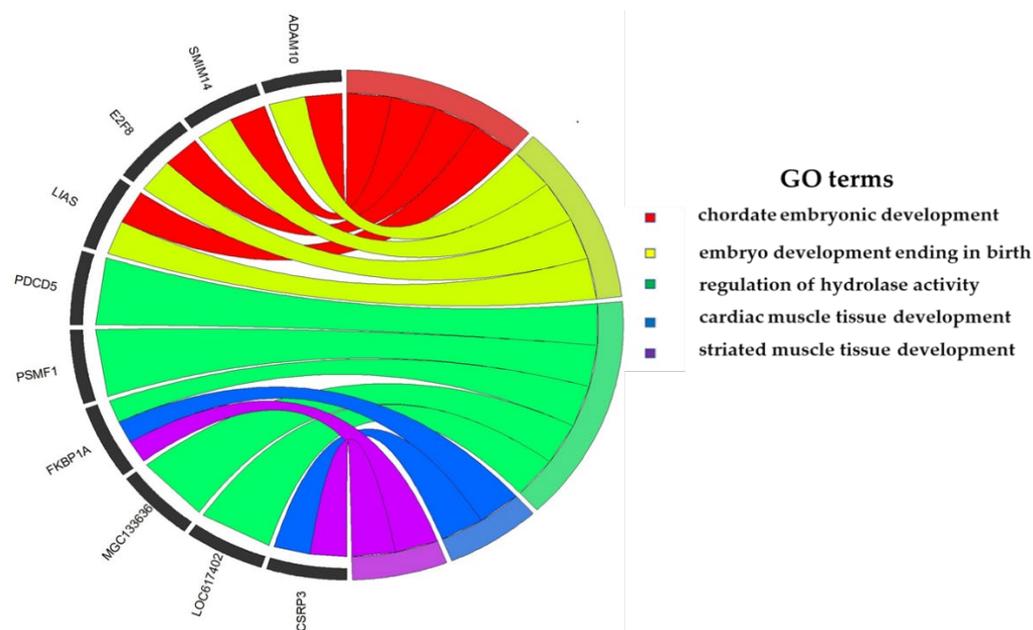
### 3.5. Annotation of the Candidate Genes

In our study, genes that were identified at a distance within 200 kb of significant SNPs were recognized as candidate genes. Among all of the 18 significant SNPs associated with udder conformation traits, 14 of them were located within the following genes: Microsomal Glutathione S-Transferase 1 (*MGST1*); LOC101903734, Microtubule-Associated Scaffold Protein 1 (*MTUS1*), *PRKN* (Parkin RBR E3ubiquitin protein ligase); *GRID2* (Glutamate Ionotropic Receptor Delta-Type Subunit 2); LOC112447118, Forkhead Box P1 (*FOXP1*), SMC5-SMC6 Complex Localization Factor 1 (*SLF1*); Transmembrane Protein 117 (*TMEM117*), SET Binding Factor 2 (*SBF2*); Galectin 2 (*LGALS2*); GC Vitamin D Binding Protein (*GC*); *ADGRB3* (Adhesion G Protein-Coupled Receptor B3); *GCLC* (Glutamate-Cysteine Ligase Catalytic Subunit). Although, one of the SNPs on Chr6 was located near (50 kb) to the Ubiquitin-Conjugating Enzyme E2 K gene (*UBE2K*), three SNPs on chromosomes 21, Chr29, Chr18 were located closely (100 kb) to the *STXBP6* (Syntaxin-Binding Protein 6), *E2F8* (E2F Transcription Factor 8), and Cadherin 11 (*CDH11*), respectively (Table 3).

### 3.6. Enrichment Analysis

The Asian server used the NCBI database and UCSC Genome Browser by using cattle assembly in April 2018 (ARC-UCD1.2/bosTau9), and a total of 141 candidate genes were

identified within the 200 kb region upstream/downstream of the significant SNPs for the udder structure traits; all of the candidate genes were used for KEGG pathway and enrichment analysis (GO). The results of the GO analysis determined 58 Gene Ontology terms (Table S1 and Figure 5).



