

Article

Genome-Wide Association Study Identifies Candidate Genes Associated with Feet and Leg Conformation Traits in Chinese Holstein Cattle

Ismail Mohamed Abdalla ¹, Xubin Lu ¹ , Mudasir Nazar ¹, Abdelaziz Adam Idriss Arbab ^{1,2}, Tianle Xu ³ , Mohammed Husien Yousif ⁴, Yongjiang Mao ¹ and Zhangping Yang ^{1,*} 

- ¹ College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China; ismailhmk@gmail.com (I.M.A.); dx120180094@yzu.edu.cn (X.L.); dr mudasirnazar457@gmail.com (M.N.); arbabor@yahoo.com (A.A.I.A.); cattle@yzu.edu.cn (Y.M.)
² Biomedical Research Institute, Darfur College, Nyala 63313, Sudan
³ Joint International Research Laboratory of Agriculture and Agri-Product Safety, Yangzhou University, Yangzhou 225009, China; tl-xu@outlook.com
⁴ Faculty of Animal Production, West Kordufan University, Alnuhud City 12942, Sudan; mohammedago@hotmail.com
* Correspondence: yzp@yzu.edu.cn; Tel.: +86-0514-87979269

Simple Summary: Feet and leg problems are among the major reasons for dairy cows leaving the herd, as well as having direct association with production and reproduction efficiency, health (e.g., claw disorders and lameness) and welfare. Hence, understanding the genetic architecture underlying feet and conformation traits in dairy cattle offers new opportunities toward the genetic improvement and long-term selection. Through a genome-wide association study on Chinese Holstein cattle, we identified several candidate genes associated with feet and leg conformation traits. These results could provide useful information about the molecular breeding basis of feet and leg traits, thus improving the longevity and productivity of dairy cattle.



Citation: Abdalla, I.M.; Lu, X.; Nazar, M.; Arbab, A.A.I.; Xu, T.; Yousif, M.H.; Mao, Y.; Yang, Z. Genome-Wide Association Study Identifies Candidate Genes Associated with Feet and Leg Conformation Traits in Chinese Holstein Cattle. *Animals* **2021**, *11*, 2259. <https://doi.org/10.3390/ani11082259>

Received: 8 June 2021
Accepted: 28 July 2021
Published: 30 July 2021

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



Copyright: © 2021 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/>).

Abstract: Feet and leg conformation traits are considered one of the most important economical traits in dairy cattle and have a great impact on the profitability of milk production. Therefore, identifying the single nucleotide polymorphisms (SNPs), genes and pathways analysis associated with these traits might contribute to the genomic selection and long-term plan selection for dairy cattle. We conducted genome-wide association studies (GWASs) using the fixed and random model circulating probability unification (FarmCPU) method to identify SNPs associated with bone quality, heel depth, rear leg side view and rear leg rear view of Chinese Holstein cows. Phenotypic measurements were collected from 1000 individuals of Chinese Holstein cattle and the GeneSeek Genomic Profiler Bovine 100 K SNP chip was utilized for individual genotyping. After quality control, 984 individual cows and 84,906 SNPs remained for GWAS work; as a result, we identified 20 significant SNPs after Bonferroni correction. Several candidate genes were identified within distances of 200 kb upstream or downstream to the significant SNPs, including *ADIPOR2*, *INPP4A*, *DNMT3A*, *ALDH1A2*, *PCDH7*, *XKR4* and *CADPS*. Further bioinformatics analyses showed 34 gene ontology terms and two signaling pathways were significantly enriched ($p \leq 0.05$). Many terms and pathways are related to biological quality, metabolism and development processes; these identified SNPs and genes could provide useful information about the genetic architecture of feet and leg traits, thus improving the longevity and productivity of Chinese Holstein dairy cattle.

Keywords: feet and leg; GWAS; FarmCPU; SNP; functional analysis; Chinese Holstein cow

1. Introduction

Body conformation traits are considered economically important traits in dairy cattle [1]. Improving the accuracy of selection for body conformation traits would enhance

the dairy industry as a whole and greatly affect the profitability of individual farms, which will automatically lead to the future profitability of the Chinese dairy industry. Thus, these traits can be used as an indirect predictor of economically important traits for an animal.

Adequate locomotion of an animal is directly associated with production and reproduction efficiency, as well as health and welfare. If an animal has feet and leg problems and bad mobility, it will be difficult to get access to feed and the milking parlor. Consequently, its longevity, welfare, productivity and reproductive performance will be affected [2] and, at most, culled before their third parity, due to such leg problems [3,4]. Several studies have reported significant genetic correlation between feet and leg conformation traits and their health problems (e.g., claw disorders and lameness), ranging between low, moderate and high, indicating that it is useful to use feet and leg conformation traits as an indirect selection criterion to control their health problems [5,6].

Over the last decades, genome-wide association studies (GWASs) have become a powerful tool in the search of potential genetic markers across the genome in dairy cattle, due to the availability of a high-density chip with single nucleotide polymorphisms (SNPs) for bovine. Phenotypic measurements and pedigree information, along with sequence variations (mainly single nucleotide polymorphisms, SNPs) in the whole genome, serve as the basic pillars for genome-wide association studies (GWAS). Altogether, they identify genes or regulatory elements necessary for each trait of interest [7].

In the last decade, various GWAS studies have been undertaken on Chinese Holstein cattle, involving milk production traits [8,9] milk composition traits [10,11], fatty acids [12,13], protein [14], body conformation traits [15,16], body size [17], pigmentation [18], fertility [19], concentration of albumin in colostrum and serum [20] and mastitis [21,22].

Regarding the traits under investigation in this paper, some studies have been conducted which identified many QTL and genes associated with feet and legs conformation traits in dairy cattle using the GWAS method [16,23–25].

Therefore, the objectives, in our study, were to detect significant single nucleotide polymorphisms (SNPs) associated with feet and legs conformation traits (heel depth, Bone quality, rear leg side view and rear leg rear view) in Chinese Holstein cattle using the genome-wide association (GWAS) approach and using the SNPs positions to identify candidate genes and pathways that may influence these traits. In addition, we will use the identified genes to conduct some bioinformatics analyses, such as gene ontology, KEGG pathway and gene network analysis for further analysis.

2. Materials and Methods

2.1. Ethics Statement

All procedures for collecting hair follicle samples and measuring phenotypic traits were carried out strictly in accordance with the guidelines proposed by the China Council on Animal Care and Ministry of Agriculture of the People's Republic of China. The study was also approved by the Institutional Animal Care and Use Committee of School of the Yangzhou University Animal Experiments Ethics Committee (License Number: SYXK (Su) IACUC 2012-0029), Yangzhou University.

2.2. Body Measurements and DNA Samples Collection

The experimental population consisted of 1000 Chinese Holstein cows raised on four farms, (199 individuals from Sihong Farm, 214 from Xuyi Farm, 224 from Xuzhou Farm and 363 from Huaxia Farm). All these farms are located in the northern part of Jiangsu province, China. Phenotypic descriptive traits (BQ, bone quality; HD, heel depth; RLSV, rear leg side view; RLRV, rear leg rear view) were measured individually and scored on a 1–9 scale score according to the National Standards of People's Republic of China Code of practice of type classification in Chinese Holstein (GBT35568-2017). The measurement of traits for each cow were performed by three trained technicians and the average of the measurements taken by the different professionals was used as the phenotype for each

trait to ensure the accuracy of the data. For the genotyping analysis, hair follicle samples were collected individually and stored in special paper envelopes for each cow to prevent it from environmental and DNA contamination; each paper envelope must contain enough hair samples, not less than 50 hairs.

2.3. Phenotypic and Genetic Parameters

Descriptive statistics and pairwise Pearson correlation coefficients of phenotypic traits were determined using the computer-based software IBM-SPSS, version 25.

Genetic analysis was carried out using the Derivative-free approach to MULTivariate analysis (DMU) software [26] to estimate heritability and genetic correlation between pairs of traits with the animal model as follows:

$$y_{ijklm} = u + \text{Herd}_i + \text{Year}_j + \text{Season}_k + \text{Parity}_l + a_m + e_{ijklm}$$

where y_{ijklm} is the phenotype in the j th year, k th season and l th parity of the m th individual from the i th herd; u is overall mean of the population, Herd_i is the herd effect according to a cow's origin from one of the four herds; Year_j is the j th year effect, Season_k is the k th season effect and parity is the effect of l th parity; a is the additive effect of the m th individual, which was evaluated by the pedigree information, and e is the residual in the j th year, k th season and l th parity of the m th individual from the i th herd. All effects were treated as random except the overall mean. The pedigree of the cows could be traced back at least three generations (2009–2020), the parities of cows were between 1 and 4 and four seasons of were defined, December–February, March–May, June–August and September–November.

2.4. Genotyping and Quality Control

Each individual from the experimental population was genotyped using the GeneSeek Genomic Profiler Bovine 100 k SNP chip (Neogen Corporation, <http://www.neogenchina.com.cn/> (accessed on 28 June 2020)) based on ARC-UCD1.2/bosTau9 as the genome reference. Genomic DNA was extracted from hair follicle samples. The GeneSeek Genomic Profiler Bovine 100 K SNP chip containing 100,000 SNPs was utilized for individual genotyping.

Quality control was conducted by using the Plink 1.90 software [27] to remove the markers which did not comply with the following criteria: (1) individual call rate lower than 95%; (2) genotype call rate of a single SNP lower than 90%; (3) minor allele frequency (MAF) > 0.05; (4) deviated from Hardy–Weinberg equilibrium ($p < 1.0 \times 10^{-6}$). After quality control, 984 cows and 84,406 markers remained for further analyses (Figure S1).

Marker intervals and linkage disequilibrium (LD) were calculated to estimate R square for all markers and plotted the marker distribution as showed in Figure 1d. LD decay fell off quickly within 100 kb physical distance, then decreased slowly afterward.

2.5. Population Structure Analysis

We used the Plink 1.90 software [27] to implement principal component analysis (PCA) on 1000 cows genotyping with 84,406 SNPs in Chinese Holstein cattle herds raised at four breeding herds to investigate the population structure and PCA was plotted using the ggplot package In R 4.0.2. The admixture software [28] was also used to study the population structure and correct population stratification.

