RESEARCH ARTICLE

Identification strategy of major genes for milk production traits in genetic selection of dairy cows

Xiayu Peng[1], Yong Wei[2, *], Aling Zou[3], Shasha Zhang[1], Tianyu Wei[1]

[1]College of Animal Science and Technology, Shihezi University, Shihezi, Xinjiang, China. [2]Xinjiang Tianrun Dairy Co., Ltd., Urumqi, Xinjiang, China. [3]Karamay Lucheng Agricultural Development Co., Ltd., Karamay, Xinjiang, China.

Received: January 30, 2024; **accepted:** April 2, 2024.

The dairy industry occupies an important position in the development of China's animal husbandry. It is a prominently developed industry in the animal husbandry industry and an important symbol of the country's agricultural development. The development of the dairy industry is of extremely important significance for promoting the transformation and upgrading of China's animal husbandry industry, increasing farmers' income, and enhancing the health quality of the entire nation. Therefore, the study aimed to explore the genetic effect of selective functional groups on milk production traits in Chinese Holstein cows to find the genes that caused mutations, and to study the identification strategies of the major genes that determined milk production traits. DNA was extracted from the frozen sperm genomes of 1,109 Holstein cows, and coagulated DNA from their blood was extracted. The mixed gene pool was sequenced, and different mutation sites were screened to classify genes. The collected data were statistically processed to analyze the correlation between genes. Haploid association analysis showed that all 35 Single Nucleotide Polymorphisms (SNP) and 3 indels were strongly associated with at least one trait, and 1, 1, 3, and 1 haplotype blocks were found in DDIT3, RPL23A, SESN2, and NR4A1 genes after continuous disequilibrium analysis. Correlation analysis showed that some haplotypes were correlated with lactation traits in dairy cows. The analysis of JASPAR software showed that 11 SNP sites and 3 binding sites had different degrees of changes, suggesting that this site would affect the transcriptional activity of transcription factor. The genetic effects of four candidate functional genes were further verified, and the four genes were identified as significantly associated with milk yield and milk composition traits in Chinese Holstein cow for the first time.

Keywords: genetic effects; milk production traits; Holstein cow; mutation sites; haplotype.

[*]**Corresponding author:** Yong Wei, Xinjiang Tianrun Dairy Co., Ltd., Urumqi 830000, Xinjiang, China. Emails: wy-260@163.com.

Introduction

Looking at the dairy cow breeding technology around the world, advanced countries such as Europe and the United States have been carrying out systematic breeding work since the 19th century and have established a complete set of mature breed breeding systems [1, 2]. In recent years, the rapid development of molecular marker-assisted selection, genome-wide association analysis, genome-wide resequencing, and transcriptomics has greatly promoted the development of dairy cow breeding, the innovation of dairy cow breeding technology, and the development of excellent germplasm resources [3, 4]. Whole-genome sequencing technology in dairy cows has injected new vitality into dairy cow breeding and

opened a new way for the genetic evaluation of dairy cows [5]. However, due to the late start of dairy cow breeding in China, there are generally problems such as low production efficiency, low genetic quality, and low breed coverage, and many problems in population genetic improvement and new breed breeding, which seriously affects the smooth development of dairy cow breeding [6]. In the field of dairy cow selection, some researchers domestically and internationally have carried out diversified research on it.