**Figure 5.** Gene Ontology term results from the udder conformation traits.

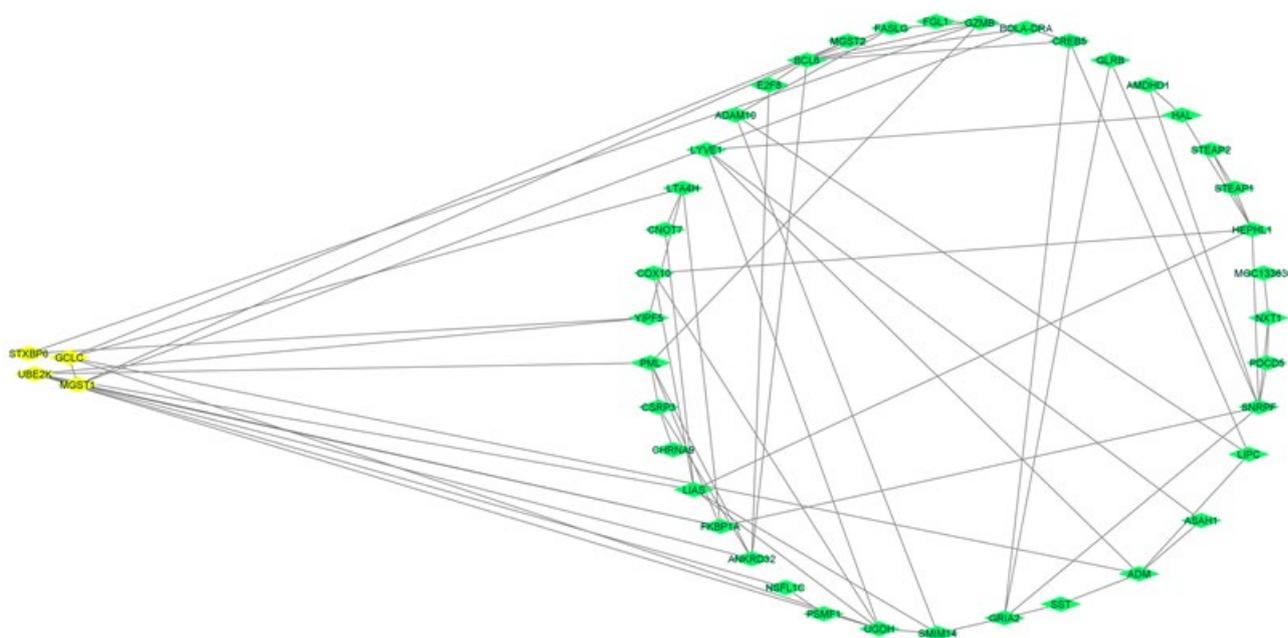
The KEGG pathways analysis revealed 18 pathways (i.e., bta04978: mineral absorption involved three genes; bta00480: glutathione metabolism comprises three genes; bta00340: histidine metabolism had two genes; bta04935: growth hormone synthesis, secretion and action involved three genes) which are shown in Table 4, according to the adjusted  $p$  value. A KEGG pathway analysis of the dot plot is presented in the Supplementary Material (Figure S4).

**Table 4.** Details of the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways analysis enriched from the nearest candidate genes and genes within 200 kb of the significant SNPs.

KEGG ID	Description	Count	Gene Name
bta04978	Mineral absorption	3	<i>STEAP1, STEAP2, HEPHL1</i>
bta00480	Glutathione metabolism	3	<i>MGST1, MGST2, GCLC</i>
bta00340	Histidine metabolism	2	<i>HAL, AMDHD1</i>
bta04935	Growth hormone synthesis, secretion and action	3	<i>CREB5, SST, CACNA1D</i>

### 3.7. Gene Network Analyses

The STRING database was used to perform a protein–protein interaction network using all the genes previously used in functional analysis. From the relationships between genes identified in this gene network, there were many interactions among the genes (consisting of 44 nodes related via 75 edges). The corresponding interactivity power among these genes was determined by the staining strength between the lines that linked one gene to the other (Figure 6).



**Figure 6.** Gene network analysis for genes associated with udder structure traits in Holstein cow. (The yellow nodes represent the significant candidate genes and their interaction with related genes).

#### 4. Discussion

In this research, the heritability of the udder traits was low and medium between 0.04 and 0.49 as shown in Table 2. The results of the heritability of udder traits were reliable with the predictions by Wu et al. [49], who described the estimation of heritability of udder structure traits ranged from 0.08 to 0.22 in Chinese Holstein Cattle, while previous studies have reported the heritability of udder structure traits 0.18 to 0.37 in dairy cattle [5,50,51].

The phenotypic and genotypic correlations result of the udder conformation traits were negatively low to medium at  $-0.37$  to  $0.26$  and  $-0.62$  to  $0.52$ , respectively, for Chinese Holstein cows (Table 2); the results of the udder structure traits are in agreement with the findings of Němcová et al. [14], who stated that the conforming phenotypic correlations were medium at 0.23 to 0.46, and genotypic correlations among the udder traits range from 0.70 to 0.75 in Czech Holstein cattle. Udder balance (UB) and fore udder attachment (FUA) showed genetic correlations of 0.40, 0.63, and 0.39 in three different French breeds (Montbeliarde, Normande, and Holstein), respectively [33].

Population stratification was a significant causative factor in GWAS. When a GWAS analysis included samples with different genetic structures, the PCA procedure [52] could identify the population structure by sorting individuals into grouped ancestry based on their genetic structure. All of the research populations were divided into two groups: one group contained a large quantity of clusters, and the other consisted of a small number of clusters (Figure 2). This identifies that every group is clustered nearly together and has a genotypic relationship. The use of Holstein semen from different overseas countries may have resulted in the division of two groups. or the cows in these farms may contain blood from other breeds, as we found that Chinese Holstein cows can be registered which are a minimum of 87.5% Holstein blood (GB/T 3157 2008, Chinese Holstein Cattle). As we performed the population structure, the PCs were adjusted as covariates to the association study for proper population stratification. A previous study found that the inflation factor ( $\lambda$ ) for colostrum and serum albumin concentrations was 0.983 and 1.004 in Chinese Holstein cattle [53], respectively, indicating that the appropriate model successfully corrected the population stratification [54]. After correcting the population structure, the lambda ( $\lambda$ ) value should be close to 1 [55]. The deviation of the observed value from the expected value is near 1.06 on the QQ plot (Figure 3), and the lambda value ( $\lambda$ ) showed

0.94 < 1.05, both representing that the population stratification was adequately adjusted with a suitable method.