2.6. Genome-Wide Association Studies

In this study, the fixed and random model circulating probability unification (FarmCPU) method was used to carry out the genome-wide association analysis [29]. The FarmCPU method uses a fixed effect model and a random effect model iteratively. In GWASs, population stratification is the main cause of false positive correlations. Therefore, the fixed effect model tests SNPs one at a time. The significant SNPs are evaluated in the

random effect model and the validated SNPs are fitted as covariates in the fixed effect model to control population structure. The model can be written as follows:

$$y_i = M_{i1}b_1 + M_{i2}b_2 + \dots + M_{it}b_t + S_{ij}d_j + e_i \quad (1)$$

where y_i is the observation of the i th individual; $M_{i1}, M_{i2} \dots M_{it}$ are the genotypes of t pseudo QTNs, initiated as an empty set; b_1, b_2, \dots, b_t are the corresponding effects of the pseudo QTNs; S_{ij} is the genotype of the i th individual and j th genetic marker; d_j is the corresponding effect of the j th genetic marker; e_i is the residual having a distribution with zero mean and variance of σ^2_e .

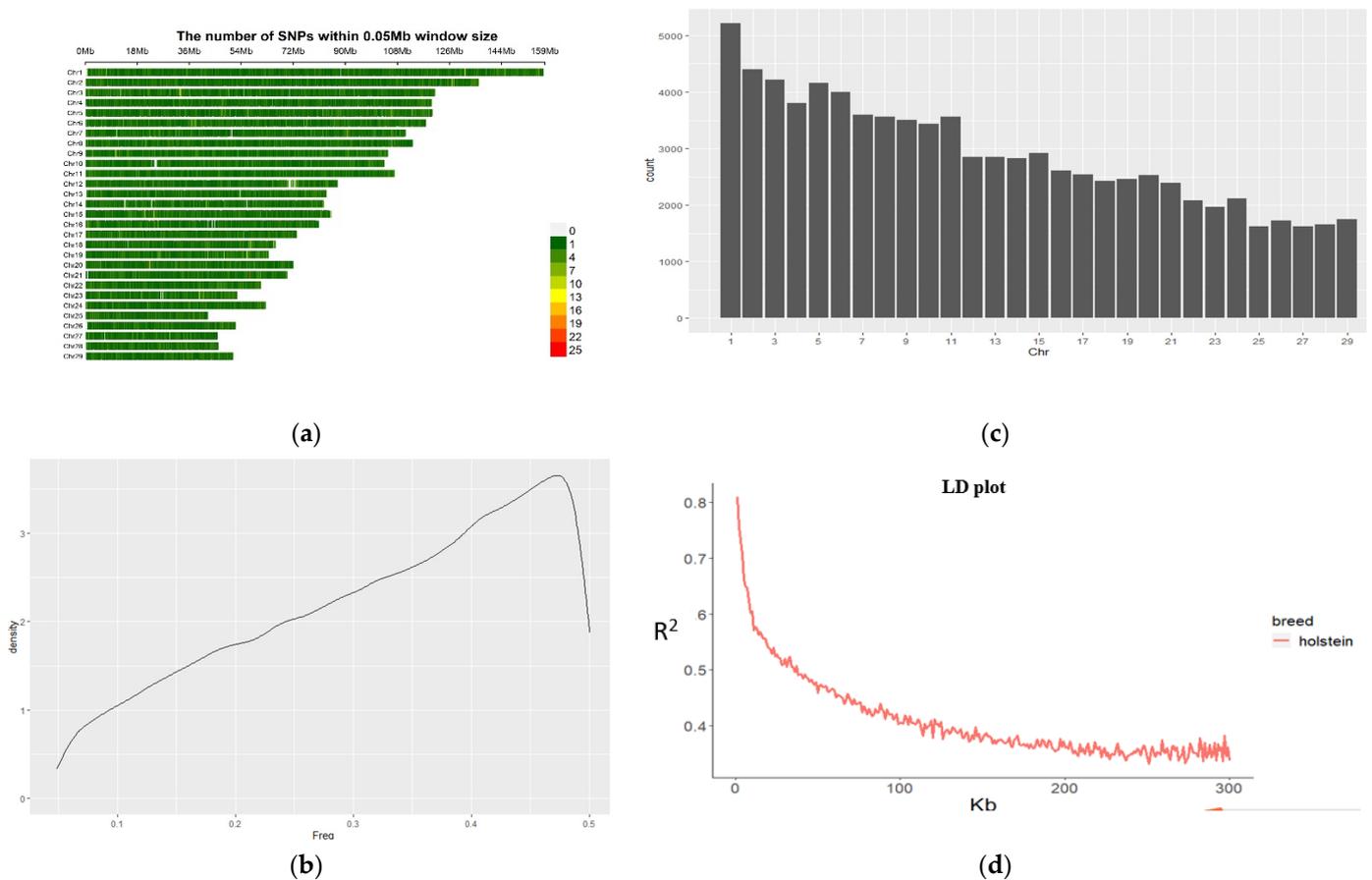


Figure 1. Properties of single nucleotide polymorphisms (SNPs). In total, 1000 individuals were genotyped by the GeneSeek Genomic Profiler Bovine 100 k bead chip; 84,609 SNPs and 984 cattle passed filters and quality control. Marker distributions are displayed in a heatmap on 29 chromosomes by minor allele frequency (MAF) (a). MAF was re-calculated after quality control. Therefore, some SNPs remained with MAFs lartherthan 0.05, as shown by the histogram (b). Marker density is displayed by the histogram on chromosome 29 (c). LD decay is shown by scatter plot according to pairwise distance and trend as a red line (d).

After substitution, every marker has its own p value. The p values and the associated marker map are used to update the selection of pseudo QTNs using the SUPER algorithm (Settlement of MLM Under Progressively Exclusive Relationship) [30] in a REM as follows:

$$y_i = u_i + e_i \quad (2)$$

where y_i and e_i are the same as in Equation (1) and u_i is the total genetic effect of the i th individual.

The total type 1 error (false positive) rate was controlled at 5% and the significance threshold of the GWAS was determined according to this formula $(0.05/N_{\text{snp}})$, where N_{snp}

is the number of SNPs remaining after quality control [31]. Subsequently, the significant threshold for the GWAS was 5.9×10^{-7} (0.05/84,406) after Bonferroni correction.

Quantile–quantile (Q–Q) and Manhattan plots were drawn using the CMplot package in the R 3.1.1 software [32].

2.7. Gene Identification

Genomic regions and candidate genes were recognized on the livestock using a genome browser (UCSC) through an Asian server for cow assembly April 2018 (ARC-UCD1.2/bosTau9). <https://genomeasia.ucsc.edu/cgi-bin/hgGateway?redirect=manual&source=www.genome.ucsc.edu> (accessed on 25 May 2021) and full NCBI (the National Center for Biotechnology Information Gene) database (<http://www.ncbi.nlm.nih.gov/gene/> (accessed on 25 May 2021)).

2.8. Function and Pathway Enrichment and Network Analysis

In our study, we submitted the candidate genes obtained by GWAS into the Database for Annotation, Visualization and Integrated Discovery (DAVID) [33] for the Gene Ontology (GO) terms [34] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis [35]. The statistically significant p value for functional analysis and pathway analysis was defined at $p \leq 0.05$. Protein–protein interactions among genes were performed using the online Search Tool for the Retrieval of Interacting Genes (STRING) database v11.0 [36] with the Cytoscape software.v3.8.2 to visualize the resultant PPI network.

3. Results

3.1. Descriptive Statistics and Heritability Estimation of Feet and Legs Traits

The feet and leg traits for the 1000 Chinese Holstein cows from four farms measured in this study included heel depth (HD), bone quality (BQ), rear leg side view (RLSV) and the rear leg rear view (RLRV). The descriptive statistics (mean, maximum and minimum values and standard deviation) of phenotypic measurements for these traits are showed in Table 1, where the mean of the HD trait had the greater score (7.04) and the RLSV trait revealed the lowest score, with 3.92. The pairwise genetic and Pearson phenotypic correlation between the traits are provided in Table 2. BQ phenotypically positively correlated with RLSV, and the other traits were low or moderate negatively correlated with RLSV. However, BQ was genetically positively correlated with HD, RLSV and RLRV. HD was negatively correlated with all traits, both genetically and phenotypically, except BQ had a highly positive genetic correlation. RLSV was highly positively genetically correlated with BQ, while it was negative with HD and RLRV, whereas it had a negative phenotypic correlation with BQ, HD and RLRV. Estimates of the heritability results were 0.15, 0.05, 0.17 and 0.15 for heel depth, bone quality, rear leg side view and rear leg rear view, respectively (Table 2).

Table 1. Descriptive statistics of feet and leg conformation traits measurements.

Traits	Mean	Std. Error	Minimum	Maximum	Std. Deviation
HD	7.04	0.051	2.00	9.00	1.61
BQ	6.03	0.028	2.00	9.00	0.87
RLSV	3.92	0.048	1.00	9.00	1.50
RLRV	5.60	0.045	2.00	9.00	1.57

HD, heel depth; BQ, bone quality; RLSV, rear leg side view; RLRV, rear leg rear view.

Table 2. Genetic (upper diagonal), Pearson phenotypic (lower diagonal) correlations and heritability (grey diagonal line) for feet and leg traits measurements.

Traits	HD	BQ	RLSV	RLRV
HD	0.15	0.67	−0.03	−0.36
BQ	−0.08 **	0.05	0.84	0.10
RLSV	−0.07 *	−0.08 *	0.17	−0.09
RLRV	−0.09 **	0.1 **	−0.39 **	0.15

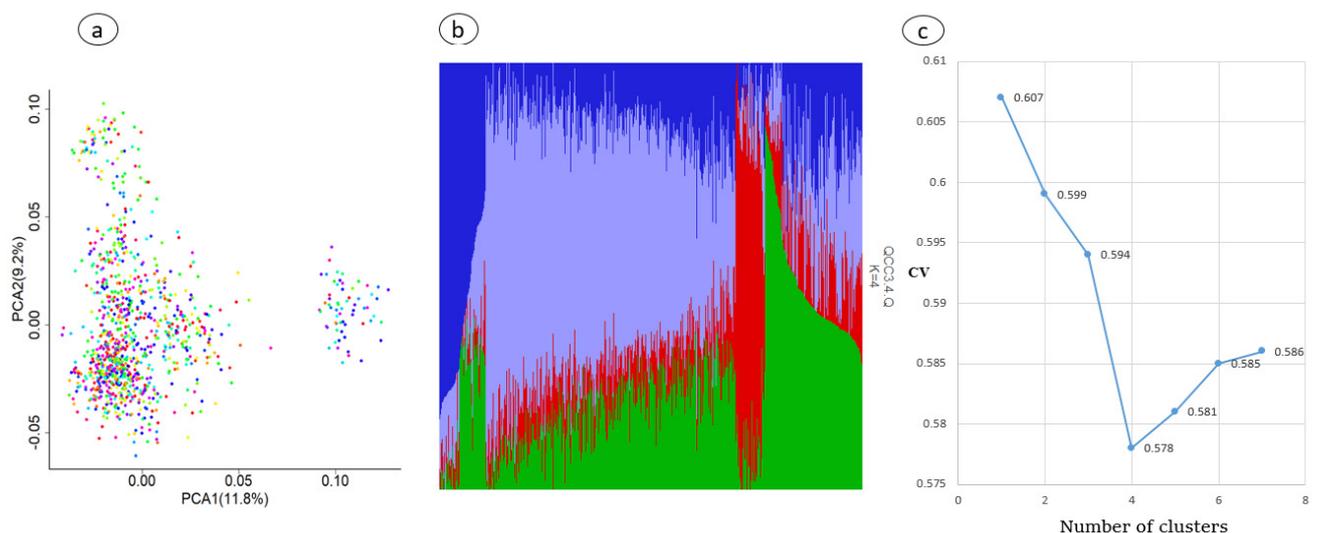
HD, heel depth; BQ, bone quality; RLSV, rear leg side view; RLRV, rear leg rear view. The upper subscript * and ** represent significant correlation at the 0.05 and 0.01, respectively.