Lam *et al*. performed three DNA sequencing analyses using DNA sequence data from 9 Holstein cows and liver tissue from 10 Jersey cows to compare the overlap with functional Single Nucleotide Polymorphism (SNP) colocalization genes in each analysis of the two breeds in the low and high Residual Feed Intake (RFI) groups. The study revealed that numerous genes were pivotal in regulating biological processes with high metabolic demand, and were involved in cell growth and regeneration, metabolism, and immune function [7]. Mion *et al*. examined transcriptional expression of interferon-stimulated genes in peripheral blood leukocytes of a group of dairy cows including 67 Salts of Trace Minerals (STM) and 73 Organic Trace Minerals (OTM). After 15 days of artificial insemination, another group of cows including 28 STM and 29 OTM were subjected to uterine irrigation to recover fertilized eggs and uterine fluid for transcriptomic and metabolomic analysis, respectively. The results showed that there were 589 distinctly expressed transcripts in the two treatments, many of which indicated faster fetal elongation and more selenoprotein expression in the OTM group. In pregnant cows, the OTM group had higher levels of 24 metabolites in uterine fluid, including spermidine, sucrose, and cholesterol. The results suggested that the replacement of STM with OTM modestly improved cyclical ovarian recovery and had an important impact on embryonic development prior to implantation but did not alter conception risk and pregnancy rate [8]. Raza *et al*. selected 317 Holstein cows for genotyping to investigate the association between different genotypes and cow milk composition. The results showed that, among the eight possible haplotype genes, four were considered to be major genes, and the estimated frequency of involvement was higher than 90%, which could be used as auxiliary markers for Chinese Holstein cattle [9]. Yang *et al*. inspected 90 SNPs and found that these SNPs were significantly associated with 5 Milk Production Traits (MPTs). These gene loci could be accurately analyzed using matrix-assisted laser desorption. The results showed that 36 of the 90 selected genotypes could be used as genetic markers, which was an effective breeding strategy [10].

Genomic selection is an important way to improve the relatively backward breeding level of dairy cows in China, and some scholars have conducted research on this. To prove that autosomal recessive deletions in dairy cows caused by loss-of-function mutations in the bovine apolipoprotein gene would lead to cholesterol deficiency in dairy cows, Wang *et al*. conducted an experiment to analyze the gene distribution and milk cholesterol content of Canadian Holstein cow. The results showed that the lack of cholesterol had no significant impact on milk composition and milk production [11]. Liu *et al*. analyzed the genetic impact and impact mechanism between the PDIA3 gene and dairy cow milk production performance. The study selected 362 Chinese Holstein cow for experiments. Fluorescein detection results showed that alleles T and C had higher motor activity and might be potential factors affecting PDIA3 gene expression. It was also found that overexpression of the PDIA3 gene induced higher levels of gene fluorescence intensity in dairy cows. The results suggested that the study revealed the significant genetic impact of the PDIA3 gene on the milk composition traits of dairy cows. The PDIA3 gene could be used as a genetic marker for dairy cow breeding [12]. However, the above-mentioned research was relatively one-sided and only focused on a certain gene or a certain experiment. Therefore,

the existing genomic breeding data cannot be directly and completely applied to breeding, and independent research is required.

To solve the current problem of lack of genetic breeding data, this study proposed to use Polymerase Chain Reaction (PCR) technology to detect and identify the main genes that determined MPTs in dairy cow genetics and analyzed the transcription factors site of action. Four target genes that were involved in regulating mammary gland development, lactation performance, coping with stress responses, and maintaining the metabolic balance of mammary gland cells were selected for this study including DNA damage inducible transcript 3 (DDIT3) (GenBank ID. AC_000162.1), ribosomal protein L23a (RPL23A) (GenBank ID. AC_000176.1), Sestrin 2 (SESN2) (GenBank ID. AC_000159), and Nuclear receptor subfamily 4 group A member 1 (NR4A1) (GenBank ID. AC_000162) [13, 14]. The matrix-assisted LIBS/TOF/MS method was selected to sequence the mixed gene pool, screen different mutation sites, and detect the genotype of the confirmed population. The transcription factor loci were then determined to obtain the correlation analysis results. Through this study, it was expected to obtain relevant genes that were significantly related to milk production and milk composition traits of Chinese Holstein cow.

Materials and methods

Experimental animals

A total of 40 Chinese Holstein cows were selected from Sanyuan Greenhe, Beijing, China with the selection criteria of no biological relationship, a certain number of daughter production performance records, and the daughters distributed in different dairy farms as much as possible for genetic polymorphism testing. As a test group, the daughters of the above 40 cow families were used as the experimental verification group for candidate gene association analysis. Based on the cow pedigree and dairy herd improvement (DHI) participation, a total of 1,109 daughter individuals were screened from 2020-2023. A group of Holstein cows in Beijing with clear pedigrees and standardized DHI records were used as research objects to analyze the genetic traits between the parents.