Population stratification is a major challenge during GWAS work analysis. Furthermore, population stratification can lead to many false positives in GWAS results due to incorrect associations [56]. Due to systematic ancestry variations in GWAS, population stratification is a significant confounding element that can lead to false positives [57]. Because there are numerous approaches for correcting population stratification, a statistical model can be a helpful tool for correcting and reducing the chances of type 1 error (false positive association) [58]. The FarmCPU method was used in our study for its benefits in fully controlling false positives, eliminating confusion, and improving the efficiency of computation by iteratively using a fixed-effect model, and random-effect models [36]. The QQ plots of the five traits are shown in Figure 4, and in this study, the effect of the cryptic interaction between the animals was effectively controlled.

In this research, GWAS revealed 18 significant SNPs (single-nucleotide polymorphism) associated with five udder structure traits in Holstein cattle using the FarmCPU model. Among all of the SNPs, two significant SNPs were ARS-BFGL-BAC-29174, located near the *STXBP6* gene, and BovineHD0700028083, located within *SLF1*. The *STXBP6* gene provides a syntoxin-binding protein which attaches to the SNARE complex components, preventing membrane fusions such as phagocytosis [59], and the *STXBP6* gene causes rheumatoid arthritis [60]. It has also been expressed in the breast cancer cells of mice [61]. The *SLF1* gene is also important in genome stability maintenance in human cells after DNA damage [62].

Our research identified genes associated with the AUA trait (i.e., *MGST1*, *MTUS1*, and *PRKN*). Some studies have determined that *MGST1* has the largest significant effect on Chr5 for fat production in U.S. Holstein cattle [17]. *MGST1* has been recognized extensively as a functional candidate gene for the QTL associated with milk composition features in Montbéliarde cows [63]. In two Holstein cattle breeds, the *MTUS1* gene was associated with meat quality and carcass traits [64]. The *PRKN* gene was associated with growth and production traits in Fuzhong Buffalo [65], and it was also associated with fatty acids in Chinese Wagyu cattle [66].

The Hapmap32447-BTC-033214 SNP located on Chr6; glutamate ionotropic receptor delta type subunit two genes (*GRID2*) is associated with the CSL trait, but, in the previous study, the *GRID2* was involved in the sexual maturity of Simmental cattle [67].

Furthermore, our results report that the *E2F8*, *CDH11*, and *FOXP1* genes are associated with the PUAH trait. E2F transcription factors, such as the *E2F8* gene control the expression of genes involved in cell cycle proliferation and progression [68]. The present study discovered a novel hormone, and developmental regulation of *E2F8* mRNA throughout bovine follicular growth [69]. The *CDH11* gene is involved in mastitis development and mammary gland growth in lactating cattle [70,71]. The *FOXP1* gene is important for mammary gland development during puberty and maturity in mice.

However, the results show that the *TMEM117*, and *SBF2* genes are associated with the PUAW trait. Zhu et al. [72] also conducted a GWAS and found that the *TMEM117* gene was related to saturated fatty acids composition in Simmental cattle. The *SBF2* gene has been identified by GWAS for the total solid yield trait in Thai dairy cows [73].

The BTA-75047-no-rs, BovineHD0600024277, BovineHD0900001933, and BovineHD2300001734 SNPs are located within genes (i.e., *LGALS2*, *GC*, *ADGRB3*, and *GCLC*, respectively), are associated with udder depth trait; meanwhile, one BovineHD0600001885 SNP located on Chr6 (50kb) was also related with this trait. Zexi et al. [74] reported that the *GC* gene has a pleiotropic effect that involves milk production and evidence of mastitis in Nordic Holstein cows. The result of the *GC* gene on milk production in Holstein cattle has already been investigated in association analysis [75]. The *ADGRB3* gene is responsible for the uniformity of yearling weight in Nellore cattle [76]. Functional pathway analyses of list candidate genes are shown; The *GCLC* gene is consistently increased as parturition approaches and lactation began in Holstein Dairy cattle [77], while our finding provides evidence that *GCLC* is associated with udder depth trait.

In our research, GO and KEGG pathways analysis identified many GO terms and KEGG pathways. These included significant candidate genes that are associated with udder traits under analysis, for example, Proteolysis, a GO:0006508 biological process which has nine genes in Table S1, among which is *PDCD5*. The protein-containing complex subunit organization, GO:0043933, had eight genes including *UGDH*, *STXBP6*, and *FKBP1A*. The cellular amide metabolic process (GO:0043603) had seven genes involving *MGST2* and *HAL*. Embryo development ending in birth (GO:0009792) consisted of four genes, including *E2F8*. Cardiac muscle tissue development (GO:0048738) consisted of two genes (*CSRP3*, and *FKBP1A*). All these identified genes were closely related to our significant SNPs. As mentioned above, some candidate genes that participated in the GO term may play an essential role in the biological function of body traits, and afterward, these candidate genes participate in udder traits. The *PDCD5* gene regulates cell proliferation, cell cycle progression, and apoptosis in bovine cancer cells (*A431*) [78]. The *MGST2* has been found as a potential gene in the bovine mammary gland [79]. The *UGDH* gene is associated with milk production traits in Chinese Holstein [80]. These four KEGG pathway analyses were enriched for udder traits; mineral absorption, glutathione metabolism, histidine metabolism, growth hormone synthesis, secretion and action. The minerals absorption pathway was significantly related to udder traits of dairy cow [81]. Glutathione metabolism affected by liver function due to high blood glutathione in early lactation of Holstein cows [82]. Histidine metabolism is involved in the mammary cell growth, protein synthesis, and milk yield of Holstein cows [83]. Prolactin is a growth hormone related to lactation and mammary growth in dairy cows [84]. The functional analysis resulted from several terms and pathways related to protein-containing compounds (amino acids), absorption, catabolism and metabolism, and liver growth. Hence, it was assumed that all significant SNPs and candidate genes might possibly be associated with udder conformation traits.