3.2. Information of SNPs

After quality control, 984 individual cows and 84,406 SNPs were used to conduct the GWAS. The filtered SNPs were distributed on all 29 chromosomes. Chromosome 1 has shown the greatest number of SNPs, whereas chromosome 25 contained the fewest. The minor allele frequency (MAF) for all SNPs was re-calculated after quality control; only MAF above 5% remained. The LD decay line dramatically decreased at the beginning then tended to be slow after 100 kb distance. (Figure 1).

3.3. Population Stratification

Principal component analysis (PCA) and the ADMIXTURE program were used to visualize family structure in this study. The results reveal all individuals were grouped into two unequal sized clusters, as shown in Figure 2a,b; as for the admixture population, using ADMIXTURE with k value ranging from 1 to 7 (Figure 2c), based on the cross-validation errors, K = 4 was identified to be the optimal number of genetic clusters defining the population structure among the four herds of Chinese Holstein cattle. Also, the results show that the first two principal components were about 21% of the variation (Figure 2a), and they were fitted as covariate variables in the association analysis using the FarmCPU model. Population stratification based on the PCA analysis results was considered and incorporated into the mixed linear model.

**Figure 2.** Population structure for experimental population (1000 cows) raised at four farms. (a) Principal component analysis. (b) Admixture analysis. (c) Cross-validation errors across 21 ancestral genetic clusters.

3.4. GWAS Results

The FarmCPU model was used to conduct the genome-wide association analysis in the present study. In Figure 3, The quantile–quantile (Q–Q) plots illustrated the model used in this study for GWAS analysis was reasonable. The lambda (inflation factor (λ))

was close to 1 ($0.9 < 1.01$) and the point at the upper right corner of (Q–Q) plots are the significant markers associated with the traits under study (Figure 3); therefore, the population stratification was adequately controlled. Meanwhile, Manhattan plots are used to visualize GWAS significance level ($-\log_{10}$ of p value of each SNP) by chromosome location (Figure 4).

In the present study, 20 SNPs (Table 3) were past the threshold and significantly associated with four traits of feet and legs (heel depth, bone quality, rear leg side view and rear leg rear view).

Table 3. Genome-wide significant SNP associated with feet and leg traits and nearest candidate genes.

Traits	SNP Name	rs. SNP Name	Chr.	Position	MAF	p Value	Nearest Gene Name	Distance
Bone quality	Hapmap54208-rs29015846	rs29015846	8	55190206	0.29	1.99×10^{-7}	<i>LOC112447952</i>	100 kb
	BovineHD0400024774	rs133088614	4	88636425	0.46	2.25×10^{-7}	<i>TMEM229A</i>	100 kb
	ARS-BFGL-NGS-81828	rs41845981	17	44543822	0.35	3.14×10^{-7}	<i>POLE</i>	within
	BovineHD1400007035	rs136017102	14	22608072	0.13	6.22×10^{-7}	<i>XKR4</i>	50 kb
	BovineHD2200011035	rs110949452	22	38556301	0.34	7.67×10^{-7}	<i>CADPS</i>	within
Heel depth	BovineHD2000006450	rs137022628	20	21510411	0.27	3.03×10^{-9}	<i>ACTBL2</i>	200 kb
	ARS-BFGL-BAC-36389	rs109652453	23	45295465	0.37	4.22×10^{-8}	<i>SYCP2L</i>	within
	BovineHD2700002886	rs42110372	27	10498902	0.10	3.11×10^{-7}	<i>LOC112444670</i>	100 kb
	ARS-BFGL-NGS-116157	rs109601642	20	16333125	0.27	6.46×10^{-7}	<i>LOC101907219</i>	100 kb
	ARS-BFGL-NGS-73835	rs41577664	15	80365537	0.48	7.43×10^{-7}	<i>LOC112441589</i>	200 bp
	BovineHD1400023839	rs134726669	14	81745515	0.31	7.59×10^{-7}	<i>MRPL13</i>	200 bp
Rear legs rear view	BovineHD0300015960	rs134130409	3	52695086	0.41	6.72×10^{-8}	<i>BARHL2</i>	100 kb
	BovineHD4100014792	rs134139959	20	57471508	0.19	1.11×10^{-7}	<i>FBXL7</i>	within
	Hapmap48798-BTA-51401	rs41638134	1	131310669	0.35	6.11×10^{-7}	<i>LOC107132214</i>	20 kb
Rear legs side view	ARS-BFGL-NGS-91167	rs41565304	5	108222067	0.19	1.11×10^{-9}	<i>ADIPOR2</i>	within
	BovineHD0300022142	rs43350216	3	76168946	0.38	4.05×10^{-9}	-	-
	BovineHD1100001372	rs43656945	11	3761839	0.07	2.32×10^{-7}	<i>INPP4A</i>	within
	BovineHD1100021155	rs136593856	11	74048825	0.44	5.07×10^{-7}	<i>DNMT3A</i>	within
	BTB-01677645	rs42791722	10	52217450	0.09	7.40×10^{-7}	<i>ALDH1A2</i>	within
	BTB-01518251	rs42639670	6	49588438	0.09	9.65×10^{-7}	<i>PCDH7</i>	200 kb

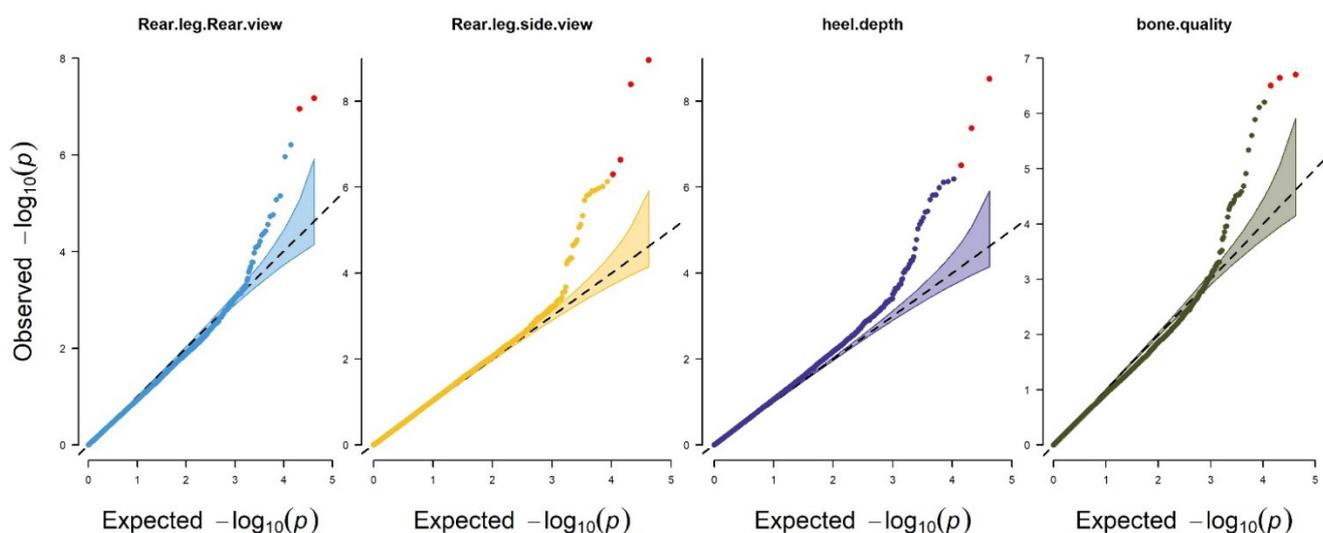


Figure 3. Quantile–quantile (Q–Q) from GWAS for feet and leg traits in Chinese Holstein cattle. Q–Q plot showing the late separation between observed and expected values. The red lines indicate the null hypothesis of no true association. Deviation from the expected p value distribution is evident only in the tail area for each trait, indicating that population stratification was properly controlled.

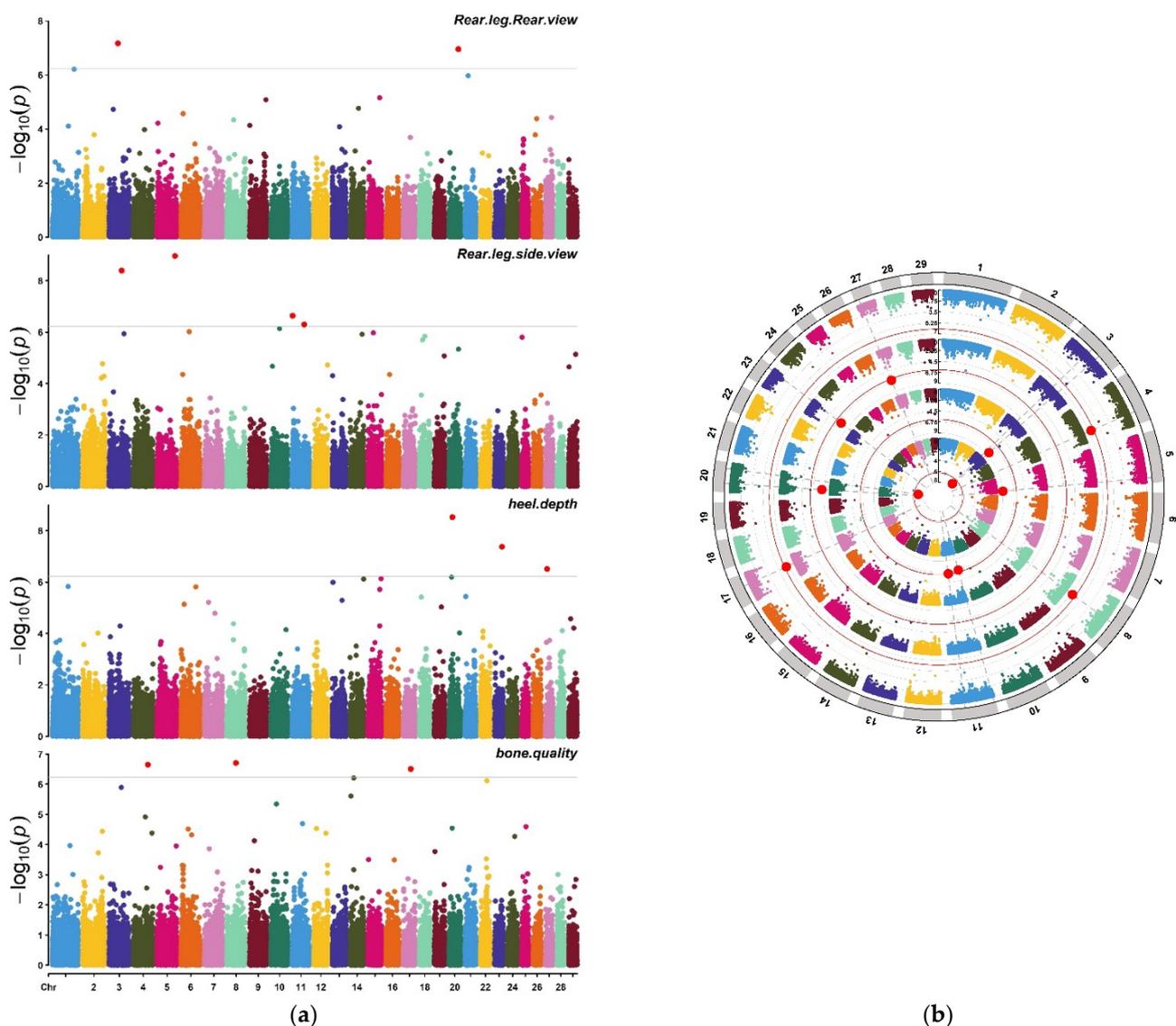


Figure 4. Manhattan from GWAS for feet and leg traits in Chinese Holstein cattle. Manhattan plot in which the genomic coordinates of SNPs are displayed along the horizontal axis, the negative logarithm of the association p value for each SNP is displayed on the vertical axis and the green line indicates the significance threshold level after Bonferroni correction (a). Circular Manhattan plot, the four feet and leg conformation traits are plotted from the outside to the inside, respectively (b).

For the bone quality trait, five SNPs were detected on the chromosomes Chr8 (Hap-map54208-rs29015846-rs29015846), Chr4 (BovineHD0400024774-rs133088614), Chr17 (ARS-BFGL-NGS-81828-rs41845981), Chr14 (BovineHD1400007035-rs136017102) and Chr22 (BovineHD2200011035-rs110949452), while six SNPs associated with heel depth trait were detected on the chromosomes Chr20 (BovineHD2000006450-rs137022628 and ARS-BFGL-NGS-116157-rs109601642), Chr23 (ARS-BFGL-BAC-36389-rs109652453), Chr27 (BovineHD2700002886-rs42110372), Chr15 (ARS-BFGL-NGS-73835-rs41577664) and Chr14 (BovineHD1400023839-rs134726669). Moreover, three significantly SNPs associated with rear leg rear view were detected on chromosomes Chr3 (BovineHD0300015960-rs134130409), Chr20 (BovineHD4100014792-rs134139959) and Chr1 (Hapmap48798-BTA-51401-rs41638134). In addition, six SNPs associated with rear leg side view were detected on the chromosomes Chr5 (ARS-BFGL-NGS-91167-rs41565304), Chr3 (BovineHD0300022142-rs43350216), Chr11 (BovineHD1100001372-rs43656945 and BovineHD1100021155-rs136593856), Chr10 (BTB-01677645-rs42791722) and Chr6 (BTB-01518251-rs42639670) (Table 3).

3.5. Identified Candidate Genes

The results of the linkage disequilibrium (LD) analysis indicate the LD decay (R^2), as showed in Figure 1d, tends to be stable when the distance is 200 kb. Therefore, genes located within this region (200 kb) of the significant SNP are defined as candidate genes.

Among all the 20 significant SNPs (Table 3), 8 of them (ARS-BFGL-NGS-81828-rs41845981, BovineHD2200011035-rs110949452, ARS-BFGL-BAC-36389-rs109652453, BovineHD4100014792-rs134139959, ARS-BFGL-NGS-91167-rs41565304, BovineHD1100001372-rs43656945, BovineHD1100021155-rs136593856 and BTB-01677645: rs42791722) are located within genes, namely, DNA polymerase epsilon (*POLE*), calcium dependent secretion activator (*CADPS*), synaptonemal complex protein 2 like (*SYCP2L*), F-box and leucine rich repeat protein 7 (*FBXL7*), adiponectin receptor 2 (*ADIPOR2*), inositol polyphosphate-4-phosphatase type I A (*INPP4A*), DNA methyltransferase 3 alpha (*DNMT3A*) and aldehyde dehydrogenase 1 family member A2 (*ALDH1A2*), respectively (Table 3).

Whilst the SNP (Hapmap54208-rs29015846-rs29015846) on Chr8 is located near (100 kb) the U6 spliceosomal RNA gene (*LOC112447952*), the SNP (BovineHD0400024774: rs133088614) on Chr4 is located close (100 kb) to the transmembrane protein 229A gene (*TMEM229A*) and the SNP (BovineHD1400007035-rs136017102) on Chr14 is located close (1 kb) to the XK related 4 gene (*XKR4*). The SNP (BovineHD2200011035-rs110949452) on Chr20 is located close (50 kb) to actin beta like 2 gene (*ACTBL2*). The SNP (BovineHD2700002886-rs42110372) on Chr27 is located close (100 kb) to the small nucleolar RNA SNORD22 gene (*LOC112444670*). The SNP (ARS-BFGL-NGS-116157-rs109601642) on Chr20 is located close (100 kb) to gene *LOC101907219*. The SNP (ARS-BFGL-NGS-73835-rs41577664) on Chr15 is located close (200 bp) to gene *LOC112441589*. The SNP (BovineHD1400023839-rs134726669) on Chr14 is located close (200 bp) to the mitochondrial ribosomal protein L13 gene (*MRPL13*). The SNP (BovineHD0300015960-rs134130409) on Chr3 is located close (100 kb) to BarH like homeobox 2 gene (*BARHL2*). The SNP (Hapmap48798-BTA-51401-rs41638134) on Chr1 is located close (2 kb) to gene *LOC107132214*. The SNP (BovineHD0300022142) on Chr3 was not harboring any gene within 200 kb. Moreover, the SNP (BTB-01518251-rs42639670) on Chr6 is located close (200 kb) to the gene protocadherin 7 (*PCDH7*).

3.6. Functional Analysis

Using the UCSC Genome browser and NCBI database through the Asian server for cow assembly April 2018 (ARC-UCD1.2/bosTau9), 105 genes were obtained within the region of 200 kb up/downstream of the significant SNPs for the feet and legs traits; these genes were used for further enrichment and pathway analysis (gene ontology and KEGG). Gene ontology enrichment analysis revealed 34 GO terms were significantly enriched ($p < 0.05$), which consist of 18 biological process terms, 1 cellular component and 15 molecular function terms (Table S1 and Figure 5).

The KEGG pathways analysis revealed that five pathways (bta00531:Glycosaminoglycan degradation involves 3 genes and the bta01100:Metabolic pathway consists of 12 genes) were enriched ($p \leq 0.05$), whereas the other three pathways (bta04020:Calcium signaling pathway, bta05414:Dilated cardiomyopathy and bta04916:Melanogenesis) did not reach the significant p value ($p \leq 0.05$) (Table 4 and Figure 5).

The STRING database was used to conduct protein–protein interaction network (PPIN) analysis using all the genes previously used in the functional analysis, to identify the interactions between these genes. Figure 6 shows that there is several interacting between genes (containing 78 nodes connected via 139 edges); the proportional interaction strength between these genes was shown by the intensity of staining between the lines that linked one gene to another (Figure 6).

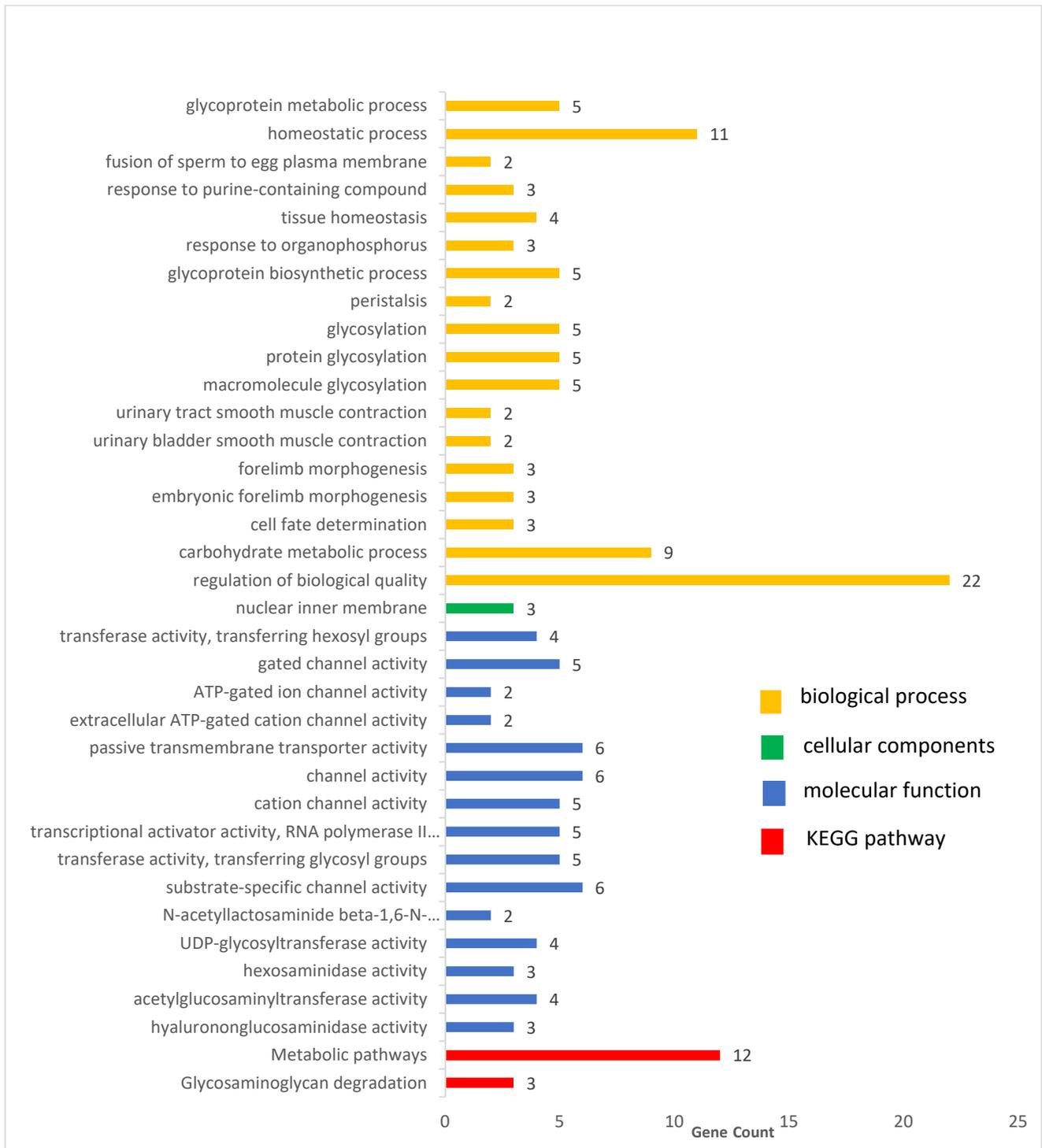


Figure 5. Significant GO terms (biological process, cellular component and molecular function) and KEGG pathways of candidate genes related to feet and leg traits.