## 5. Conclusions

In conclusion, our research found 18 significant SNPs associated with udder conformation traits in Chinese Holstein cattle. Several candidate genes harbored SNPs (i.e., *MGST1*, *MGST2*, *MTUS1*, *PRKN*, *STXBP6*, *GRID2*, *E2F8*, *CDH11*, *FOXP1*, *SLF1*, *TMEM117*, *SBF2*, *GC*, *ADGRB3* and *GCLC*) identified mostly to participate in biological information, metabolism, and development processes. Our findings provide useful biological information for understanding the genetic architecture for improving udder traits and will therefore contribute to the genetic selection of Chinese Holstein cattle.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani12192542/s1>, Figure S1: The distribution of the SNPs information on 29 chromosomes within 1mb window size, Figure S2: The distribution and correlation of the phenotype among the udder traits in Chinese Holstein cattle, Figure S3: Genotype correlation among the udder traits, Figure S4: KEGG dot plot, Table S1: GO Results.

**Author Contributions:** Conceptualization, X.L., Z.Y. and M.N.; methodology, M.N.; software, M.N. and X.L.; validation, X.L. and Z.Y.; formal analysis, M.N. and X.L.; investigation, T.X., Z.C., M.N., X.L. and I.M.A.; resources, X.L.; data curation, X.L., M.N. and I.M.A.; writing—original draft preparation, M.N.; writing—review and editing, N.U., Y.L. and S.C.; visualization, Z.Y.; supervision, Z.Y.; project administration, Z.Y., Y.M. and Z.C.; funding acquisition, Z.Y. and X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Earmarked Fund for Jiangsu Agricultural Industry Technology System (JATS [2021]486), the National Natural Science Foundation of China (31872324), and Jiangsu Province Seed Industry Revitalization Revealing the Leaders Project (JBGS [2021]117).

**Institutional Review Board Statement:** The research was conducted according to the strategies of the Institutional Animal Care and Use Committee of the School of the Yangzhou University Animal Experiments Ethics Committee (License Number: SYXK (Su) IACUC 2012-0029).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** In this study, the data presented are available on request from the corresponding author.

**Acknowledgments:** The authors would like to thank all those who assisted us during this project for important guidance and technical support.