Table 4. KEGG analysis for the regional candidate genes with genome-wide significant association.

Term	Description	Gene Count	%	p-Value	Genes
bta00531:	Glycosaminoglycan degradation	3	2.9	0.00469386	<i>LOC527125, LOC509761, SPAM1</i>
bta01100:	Metabolic pathways	12	11.4	0.016110131	<i>INPP4A, LOC527125, LIPC, LOC520336, MGAT4A, ALDH1A2, DNMT3A, LOC100848875, LOC509761, POLE, GALNT9, SPAM1</i>
bta04020:	Calcium signaling pathway	4	3.8	0.052382966	<i>P2RX3, P2RX2, ADCY3, CACNA1C</i>
bta05414:	Dilated cardiomyopathy	3	2.9	0.056875544	<i>ADCY3, CACNA2D4, CACNA1C</i>
bta04916:	Melanogenesis	3	2.9	0.072824808	<i>POMC, WNT5B, DCY3</i>

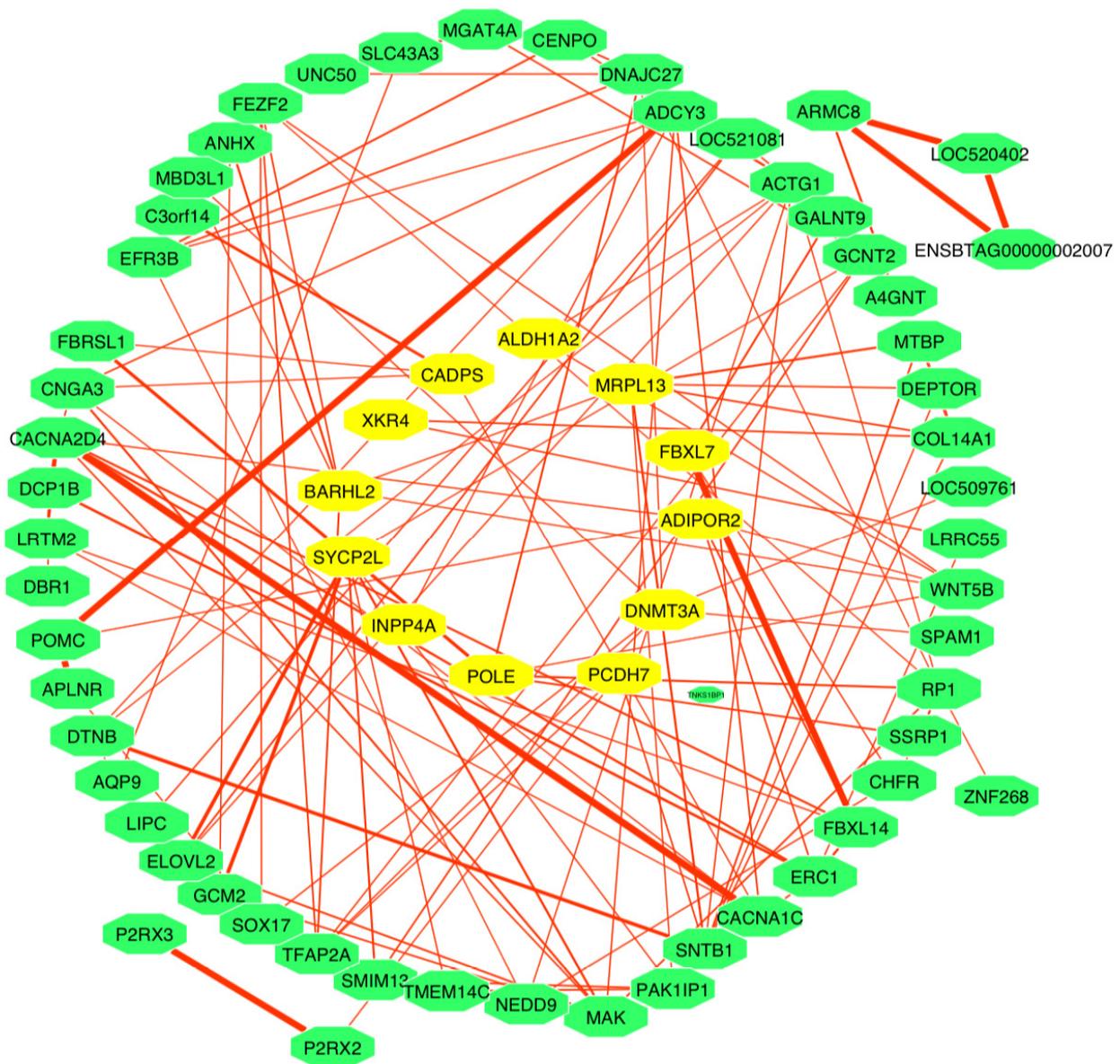


Figure 6. Gene interaction network for genes associated with feet and leg traits in Chinese Holstein cattle. (The nodes in yellow represent the significant candidate genes and their interacting partners).

4. Discussion

The estimated genetic correlations in our study ranged between -0.35 and 0.84 (Table 2); these results are in agreement with the study by Olasege et al. [37], who reported that between moderate negative and strong positive genetic correlation existed among feet and leg traits, ranging from -0.35 (between rear leg view and set of rear legs) to 0.74 (between foot angle and bone quality) in Chinese Holstein cattle. Among the feet and leg traits, the strongest genetic correlation was found between RLSV and BQ (0.84) (Table 2); if two traits have a high and positive genetic correlation, it implies that most QTL affect them both in the same direction [38]. For instance, most SNPs with significant influences affect weight and height in the same direction, thus helping to explain the known high genetic correlation between these two traits [39].

Population stratification or cryptic relatedness are considered major challenges that face genome-wide association analysis. Thus, the presence of population stratification can cause spurious association due to systematic ancestry differences [40] that can lead to false positives in GWAS results [41]; therefore, stratification must be well controlled across the experimental population to maximize power to detect true associations [42,43].

There are several methods to correct population stratification; carefully choosing the statistical model can be a useful method to correct and minimize the chances of type 1 error (false positive associations) [44].

To eliminate false positives, the most effective strategy is either (1) fitting population structure as covariates in a General Linear Model (GLM) [45] or (2) fitting both population structure and each individual's total genetic effect as covariates in a Mixed Linear Model (MLM) [46] to make adjustments for testing markers. As MLM can miss some potentially important discoveries and lead to false negatives, due to the confounding among population structures, quantitative trait nucleotides (QTNs) and kinship. Therefore, in the present study, we conducted Fixed and random model Circulating Probability Unification (FarmCPU) model analysis because of its advantages to completely control false positives, eliminate confounding and improve computational efficiency by using the fixed effect model and random effect model iteratively [29].

The inflation factor (λ) should be close to 1, after correcting the population stratification [45]. In the present study, the Q-Q plot in Figure 3 shows that the deviation of the observed value from the expected value is near to 1 and the inflation factor (λ) is $0.9 < 1.01$, both indicating that the population stratification was properly corrected by using an appropriate model.

The principal component analysis (PCA) method [47] and ADMIXTURE program analysis [28] are able to detect the population structure by classifying individuals into grouped ancestry based on their genetic makeup. Figure 2a shows all the experimental population was clustered into two groups (one large and one small). This indicates that each group clustered closely together has a genetic relationship. The division in two groups might be attributed to using Holstein semen from different countries or the cows in these farms may contain blood from other breeds, as it is known one of the requirements for registering Holstein cattle in China is that the cattle have at least 87.5 percent blood of Holstein (Chinese Holstein, GB/T 3157 2008).

In the present study, genome-wide association analysis identified 20 significant SNPs associated with four traits belonging to feet and legs trait in Chinese Holstein cattle using the FarmCPU model; among them, the most three significant SNPs were BovineHD2000006450-rs137022628, located near the *ACTBL2* gene, ARS-BFGL-NGS-91167-rs41565304, located within the *ADIPOR2* gene and no genes were found within 200 kb upstream or downstream of the BovineHD0300022142-rs43350216 SNP.

Until now, only one report has discussed the functional role of the *ACTBL2* gene. Hödebeck et al. [48] reported that the silencing of *ACTBL2* leads to diminished motility of human arterial smooth muscle cells. These authors also demonstrated that the expression of *ACTBL2* in smooth muscle cells under stretch conditions depends on the nuclear factor 5 of activated T-cells (NFAT5).

Adiponectin receptor 2 (*ADIPOR2*) is a member of the adipocytokines gene family. Generally, adiponectin receptor genes play an important role in bone and whole-body energy homeostasis that could be induced by activation of the AMPK signaling pathway; *ADIPOR1* and *ADIPOR2* also play a crucial role in metabolic pathways, which lead to the regulation of lipid and glucose metabolism, oxidative stress and inflammation [49]. Ouyang et al. [50] reported *ADIPOR2* has a role in intramuscular fat content in the longissimus dorsi (LD) muscle of different pig breeds. Lewis et al. [49] discussed in detail the role of adiponectin signaling in bone homeostasis and its effect on the lifestyle of human healthy aging and disease. Moreover, *ADIPOR2* is considered to be one of the genes associated with conformation and reproductive performance traits (hip height, rump length, calving ease, height, ovulation, type and rump angle) in one of the South African cattle breeds [51].

In addition to the *ADIPOR2* gene, our study identified other genes that were significantly associated with rear-leg side view traits (*INPP4A*, *DNMT3A*, *ALDH1A2* and *PCDH7*).

Some studies concerning the inositol polyphosphate-4-phosphatase type I A gene (*INPP4A*) reported it to have an association with the thoracic vertebrae number in sheep [52] and a role in regulating the normal functioning of the spinal cord neurons [53]. Additionally, the metabolic process [54] also presented different isoforms in transcriptomic investigation of meat tenderness in two Italian cattle breeds [55].

The DNA methyltransferase 3 alpha gene (*DNMT3A*) was associated to obese adipose tissue in transgenic mice [56] and significantly associated with beef quality traits [57]. Meanwhile, the aldehyde dehydrogenase 1 family member A2 gene (*ALDH1A2*) was associated with the growth trait in the Blanco Orejinegro cattle breed from Colombia [58], as well as intramuscular fatty acid composition in rabbits [59]. *ALDH1A2* was also associated with lactation persistency in Canadian Holstein cattle [60], carcass traits in Chinese Simmental beef cattle [61] and litter traits in Landrace and Large White pigs [62]. Moreover, *ALDH1A2* has been reported to play critical roles in the synthesis of retinoic acid, the active derivative of vitamin A, which is important for limb and organ development [63].

The protocadherin 7 gene (*PCDH7*) might be related to the body size trait of Chinese Simmental beef cattle [64], residual feed intake in Nelore cattle steers [65], feed efficiency trait in pig [66], internal organ traits in chickens [67] and sperm traits in Assaf rams [68].

The BovineHD4100014792-rs134139959 SNP located within F-box and leucine rich repeat protein 7 gene (*FBXL7*) was associated with the rear leg rear view trait in the present study, but, in previous studies, was associated with body length in water buffaloes (*Bubalus bubalis*) [69], clinical mastitis in first lactating US Holstein dairy cattle [70], subclinical ketosis in second and later lactations in Canadian Holstein dairy cattle [71] and also litter traits in Landrace and Large White pigs [62].

Furthermore, the results indicate the *TMEM229A*, *POLE*, *XKR4* and *CADPS* genes were harboring significant SNPs correlated to the bone quality trait. In relation to the BovineHD0400024774-rs133088614 SNP, located close to the transmembrane protein 229A gene (*TMEM229A*), Juan Carlos [72] described the polymorphism of this gene as having a great effect on dairy yield per lactation in Colombian Holstein cattle. There are no further publications concerning this gene in cattle.

Functional analyses of list candidate genes showed a DNA polymerase epsilon (*POLE*) gene potentially associated with reproductive traits [73]; they might be implicated in milk yield regulation by performing certain biological functions [74], while our findings suggest that *POLE* is associated with the bone quality trait.

The BovineHD1400007035-rs136017102 SNP on Chr14, located close to the *XKR4* gene, in previous studies, was reported to be associated with feed intake and growth phenotypes in cattle [75], residual feed intake in Australian Angus cattle [76], carcass weight (CWT) and eye muscle area (EMA) in Korean Hanwoo cattle [77], carcass trait [78], fat deposition in genome-wide association study for carcass trait in Brazilian Nelore cattle [79] and subcutaneous rump fat thickness in indicine and composite cattle [80].

The significant SNP BovineHD2200011035-rs110949452 on Chr22 is positioned within the calcium dependent secretion activator gene (*CADPS*); this gene is a member of the

Ca²⁺-dependent activator for the secretion protein family [81]. Interestingly, our study detected this gene is associated with the bone quality trait, while Vargas et al. [25] annotated the same gene within the top windows related to the overall quality of feet and leg in Nellore cattle. Kim et al. [82] reported this gene as selection signatures in Korean Brindle Hanwoo cattle obtained from genome-wide SNP analysis. Another study in two pig breeds identified the *CADPS* gene as associated with body weight [83].

In the present study, we found the ARS-BFGL-BAC-36389-rs109652453 SNP, which is associated with heel depth trait, to be located within synaptonemal complex protein 2, such as *SYCP2L*. Previously, most *SYCP2L* gene studies related to fertility. He et al. [84] reported a novel gene responsible for human premature ovarian insufficiency. For instance, *SYCP2L* was demonstrated to play a significant role in the survival of oocytes [85]. This gene is a paralogue of the synaptonemal complex protein gene (*SYCP2*), regulating females' reproductive aging. Therefore, it was suggested to have an association with the age of natural menopause in humans [85]. In addition, this gene is involved in the goat fecundity trait of Chinese Guangfeng goat [86].

Gene ontology enrichment analysis and KEGG pathways, in our study, revealed a number of GO terms and KEGG pathways. These involved significant genes related to the traits under examination, for example, the regulation of biological quality term (GO:0065008) in the biological process category, which has 22 genes (Figure 5 and Table S1); among them, the *CADPS*, *BARHL2*, *ADIPOR2* and *ALDH1A2* genes were considered as nearest genes to our significant SNPs. According to the Mouse Genome Database (MGD) <http://www.informatics.jax.org/> (accessed on 25 May 2021), this term and its sub-term (homeostatic process; GO:0042592) are related to any process that modulates a qualitative or quantitative trait of a biological quality, such as size, mass, shape, color, etc. As discussed earlier, some of the candidate genes involved in this term might play an important role in the biological regulation of body conformation traits and, subsequently, contribute to feet and legs traits.

In addition, the *LOC520336*, *MGAT4A*, *A4GNT*, *GCNT2* and *GALNT9* genes were clustered in certain biological processes and molecular functions. The terms (including carbohydrate metabolic process, GO:0005975, macromolecule glycosylation, GO:0043413, protein glycosylation, GO:0006486, glycosylation, GO:0070085, glycoprotein biosynthetic process, GO:0009101, glycoprotein metabolic process, GO:0009100, acetylglucosaminyl-transferase activity, GO:0008375, UDP-glycosyltransferase activity, GO:0008194, transferase activity, transferring glycosyl groups, GO:0016757, and transferase activity, transferring hexosyl groups, GO:0016758) in our study revealed to be related to metabolic processes.

Moreover, there were two terms (embryonic forelimb morphogenesis, GO:0035115, and forelimb morphogenesis, GO:0035136) that were related to the limb development term.

On another hand, the *P2RX3* and *P2RX2* genes were noticed in many biological process terms, such as urinary bladder smooth muscle contraction, GO:0014832, urinary tract smooth muscle contraction, GO:0014848, response to organophosphorus, GO:0046683, response to purine-containing compound, GO:0014074, peristalsis, GO:0022803, and in the molecular function category, such as channel activity (GO:0015267) and terms related with it (cation channel activity, gated channel activity, ATP-gated ion channel activity, extracellular ATP-gated cation channel activity and substrate-specific channel activity), and passive transmembrane transporter activity, GO:0022803. The *P2RX3* and *P2RX2* genes are subunits of the *P2X* gene family. *P2X* receptors are non-selective cation channels gated by extracellular ATP and exhibit relatively high Ca²⁺ permeability [87]; the *P2rx2* receptor mainly functions in sensory neurons, neuromuscular junction formation, whereas *P2RX3*, with multiple copies, were found in bony fishes [88].

The top two KEGG analysis pathways were bta00531: Glycosaminoglycan degradation, which has three genes, and bta01100: Metabolic pathways, which comprise 12 genes.

The functional analysis yielded many terms and pathways related to biological quality, metabolism and development. Therefore, it is reasonable to presume that all significant SNPs and candidate genes might be associated with feet and legs conformation.

5. Conclusions

In conclusion, this study identified 20 significant SNPs associated with feet and legs traits (bone quality, heel depth, rear leg side view and rear leg rear view) in Chinese Holstein cattle. Serval genes harbor SNPs (*ADIPOR2*, *INPP4A*, *DNMT3A*, *ALDH1A2*, *PCDH7*, *XKR4* and *CADPS*), identified mainly to participate in biological quality, metabolism and development processes. Our findings provide useful information to understand genetic architecture and give some fundamentals for molecular-based breeding, leading to genetic improvement programs on feet and leg traits. Further investigations to validate the biological functions of these genes are recommended.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ani11082259/s1>, Figure S1: Data Quality Control: (a). Histogram of individual filtering (b). Histogram of SNP filtering (c). Hardy–Weinberg equilibrium (HWE) (d). Histogram SNP divination from HWE, Table S1: Gene ontology terms for feet and leg traits title were significantly enriched ($p < 0.05$).

Author Contributions: Conceptualization, Z.Y., X.L. and I.M.A.; methodology, I.M.A.; software, I.M.A. and X.L.; validation, X.L. and Z.Y.; formal analysis, I.M.A. and X.L.; investigation, T.X., I.M.A., X.L. and M.N.; resources, X.L.; data curation, X.L., I.M.A. and M.N.; writing—original draft preparation, I.M.A.; writing—review and editing, A.A.I.A. and M.H.Y.; visualization, Z.Y.; supervision, Z.Y.; project administration, Z.Y., Y.M. and T.X.; 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 [2019]446), the national Natural Science Foundation of China (31872324) and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (XKYCX20_026).

Institutional Review Board Statement: The study was conducted according to the guidelines 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).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank all those who helped us during this project for their valuable advice and technical support.

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

References

1. Miglior, F.; Fleming, A.; Malchiodi, F.; Brito, L.F.; Martin, P.; Baes, C.F. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *J. Dairy Sci.* **2017**, *100*, 10251–10271. [[CrossRef](#)]
2. Pérez-Cabal, M.A.; García, C.; González-Recio, O.; Alenda, R. Genetic and phenotypic relationships among locomotion type traits, profit, production, longevity, and fertility in Spanish dairy cows. *J. Dairy Sci.* **2006**, *89*, 1776–1783. [[CrossRef](#)]
3. Wu, J.J.; Wathes, D.C.; Brickell, J.S.; Yang, L.G.; Cheng, Z.; Zhao, H.Q.; Xu, Y.J.; Zhang, S.J. Reproductive performance and survival of Chinese Holstein dairy cows in central China. *Anim. Prod. Sci.* **2012**, *52*, 11–19. [[CrossRef](#)]
4. Liu, S.; Yi, J.; Yang, L. Genetic parameters estimates for locomotion score, body condition score and final type score of Holstein cattle in southern China. *J. Appl. Anim. Res.* **2013**, *41*, 240–243. [[CrossRef](#)]
5. Heringstad, B.; Egger-Danner, C.; Charfeddine, N.; Pryce, J.E.; Stock, K.F.; Kofler, J.; Sogstad, A.M.; Holzhauser, M.; Fiedler, A.; Müller, K.; et al. Invited review: Genetics and claw health: Opportunities to enhance claw health by genetic selection. *J. Dairy Sci.* **2018**, *101*, 4801–4821. [[CrossRef](#)]
6. Chapinal, N.; Koeck, A.; Sewalem, A.; Kelton, D.F.; Mason, S.; Cramer, G.; Miglior, F. Genetic parameters for hoof lesions and their relationship with feet and leg traits in Canadian Holstein cows. *J. Dairy Sci.* **2013**, *96*, 2596–2604. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, H.; Wang, Z.; Wang, S.; Li, H. Retraction: Progress of genome wide association study in domestic animals. *J. Anim. Sci. Biotechnol.* **2013**, *4*, 1–10. [[CrossRef](#)] [[PubMed](#)]
8. Wang, D.; Ning, C.; Liu, J.F.; Zhang, Q.; Jiang, L. Short communication: Replication of genome-wide association studies for milk production traits in Chinese Holstein by an efficient rotated linear mixed model. *J. Dairy Sci.* **2019**, *102*, 2378–2383. [[CrossRef](#)] [[PubMed](#)]
9. 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](#)] [[PubMed](#)]