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

## References

1. Devani, K.; Plastow, G.; Orsel, K.; Valente, T.S. Genome-wide association study for mammary structure in Canadian Angus cows. *PLoS ONE* **2020**, *15*, e0237818. [[CrossRef](#)] [[PubMed](#)]
2. Vukašinović, N.; Moll, J.; Künzi, N. Genetic relationships among longevity, milk production, and type traits in Swiss Brown cattle. *Livest. Prod. Sci.* **1995**, *41*, 11–18. [[CrossRef](#)]
3. Wiggans, G.R.; VanRaden, P.M.; Cooper, T.A. The genomic evaluation system in the United States: Past, present, future. *J. Dairy Sci.* **2011**, *94*, 3202–3211. [[CrossRef](#)] [[PubMed](#)]
4. Foster, M.K.; Caldwell, K.R.; Arnold, A.N.; Griffin, D.B.; Gehring, K.B.; Savell, J.W. Palatability assessments of beef strip loin steaks portioned by weight or by thickness sourced from various carcass weight/ribeye area size combinations. *Meat Sci.* **2021**, *172*, 108319. [[CrossRef](#)] [[PubMed](#)]
5. Persson Waller, K.; Persson, Y.; Nyman, A.K.; Stengårde, L. Udder health in beef cows and its association with calf growth. *Acta Vet. Scand.* **2014**, *56*, 9. [[CrossRef](#)]
6. Riley, D.G.; Sanders, J.O.; Knutson, R.E.; Lunt, D.K. Comparison of F1 Bos indicus x Hereford cows in central Texas: II. Udder, mouth, longevity, and lifetime productivity. *J. Anim. Sci.* **2001**, *79*, 1439–1449. [[CrossRef](#)]
7. Lund, T.; Miglior, F.; Dekkers, J.C.M.; Burnside, E.B. Genetic relationships between clinical mastitis, somatic cell count, and udder conformation in Danish Holsteins. *Livest. Prod. Sci.* **1994**, *39*, 243–251. [[CrossRef](#)]
8. Rupp, R.; Boichard, D. Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins. *J. Dairy Sci.* **1999**, *82*, 2198–2204. [[CrossRef](#)]
9. Carlström, C.; Pettersson, G.; Johansson, K.; Strandberg, E.; Stålhammar, H.; Philipsson, J. Feasibility of using automatic milking system data from commercial herds for genetic analysis of milkability. *J. Dairy Sci.* **2013**, *96*, 5324–5332. [[CrossRef](#)]
10. Klein, D.; Flöck, M.; Khol, J.L.; Franz, S.; Stüger, H.P.; Baumgartner, W. Ultrasonographic measurement of the bovine teat: Breed differences, and the significance of the measurements for udder health. *J. Dairy Res.* **2005**, *72*, 296–302. [[CrossRef](#)]
11. Vanraden, P.M.; Jensen, E.L.; Lawlor, T.J.; Funk, D.A. Prediction of Transmitting Abilities for Holstein Type Traits. *J. Dairy Sci.* **1990**, *73*, 191–197. [[CrossRef](#)]
12. Dal Zotto, R.; De Marchi, M.; Dalvit, C.; Cassandro, M.; Gallo, L.; Carnier, P.; Bittante, G. Heritabilities and genetic correlations of body condition score and calving interval with yield, somatic cell score, and linear type traits in brown swiss cattle. *J. Dairy Sci.* **2007**, *90*, 5737–5743. [[CrossRef](#)]
13. Campos, R.V.; Cobuci, J.A.; Kern, E.L.; Costa, C.N.; McManus, C.M. Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 476–484. [[CrossRef](#)] [[PubMed](#)]
14. Němcová, E.; Stípková, M.; Zavadilová, L. Genetic parameters for linear type traits in Czech Holstein cattle. *Czech J. Anim. Sci.* **2011**, *56*, 157–162. [[CrossRef](#)]
15. Hogeveen, H.; Huijps, K.; Lam, T.J.G.M. Economic aspects of mastitis: New developments. *N. Z. Vet. J.* **2011**, *59*, 16–23. [[CrossRef](#)] [[PubMed](#)]
16. Nakov, D.; Hristov, S.; Andonov, S.; Trajchev, M. Udder-related risk factors for clinical mastitis in dairy cows. *Vet. Arh.* **2014**, *84*, 111–127.
17. Jiang, J.; Ma, L.; Prakapenka, D.; VanRaden, P.M.; Cole, J.B.; Da, Y. A large-scale genome-wide association study in U.S. Holstein cattle. *Front. Genet.* **2019**, *10*, 412. [[CrossRef](#)]
18. 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](#)]
19. 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](#)]
20. Li, C.; Sun, D.; Zhang, S.; Wang, S.; Wu, X.; Zhang, Q.; Liu, L.; Li, Y.; Qiao, L. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. *PLoS ONE* **2014**, *9*, e0243343. [[CrossRef](#)]
21. Li, X.; Buitenhuis, A.J.; Lund, M.S.; Li, C.; Sun, D.; Zhang, Q.; Poulsen, N.A.; Su, G. Joint genome-wide association study for milk fatty acid traits in Chinese and Danish Holstein populations. *J. Dairy Sci.* **2015**, *98*, 8152–8163. [[CrossRef](#)] [[PubMed](#)]
22. 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](#)]
23. Song, J.; Sapi, E.; Brown, W.; Nilsen, J.; Tartaro, K.; Kacinski, B.M.; Craft, J.; Naftolin, F.; Mor, G. Roles of Fas and Fas ligand during mammary gland remodeling. *J. Clin. Investig.* **2000**, *106*, 1209–1220. [[CrossRef](#)] [[PubMed](#)]
24. Sahana, G.; Guldbbrandtsen, B.; Lund, M.S. Genome-wide association study for calving traits in Danish and Swedish Holstein cattle. *J. Dairy Sci.* **2011**, *94*, 479–486. [[CrossRef](#)]