10. Wang, T.; Li, J.; Gao, X.; Song, W.; Chen, C.; Yao, D.; Ma, J.; Xu, L.; Ma, Y. Genome-wide association study of milk components in Chinese Holstein cows using single nucleotide polymorphism. *Livest. Sci.* **2020**, *233*. [[CrossRef](#)]
11. Liu, L.; Zhou, J.; Chen, C.J.; Zhang, J.; Wen, W.; Tian, J.; Zhang, Z.; Gu, Y. GWAS-based identification of new loci for milk yield, fat, and protein in holstein cattle. *Animals* **2020**, *10*, 2048. [[CrossRef](#)]
12. 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*, 103951. [[CrossRef](#)]
13. Li, X.; Buitenhuis, A.J.; Lund, M.S.; Li, C.; Sun, D.; Zhang, Q.; Poulsen, N.A.; Su, G. Inbred and Danish Holstein population genome-wide association study for milk fatty acid traits in Chateaus. *J. Dairy Sci.* **2015**, *98*, 8152–8163. [[CrossRef](#)] [[PubMed](#)]
14. 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*, 1–17. [[CrossRef](#)]
15. Yan, Z.; Wang, Z.; Zhang, Q.; Yue, S.; Yin, B.; Jiang, Y.; Shi, K. Identification of whole-genome significant single nucleotide polymorphisms in candidate genes associated with body conformation traits in Chinese Holstein cattle. *Anim. Genet.* **2020**, *51*, 141–146. [[CrossRef](#)]
16. 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*. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, X.; Chu, Q.; Guo, G.; Dong, G.; Li, X.; Zhang, Q.; Zhang, S.; Zhang, Z.; Wang, Y. Genome-wide association studies identified multiple genetic loci for body size at four growth stages in Chinese Holstein cattle. *PLoS ONE* **2017**, *12*, e0175971. [[CrossRef](#)]
18. 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](#)]
19. 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](#)] [[PubMed](#)]
20. Lin, S.; Wan, Z.; Zhang, J.; Xu, L.; Han, B.; Sun, D. Genome-wide association studies for the concentration of albumin in colostrum and serum in Chinese Holstein. *Animals* **2020**, *10*, 2211. [[CrossRef](#)]
21. Cai, Z.; Guldbrandsen, 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](#)]
22. 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*, 1–9. [[CrossRef](#)]
23. Cole, J.B.; Wiggans, G.R.; Ma, L.; Sonstegard, T.S.; Lawlor, T.J.; Crooker, B.A.; Van Tassell, C.P.; Yang, J.; Wang, S.; Matukumalli, L.K.; et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genom.* **2011**, *12*, 111. [[CrossRef](#)] [[PubMed](#)]
24. Abo-Ismael, M.K.; Brito, L.F.; Miller, S.P.; Sargolzaei, M.; Grossi, D.A.; Moore, S.S.; Plastow, G.; Stothard, P.; Nayeri, S.; Schenkel, F.S. Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle. *Genet. Sel. Evol.* **2017**, *49*, 82. [[CrossRef](#)]
25. Vargas, G.; Neves, H.H.R.; Camargo, G.M.F.; Cardoso, V.; Munari, D.P.; Carvalheiro, R. Genome-wide association study and functional analysis of feet and leg conformation traits in Nelore cattle. *J. Anim. Sci.* **2018**, *96*, 1617–1627. [[CrossRef](#)] [[PubMed](#)]
26. Madsen, P.; Milkevych, V.; Gao, H.; Christensen, O.F.; Jensen, J. DMU—a package for analyzing multivariate mixed models in quantitative genetics and genomics. In Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, Vancouver, BC, Canada, 17–22 August 2014; pp. 18–22.
27. 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](#)]
28. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)]
29. 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*, 1005957. [[CrossRef](#)]
30. 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](#)]
31. 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](#)]
32. Srikanth, K.; Lee, S.H.; Chung, K.Y.; Park, J.E.; Jang, G.W.; Park, M.R.; Kim, N.Y.; Kim, T.H.; Chai, H.H.; Park, W.C.; et al. A gene-set enrichment and protein–protein interaction network-based GWAS with regulatory SNPs identifies candidate genes and pathways associated with carcass traits in Hanwoo cattle. *Genes* **2020**, *11*, 316. [[CrossRef](#)]
33. 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*, 183. [[CrossRef](#)]

34. 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](#)]
35. 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](#)]
36. 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](#)]
37. Olasege, B.S.; Zhang, S.; Zhao, Q.; Liu, D.; Sun, H.; Wang, Q.; Ma, P.; Pan, Y. Genetic parameter estimates for body conformation traits using composite index, principal component, and factor analysis. *J. Dairy Sci.* **2019**, *102*, 5219–5229. [[CrossRef](#)] [[PubMed](#)]
38. Bolormaa, S.; Pryce, J.E.; Reverter, A.; Zhang, Y.; Barendse, W.; Kemper, K.; Tier, B.; Savin, K.; Hayes, B.J.; Goddard, M.E. A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for Stature, Fatness and Reproduction in Beef Cattle. *PLoS Genet.* **2014**, *10*, e1004198. [[CrossRef](#)]
39. Vargas, C.A.; Elzo, M.A.; Chase, C.C.; Olson, T.A. Genetic parameters and relationships between hip height and weight in Brahman cattle. *J. Anim. Sci.* **2000**, *78*, 3045–3052. [[CrossRef](#)]
40. 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](#)]
41. 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](#)]
42. 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](#)] [[PubMed](#)]
43. Bittles, A.H. Population stratification and genetic association studies in South Asia. *J. Mol. Genet. Med.* **2005**, *1*, 43–48. [[CrossRef](#)] [[PubMed](#)]
44. Sharma, A.; Lee, J.S.; Dang, C.G.; Sudrajad, 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](#)]
45. Zhou, X.; Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **2014**, *11*, 407–409. [[CrossRef](#)]
46. Devlin, B.; Roeder, K. Genomic control for association studies. *Biometrics* **1999**, *55*, 997–1004. [[CrossRef](#)]
47. 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*, 109. [[CrossRef](#)] [[PubMed](#)]
48. Hödebeck, M.; Scherer, C.; Wagner, A.; Hecker, M.; Korff, T. TonEBP/NFAT5 regulates ACTBL2 expression in biomechanically activated vascular smooth muscle cells. *Front. Physiol.* **2014**, *5*, 467. [[CrossRef](#)]
49. Lewis, J.W.; Edwards, J.R.; Naylor, A.J.; McGettrick, H.M. Adiponectin signalling in bone homeostasis, with age and in disease. *Bone Res.* **2021**, *9*, 1. [[CrossRef](#)] [[PubMed](#)]
50. Ouyang, H.; Yao, C.; Pang, D.; Lu, C.; Xu, A.; Huang, P.; Yu, H. Data mining and validation of ampk pathway as a novel candidate role affecting intramuscular fat content in pigs. *Animals* **2019**, *9*, 137. [[CrossRef](#)]
51. Makina, S.O.; Muchadeyi, F.C.; Van Marle-Köster, E.; Taylor, J.F.; Makgahlela, M.L.; Maiwashe, A. Genome-wide scan for selection signatures in six cattle breeds in South Africa. *Genet. Sel. Evol.* **2015**, *47*, 1–14. [[CrossRef](#)]
52. Shengwei, H.; Cunyuan, L.; Ming, L.; Xiaoyue, L.; Wei, N.; Yueren, X.; Rui, Y.; Bin, W.; Mengdan, Z.; Huixiang, L.; et al. Whole-genome resequencing reveals loci associated with thoracic vertebrae number in sheep. *Front. Genet.* **2019**, *10*, 674. [[CrossRef](#)]
53. Sasaki, J.; Kofuji, S.; Itoh, R.; Momiyama, T.; Takayama, K.; Murakami, H.; Chida, S.; Tsuya, Y.; Takasuga, S.; Eguchi, S.; et al. The PtdIns(3,4)P 2 phosphatase INPP4A is a suppressor of excitotoxic neuronal death. *Nature* **2010**, *465*, 497–501. [[CrossRef](#)] [[PubMed](#)]
54. Edea, Z.; Dadi, H.; Kim, S.W.; Park, J.H.; Shin, G.H.; Dessie, T.; Kim, K.S. Linkage disequilibrium and genomic scan to detect selective loci in cattle populations adapted to different ecological conditions in Ethiopia. *J. Anim. Breed. Genet.* **2014**, *131*, 358–366. [[CrossRef](#)] [[PubMed](#)]
55. Bongiorno, S.; Gruber, C.E.M.; Bueno, S.; Chillemi, G.; Ferrè, F.; Failla, S.; Moiola, B.; Valentini, A. Transcriptomic investigation of meat tenderness in two Italian cattle breeds. *Anim. Genet.* **2016**, *47*, 273–287. [[CrossRef](#)]
56. Kamei, Y.; Suganami, T.; Ehara, T.; Kanai, S.; Hayashi, K.; Yamamoto, Y.; Miura, S.; Ezaki, O.; Okano, M.; Ogawa, Y. Increased expression of DNA methyltransferase 3a in obese adipose tissue: Studies with transgenic mice. *Obesity* **2010**, *18*, 314–321. [[CrossRef](#)] [[PubMed](#)]
57. Liu, X.; Usman, T.; Wang, Y.; Wang, Z.; Xu, X.; Wu, M.; Zhang, Y.; Zhang, X.; Li, Q.; Liu, L.; et al. Polymorphisms in epigenetic and meat quality related genes in fourteen cattle breeds and association with beef quality and carcass traits. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 467–475. [[CrossRef](#)]
58. Londoño-Gil, M.; Rincón Flórez, J.C.; Lopez-Herrera, A.; Gonzalez-Herrera, L.G. Genome-Wide Association Study for Growth Traits in Blanco Orejinegro (Bon) Cattle from Colombia. *Livest. Sci.* **2021**, *243*, 4366. [[CrossRef](#)]

59. Laghouaouta, H.; Sosa-Madrid, B.S.; Zubiri-Gaitán, A.; Hernández, P.; Blasco, A. Novel genomic regions associated with intramuscular fatty acid composition in rabbits. *Animals* **2020**, *10*, 2090. [[CrossRef](#)]
60. Do, D.N.; Bissonnette, N.; Lacasse, P.; Miglior, F.; Sargolzaei, M.; Zhao, X.; Ibeagha-Awemu, E.M. Genome-wide association analysis and pathways enrichment for lactation persistency in Canadian Holstein cattle. *J. Dairy Sci.* **2017**, *100*, 1955–1970. [[CrossRef](#)]
61. Chang, T.; Xia, J.; Xu, L.; Wang, X.; Zhu, B.; Zhang, L.; Gao, X.; Chen, Y.; Li, J.; Gao, H. A genome-wide association study suggests several novel candidate genes for carcass traits in Chinese Simmental beef cattle. *Anim. Genet.* **2018**, *49*, 312–316. [[CrossRef](#)]
62. Wu, P.; Wang, K.; Yang, Q.; Zhou, J.; Chen, D.; Ma, J.; Tang, Q.; Jin, L.; Xiao, W.; Jiang, A.; et al. Identifying SNPs and candidate genes for three litter traits using single-step GWAS across six parities in landrace and large white pigs. *Physiol. Genom.* **2018**, *50*, 1026–1035. [[CrossRef](#)]
63. Sandell, L.L.; Sanderson, B.W.; Moiseyev, G.; Johnson, T.; Mushegian, A.; Young, K.; Rey, J.P.; Ma, J.X.; Staehling-Hampton, K.; Trainor, P.A. RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. *Genes Dev.* **2007**, *21*, 1113–1124. [[CrossRef](#)]
64. An, B.; Xu, L.; Xia, J.; Wang, X.; Miao, J.; Chang, T.; Song, M.; Ni, J.; Xu, L.; Zhang, L.; et al. Multiple association analysis of loci and candidate genes that regulate body size at three growth stages in Simmental beef cattle. *BMC Genet.* **2020**, *21*, 32. [[CrossRef](#)]
65. Tizioto, P.C.; Coutinho, L.L.; Decker, J.E.; Schnabel, R.D.; Rosa, K.O.; Oliveira, P.S.N.; Souza, M.M.; Mourão, G.B.; Tullio, R.R.; Chaves, A.S.; et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. *BMC Genom.* **2015**, *16*, 242. [[CrossRef](#)] [[PubMed](#)]
66. Sahana, G.; Kadlecová, V.; Hornshøj, H.; Nielsen, B.; Christensen, O.F. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. *J. Anim. Sci.* **2013**, *91*, 1041–1050. [[CrossRef](#)]
67. Moreira, G.C.M.; Salvian, M.; Boschiero, C.; Cesar, A.S.M.; Reecy, J.M.; Godoy, T.F.; Ledur, M.C.; Garrick, D.; Mourão, G.B.; Coutinho, L.L. Genome-wide association scan for QTL and their positional candidate genes associated with internal organ traits in chickens. *BMC Genom.* **2019**, *20*, 669. [[CrossRef](#)] [[PubMed](#)]
68. Serrano, M.; Ramón, M.; Calvo, J.H.; Jiménez, M.Á.; Freire, F.; Vázquez, J.M.; Arranz, J.J. Genome-wide association studies for sperm traits in Assaf sheep breed. *Animal* **2020**, *15*, 100065. [[CrossRef](#)] [[PubMed](#)]
69. Gonzalez Guzman, J.L.; Lázaro, S.F.; do Nascimento, A.V.; de Abreu Santos, D.J.; Cardoso, D.F.; Becker Scalez, D.C.; Galvão de Albuquerque, L.; Hurtado Lugo, N.A.; Tonhati, H. Genome-wide association study applied to type traits related to milk yield in water buffaloes (*Bubalus bubalis*). *J. Dairy Sci.* **2020**, *103*, 1642–1650. [[CrossRef](#)] [[PubMed](#)]
70. Tiezzi, F.; Parker-Gaddis, K.L.; Cole, J.B.; Clay, J.S.; Maltecca, C. A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure. *PLoS ONE* **2015**, *10*, e0114919. [[CrossRef](#)]
71. Nayeri, S.; Kroezen, V.; Sargolzaei, M.; Fleming, A.; Schenkel, F.S.; Baes, C.; Canovas, A.; Squires, J.; Miglior, F. Genome-wide Association Analysis for β -hydroxybutyrate Concentration in Milk Using Mid-Infrared Spectroscopy in North American Holstein Cattle. *BMC Genet.* **2019**, *20*, 58. [[CrossRef](#)] [[PubMed](#)]
72. Flórez, J.C.R.; Herrera, A.L.; Zuluaga, J.J.E. Genome-wide association study using the Bayes C method for important traits in dairy yield in Colombian Holstein cattle. *Acta Sci. Anim. Sci.* **2018**, *40*, 1–12. [[CrossRef](#)]
73. Saravanan, K.A.; Panigrahi, M.; Kumar, H.; Parida, S.; Bhushan, B.; Gaur, G.K.; Dutt, T.; Mishra, B.P.; Singh, R.K. Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics* **2021**, *113*, 955–963. [[CrossRef](#)] [[PubMed](#)]
74. Chen, Z.; Yao, Y.; Ma, P.; Wang, Q.; Pan, Y. Haplotype-based genome-wide association study identifies loci and candidate genes for milk yield in Holsteins. *PLoS ONE* **2018**, *13*, e0192695. [[CrossRef](#)] [[PubMed](#)]
75. Lindholm-Perry, A.K.; Kuehn, L.A.; Smith, T.P.L.; Ferrell, C.L.; Jenkins, T.G.; Freetly, H.C.; Snelling, W.M. A region on BTA14 that includes the positional candidate genes LYPLA1, XKR4 and TMEM68 is associated with feed intake and growth phenotypes in cattle. *Anim. Genet.* **2012**, *43*, 216–219. [[CrossRef](#)] [[PubMed](#)]
76. De Las Heras-Saldana, S.; Clark, S.A.; Duijvesteijn, N.; Gondro, C.; Van Der Werf, J.H.J.; Chen, Y. Combining information from genome-wide association and multi-tissue gene expression studies to elucidate factors underlying genetic variation for residual feed intake in Australian Angus cattle. *BMC Genom.* **2019**, *20*, 939. [[CrossRef](#)] [[PubMed](#)]
77. Bhuiyan, M.S.A.; Lim, D.; Park, M.; Lee, S.; Kim, Y.; Gondro, C.; Park, B.; Lee, S. Functional partitioning of genomic variance and genome-wide association study for carcass traits in Korean hanwoo cattle using imputed sequence level SNP data. *Front. Genet.* **2018**, *9*, 217. [[CrossRef](#)]
78. Srivastava, S.; Srikanth, K.; Won, S.; Son, J.H.; Park, J.E.; Park, W.; Chai, H.H.; Lim, D. Haplotype-based genome-wide association study and identification of candidate genes associated with carcass traits in Hanwoo cattle. *Genes* **2020**, *11*, 551. [[CrossRef](#)]
79. Medeiros De Oliveira Silva, R.; Stafuzza, N.B.; Fragomeni, B.D.O.; Ferreira De Camargo, G.M.; Ceacero, T.M.; Cyrillo, J.N.D.S.G.; Baldi, F.; Boligon, A.A.; Mercadante, M.E.Z.; Lourenco, D.L.; et al. Genome-wide association study for carcass traits in an experimental nelore cattle population. *PLoS ONE* **2017**, *12*, e169860. [[CrossRef](#)]
80. Porto Neto, L.R.; Bunch, R.J.; Harrison, B.E.; Barendse, W. Variation in the XKR4 gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. *Anim. Genet.* **2012**, *43*, 785–789. [[CrossRef](#)]
81. Cisternas, F.A.; Vincent, J.B.; Scherer, S.W.; Ray, P.N. Cloning and characterization of human CADPS and CADPS2, new members of the Ca²⁺-dependent activator for secretion protein family. *Genomics* **2003**, *81*, 279–291. [[CrossRef](#)]

82. Kim, S.; Cheong, H.S.; Shin, H.D.; Lee, S.S.; Roh, H.J.; Jeon, D.Y.; Cho, C.Y. Genetic diversity and divergence among Korean cattle breeds assessed using a BovineHD single-nucleotide polymorphism chip. *Asian-Australas. J. Anim. Sci.* **2018**, *31*, 1691–1699. [[CrossRef](#)]
83. Yang, Q.; Wu, P.; Wang, K.; Chen, D.; Zhou, J.; Ma, J.; Li, M.; Xiao, W.; Jiang, A.; Jiang, Y.; et al. SNPs associated with body weight and backfat thickness in two pig breeds identified by a genome-wide association study. *Genomics* **2019**, *111*, 1583–1589. [[CrossRef](#)] [[PubMed](#)]
84. He, W.B.; Tan, C.; Zhang, Y.X.; Meng, L.L.; Gong, F.; Lu, G.X.; Lin, G.; Du, J.; Tan, Y.Q. Homozygous variants in SYCP2L cause premature ovarian insufficiency. *J. Med. Genet.* **2020**, *58*, 168–172. [[CrossRef](#)] [[PubMed](#)]
85. Zhou, J.; Stein, P.; Leu, N.A.; Chmátal, L.; Xue, J.; Ma, J.; Huang, X.; Lampson, M.A.; Schultz, R.M.; Wang, P.J. Accelerated reproductive aging in females lacking a novel centromere protein SYCP2L. *Hum. Mol. Genet.* **2015**, *24*, 6505–6514. [[CrossRef](#)]
86. Islam, R.; Li, Y.; Liu, X.; Berihulay, H.; Abied, A.; Gebreselassie, G.; Ma, Q.; Ma, Y. Genome-wide runs of homozygosity, effective population size, and detection of positive selection signatures in six chinese goat breeds. *Genes* **2019**, *10*, 938. [[CrossRef](#)] [[PubMed](#)]
87. Kucenas, S.; Li, Z.; Cox, J.A.; Egan, T.M.; Voigt, M.M. Molecular characterization of the zebrafish P2X receptor subunit gene family. *Neuroscience* **2003**, *121*, 935–945. [[CrossRef](#)]
88. Zhang, Y.; Gao, H.; Li, H.; Guo, J.; Ouyang, B.; Wang, M.; Xu, Q.; Wang, J.; Lv, M.; Guo, X.; et al. *The White-Spotted Bamboo Shark Genome Reveals Chromosome Rearrangements and Fast-Evolving Immune Genes of Cartilaginous Fish*; Elsevier: Amsterdam, The Netherlands, 2020; Volume 23, ISBN 0000017191.