25. Fan, Y.; Wang, P.; Fu, W.; Dong, T.; Qi, C.; Liu, L.; Guo, G.; Li, C.; Cui, X.; Zhang, S.; et al. Genome-wide association study for pigmentation traits in Chinese Holstein population. *Anim. Genet.* **2014**, *45*, 740–744. [[CrossRef](#)]
26. Liu, A.; Wang, Y.; Sahana, G.; Zhang, Q.; Liu, L.; Lund, M.S.; Su, G. Genome-wide Association Studies for Female Fertility Traits in Chinese and Nordic Holsteins. *Sci. Rep.* **2017**, *7*, 8487. [[CrossRef](#)]
27. Cai, Z.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Prioritizing candidate genes post-GWAS using multiple sources of data for mastitis resistance in dairy cattle. *BMC Genom.* **2018**, *19*, 656. [[CrossRef](#)]
28. Wang, X.; Ma, P.; Liu, J.; Zhang, Q.; Zhang, Y.; Ding, X.; Jiang, L.; Wang, Y.; Zhang, Y.; Sun, D.; et al. Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC Genet.* **2015**, *16*, 111. [[CrossRef](#)]
29. Vallée, A.; Daure, J.; Van Arendonk, J.A.M.; Bovenhuis, H. Genome-wide association study for behavior, type traits, and muscular development in Charolais beef cattle. *J. Anim. Sci.* **2016**, *94*, 2307–2316. [[CrossRef](#)]
30. Tolleson, M.W.; Gill, C.A.; Herring, A.D.; Riggs, P.K.; Sawyer, J.E.; Sanders, J.O.; Riley, D.G. Association of udder traits with single nucleotide polymorphisms in crossbred bos indicus–bos taurus cows. *J. Anim. Sci.* **2017**, *95*, 2399–2407. [[CrossRef](#)]
31. Pausch, H.; Emmerling, R.; Schwarzenbacher, H.; Fries, R. A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genet. Sel. Evol.* **2016**, *48*, 47. [[CrossRef](#)] [[PubMed](#)]
32. Soares, A.C.C.; Guimaraes, S.E.F.; Kelly, M.J.; Fortes, M.R.S.; e Silva, F.F.; Verardo, L.L.; Mota, R.; Moore, S. Multiple-trait genomewide mapping and gene network analysis for scrotal circumference growth curves in Brahman cattle. *J. Anim. Sci.* **2017**, *95*, 3331. [[CrossRef](#)] [[PubMed](#)]
33. Marete, A.; Lund, M.S.; Boichard, D.; Ramayo-Caldas, Y. A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds. *PLoS ONE* **2018**, *13*, e0199931. [[CrossRef](#)] [[PubMed](#)]
34. Madsen, P.; Jensen, J.A. A User'S Guide to DMU. A Package for Analysing Multivariate Mixed Models. Available online: <https://dmu.ghpc.au.dk/dmu/index.html> (accessed on 25 June 2022).
35. 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](#)]
36. 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*, 427–433. [[CrossRef](#)] [[PubMed](#)]
37. Reich, D.; Price, A.L.; Patterson, N. Principal component analysis of genetic data. *Nat. Genet.* **2008**, *40*, 491–492. [[CrossRef](#)]
38. Wang, Q.; Tian, F.; Pan, Y.; Buckler, E.S.; Zhang, Z. A SUPER powerful method for genome wide association study. *PLoS ONE* **2014**, *9*, e107684. [[CrossRef](#)]
39. Duggal, P.; Gillanders, E.M.; Holmes, T.N.; Bailey-Wilson, J.E. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genom.* **2008**, *9*, 516. [[CrossRef](#)] [[PubMed](#)]
40. Lili-yin, A. Package 'CMplot: Circle Manhattan Plot'. 2022. Available online: <https://github.com/YinLiLin/CMplot> (accessed on 30 June 2022).
41. Sanchez, M.P.; Govignon-Gion, A.; Croiseau, P.; Fritz, S.; Hozé, C.; Miranda, G.; Martin, P.; Barbat-Leterrier, A.; Letaïef, R.; Rocha, D.; et al. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genet. Sel. Evol.* **2017**, *49*, 68. [[CrossRef](#)]
42. Mota, L.F.M.; Lopes, F.B.; Fernandes Júnior, G.A.; Rosa, G.J.M.; Magalhães, A.F.B.; Carvalheiro, R.; Albuquerque, L.G. Genome-wide scan highlights the role of candidate genes on phenotypic plasticity for age at first calving in Nellore heifers. *Sci. Rep.* **2020**, *10*, 6481. [[CrossRef](#)]
43. Navarro Gonzalez, J.; Zweig, A.S.; Speir, M.L.; Schmelter, D.; Rosenbloom, K.R.; Raney, B.J.; Powell, C.C.; Nassar, L.R.; Maulding, N.D.; Lee, C.M.; et al. The UCSC genome browser database: 2021 update. *Nucleic Acids Res.* **2021**, *49*, D1046–D1057. [[CrossRef](#)]
44. Huang, D.W.; Sherman, B.T.; Tan, Q.; Collins, J.R.; Alvord, W.G.; Roayaei, J.; Stephens, R.; Baseler, M.W.; Lane, H.C.; Lempicki, R.A. The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* **2007**, *8*, R183. [[CrossRef](#)] [[PubMed](#)]
45. Carbon, S.; Dietze, H.; Lewis, S.E.; Mungall, C.J.; Munoz-Torres, M.C.; Basu, S.; Chisholm, R.L.; Dodson, R.J.; Fey, P.; Thomas, P.D.; et al. Expansion of the gene ontology knowledgebase and resources: The gene ontology consortium. *Nucleic Acids Res.* **2017**, *45*, D331–D338. [[CrossRef](#)]
46. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [[CrossRef](#)] [[PubMed](#)]
47. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
48. Heberle, H.; Meirelles, V.G.; da Silva, F.R.; Telles, G.P.; Minghim, R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform.* **2015**, *16*, 169. [[CrossRef](#)]
49. Wu, X.; Fang, M.; Liu, L.; Wang, S.; Liu, J.; Ding, X.; Zhang, S.; Zhang, Q.; Zhang, Y.; Qiao, L.; et al. Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. *BMC Genom.* **2013**, *14*, 897. [[CrossRef](#)]

50. DeGroot, B.J.; Keown, J.F.; Van Vleck, L.D.; Marotz, E.L. Genetic parameters and responses of linear type, yield traits, and somatic cell scores to divergent selection for predicted transmitting ability for type in Holsteins. *J. Dairy Sci.* **2002**, *85*, 1578–1585. [[CrossRef](#)]
51. Royal, M.D.; Pryce, J.E.; Woolliams, J.A.; Flint, A.P.F. The genetic relationship between commencement of luteal activity and calving interval, body condition score, production, and linear type traits in holstein-friesian dairy cattle. *J. Dairy Sci.* **2002**, *85*, 3071–3080. [[CrossRef](#)]
52. Wang, D.; Sun, Y.; Stang, P.; Berlin, J.A.; Wilcox, M.A.; Li, Q. Comparison of methods for correcting population stratification in a genome-wide association study of rheumatoid arthritis: Principal-component analysis versus multidimensional scaling. *BMC Proc.* **2009**, *3*, S109. [[CrossRef](#)]
53. Lin, S.; Ke, C.; Liu, L.; Gao, Y.; Xu, L.; Han, B.; Zhao, Y.; Zhang, S.; Sun, D. Genome-wide association studies for immunoglobulin concentrations in colostrum and serum in Chinese Holstein. *BMC Genom.* **2022**, *23*, 840. [[CrossRef](#)] [[PubMed](#)]
54. Price, A.L.; Zaitlen, N.A.; Reich, D.; Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **2010**, *11*, 459–463. [[CrossRef](#)] [[PubMed](#)]
55. Devlin, B.; Roeder, K. Genomic control for association studies. *Biometrics* **1999**, *55*, 997–1004. [[CrossRef](#)] [[PubMed](#)]
56. Haldar, T.; Ghosh, S. Effect of Population Stratification on False Positive Rates of Population-Based Association Analyses of Quantitative Traits. *Ann. Hum. Genet.* **2012**, *76*, 237–245. [[CrossRef](#)] [[PubMed](#)]
57. 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](#)]
58. Sharma, A.; Lee, J.S.; Dang, C.G.; Sudrajat, P.; Kim, H.C.; Yeon, S.H.; Kang, H.S.; Lee, S.H. Stories and challenges of genome wide association studies in livestock—A review. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 1371–1379. [[CrossRef](#)]
59. Scales, S.J.; Hesser, B.A.; Masuda, E.S.; Scheller, R.H. Amisyn, a novel syntaxin-binding protein that may regulate SNARE complex assembly. *J. Biol. Chem.* **2002**, *277*, 28271–28279. [[CrossRef](#)]
60. Krintel, S.B.; Essioux, L.; Wool, A.; Johansen, J.S.; Schreiber, E.; Zekharya, T.; Akiva, P.; Østergaard, M.; Hetland, M.L. CD6 and syntaxin binding protein 6 variants and response to tumor necrosis factor alpha inhibitors in danish patients with rheumatoid arthritis. *PLoS ONE* **2012**, *7*, e38539. [[CrossRef](#)]
61. Lenka, G.; Shan, J.; Halabi, N.; Abuaqel, S.W.J.; Goswami, N.; Schmidt, F.; Zaghlool, S.; Romero, A.R.; Subramanian, M.; Boujassoum, S.; et al. STXBP6, reciprocally regulated with autophagy, reduces triple negative breast cancer aggressiveness. *Clin. Transl. Med.* **2020**, *10*, e147. [[CrossRef](#)]
62. Räschle, M.; Smeenk, G.; Hansen, R.K.; Temu, T.; Oka, Y.; Hein, M.Y.; Nagaraj, N.; Long, D.T.; Walter, J.C.; Hofmann, K.; et al. Proteomics reveals dynamic assembly of Repair complexes during bypass of DNA cross-links. *Science* **2015**, *348*, 1253671. [[CrossRef](#)]
63. Sanchez, M.P.; Ramayo-Caldas, Y.; Wolf, V.; Laithier, C.; El Jabri, M.; Michenet, A.; Boussaha, M.; Taussat, S.; Fritz, S.; Delacroix-Buchet, A.; et al. Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. *Genet. Sel. Evol.* **2019**, *51*, 34. [[CrossRef](#)] [[PubMed](#)]
64. Albrecht, E.; Komolka, K.; Ponsuksili, S.; Gotoh, T.; Wimmers, K.; Maak, S. Transcriptome profiling of Musculus longissimus dorsi in two cattle breeds with different intramuscular fat deposition. *Genomics Data* **2016**, *7*, 109–111. [[CrossRef](#)] [[PubMed](#)]
65. Sun, T.; Huang, G.Y.; Wang, Z.H.; Teng, S.H.; Cao, Y.H.; Sun, J.L.; Hanif, Q.; Chen, N.B.; Lei, C.Z.; Liao, Y.Y. Selection signatures of Fuzhong Buffalo based on whole-genome sequences. *BMC Genom.* **2020**, *21*, 674. [[CrossRef](#)]
66. Wang, Z.; Ma, H.; Xu, L.; Zhu, B.; Liu, Y.; Bordbar, F.; Chen, Y.; Zhang, L.; Gao, X.; Gao, H.; et al. Genome-wide scan identifies selection signatures in chinese wagyu cattle using a high-density SNP array. *Animals* **2019**, *9*, 296. [[CrossRef](#)] [[PubMed](#)]
67. Braz, C.U.; Rowan, T.N.; Schnabel, R.D.; Decker, J.E. Genome-wide association analyses identify genotype-by-environment interactions of growth traits in Simmental cattle. *Sci. Rep.* **2021**, *11*, 13335. [[CrossRef](#)]
68. Christensen, J.; Cloos, P.; Toftegaard, U.; Klinkenberg, D.; Bracken, A.P.; Trinh, E.; Heeran, M.; Di Stefano, L.; Helin, K. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Res.* **2005**, *33*, 5458–5470. [[CrossRef](#)] [[PubMed](#)]
69. Morrell, B.C.; Zhang, L.; Schütz, L.F.; Perego, M.C.; Maylem, E.R.S.; Spicer, L.J. Regulation of the transcription factor E2F8 gene expression in bovine ovarian cells. *Mol. Cell. Endocrinol.* **2019**, *498*, 110572. [[CrossRef](#)]
70. Yang, B.; Jiao, B.; Ge, W.; Zhang, X.; Wang, S.; Zhao, H.; Wang, X. Transcriptome sequencing to detect the potential role of long non-coding RNAs in bovine mammary gland during the dry and lactation period. *BMC Genom.* **2018**, *19*, 605. [[CrossRef](#)]
71. Bakhtiarizadeh, M.R.; Mirzaei, S.; Norouzi, M.; Sheybani, N.; Vafaei Sadi, M.S. Identification of Gene Modules and Hub Genes Involved in Mastitis Development Using a Systems Biology Approach. *Front. Genet.* **2020**, *11*, 722. [[CrossRef](#)]
72. Zhu, B.; Niu, H.; Zhang, W.; Wang, Z.; Liang, Y.; Guan, L.; Guo, P.; Chen, Y.; Zhang, L.; Guo, Y.; et al. Genome wide association study and genomic prediction for fatty acid composition in Chinese Simmental beef cattle using high density SNP array. *BMC Genom.* **2017**, *18*, 464. [[CrossRef](#)]
73. Buaban, S.; Lengnudum, K.; Boonkum, W.; Phakdeedindan, P. Genome-wide association study on milk production and somatic cell score for Thai dairy cattle using weighted single-step approach with random regression test-day model. *J. Dairy Sci.* **2022**, *105*, 468–494. [[CrossRef](#)]
74. Cai, Z.; Dusza, M.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. *Genet. Sel. Evol.* **2020**, *52*, 19. [[CrossRef](#)] [[PubMed](#)]

75. Freebern, E.; Santos, D.J.A.; Fang, L.; Jiang, J.; Parker Gaddis, K.L.; Liu, G.E.; Vanraden, P.M.; Maltecca, C.; Cole, J.B.; Ma, L. GWAS and fine-mapping of livability and six disease traits in holstein cattle. *BMC Genom.* **2022**, *21*, 41. [[CrossRef](#)] [[PubMed](#)]
76. de Souza Lung, L.H.; Mulder, H.A.; de Rezende Neves, H.H.; Carvalheiro, R. Genomic regions underlying uniformity of yearling weight in Nellore cattle evaluated under different response variables. *BMC Genom.* **2018**, *19*, 619. [[CrossRef](#)] [[PubMed](#)]
77. Han, L.Q.; Zhou, Z.; Ma, Y.; Batistel, F.; Osorio, J.S.; Looor, J.J. Phosphorylation of nuclear factor erythroid 2-like 2 (NFE2L2) in mammary tissue of Holstein cows during the periparturient period is associated with mRNA abundance of antioxidant gene networks. *J. Dairy Sci.* **2018**, *101*, 6511–6522. [[CrossRef](#)]
78. Li, P.; Fei, H.; Wang, L.; Xu, H.; Zhang, H.; Zheng, L. Pdc5 regulates cell proliferation, cell cycle progression and apoptosis. *Oncol. Lett.* **2018**, *15*, 1177–1183. [[CrossRef](#)]
79. Casey, T.; Patel, O.V.; Plaut, K. Transcriptomes reveal alterations in gravity impact circadian clocks and activate mechanotransduction pathways with adaptation through epigenetic change. *Physiol. Genom.* **2015**, *47*, 113–128. [[CrossRef](#)]
80. Xu, Q.; Mei, G.; Sun, D.; Zhang, Q.; Zhang, Y.; Yin, C.; Chen, H.; Ding, X.; Liu, J. Detection of genetic association and functional polymorphisms of UGDH affecting milk production trait in Chinese Holstein cattle. *BMC Genom.* **2012**, *13*, 661. [[CrossRef](#)]
81. Libera, K.; Konieczny, K.; Witkowska, K.; Żurek, K.; Szumacher-Strabel, M.; Cieslak, A.; Smulski, S. The association between selected dietary minerals and mastitis in dairy cows—A review. *Animals* **2021**, *11*, 2330. [[CrossRef](#)]
82. Necasova, A.; Pechova, A.; Bodor, R.; Masar, M. Evaluation of the glutathione concentration in whole blood of dairy Holstein cows. *Vet. Med.* **2021**, *66*, 179–188. [[CrossRef](#)]
83. Moro, J.; Tome, D.; Schmidely, P.; Demersay, T.; Azzout-Marniche, D. Histidine: A Systematic Review on Metabolism and. *Nutrients* **2020**, *12*, 1414. [[CrossRef](#)] [[PubMed](#)]
84. Akers, R.M. Major advances associated with hormone and growth factor regulation of mammary growth and lactation in dairy cows. *J. Dairy Sci.* **2006**, *89*, 1222–1234. [[CrossRef](#)]