Design and synthesis of primers

Genomic DNAs were extracted from the frozen cow sperm and anticoagulation cow blood using DNA extraction kit (Beijing Tiangen Biochemical Technology Co., Ltd., Beijing, China). PCR primers for DDIT3, RPL23A, SESN2, and NR4A1 genes were designed using primer3.0 (Whitehead Institute, Cambridge, Massachusetts, USA) and Oligo6.0 software (Molecular Biology Insights, Inc., Boulder, Colorado, USA) with 23, 21, 13, and 14 pairs of primers for SESN2, NR4AI, RPL23A, and DDIT3 genes, respectively (Table 1). The primer sequences covered all coding regions, some intronic subregions, and regulatory regions. The primers were synthesized by Beijing Liuhe Huada Gene Technology Co., Ltd. 9Beijing, China). The primers were dissolved in ddH_2O to 10 pmol/μL.

Construction of DNA pool and sequencing for bulk segregation analysis and genotyping

The mixed pool sequencing, also known as bulk segregation analysis (BSA), is a fast method for locating target trait genes. The DNA samples (each at 50 ng/μL) from 40 Chinese Holstein cows were randomly divided and mixed into two groups to obtain two DNA pools. PCR was performed to amplify DDIT3, RPL23A, SESN2, and NR4A1 genes. The PCR reaction mixture consisted of 1.0 μL of template DNA (50 ng/μL), 1.25 μL of each 10 pmol/μL forward and reverse primers, 12.5 μL of 2× Taq Msdter Mix, and 9.0 μL of ddH_2O to a total volume of 25 μL. The PCR was performed using Mastercycler (Eppendorf, Enfield, CT, USA) with the program as 94°C for 5 mins followed by 35 cycles of 94°C for 30 s, melting temperature for 30 s, 72°C for 40 s, and then 72°C for 7 mins before stored at 4°C. The purified PCR products were subjected to mixed pool sequencing using an ABI3730XL sequencer (Applied Biosystems, Waltham, Massachusetts,

Table 1. PCR amplification primer sequence information of candidate genes.

Primer Name	Primer sequence	Primer Name	Primer sequence
DDIT3-1F	CCATCCTCTC CACGTTCAGT	RPL23A-1F	CATCTTGCCT GCGTCCTT
DDIT3-1R	TCTTCACCTC CCTCCTCTCA	RPL23A-1R	TTCCGGGGAT CAAGTTACAG
DDIT3-2F	TGAGAGGAGG AGGTGAAGA	RPL23A-2F	TGTAACTTGA TCCCCGGAAA
DDIT3-2R	TTGGACCAGA TGGTCTCTCC	RPL23A-2R	GTCCTAGGCC TGCAGTTGAC
DDIT3-3F	GGAGAGACCA CTGGTCCAA	RPL23A-3F	AGTCTCTCCC CAGCATCTCC
DDIT3-3R	ATGCCAGAAT TCGTGCTCTT	RPL23A-3R	TTTGCATCCC CAAATTCAGT
DDIT3-4F	CCGAGGGCTC ATCAGATACT	RPL23A-4F	TTGGGGATGC AAAGATAAGC
DDIT3-4R	GGGGTCGACT TAGTGTTGGA	RPL23A-4R	TCAGTCGTGT CCGACTCTTG
DDIT3-5F	TCCGGGTCCA ACACTAAGTC	RPL23A-5F	GGAGAATCCC GTGAACAGAG
DDIT3-5R	GACGGGGTGT CTTTCTCGTA	RPL23A-5R	CTTCACTAGA GCCCCAACCA
DDIT3-6F	AGTTGGCCAG GATTTTGCTA	RPL23A-6F	AATCGTGCAC ATGATGGAAA
DDIT3-6R	GCACTCAAGA CCCAGCTTTC	RPL23A-6R	CAGGTCCCTA ACCTGGAACTC
DDIT3-7F	TAGACTGACC GGTGGAGCTT	RPL23A-7F	TGTTCAGCGA CCAAAGTCTG
DDIT3-7R	GGGAGGTGTG TGTGACCTCT	RPL23A-7R	GGGATCAGGA AAATGAGCAA
DDIT3-8F	TGCCCTTCCC TTCTCAACTA	RPL23A-8F	TCTGGGTGGA AGAGGGTATG
DDIT3-8R	GCTGCAGAGG TGGCTAAACT	RPL23A-8R	GGCTGAGTGTTCCAACAGGT
DDIT3-9F	ACCTCTGCAG CAATCTGGTT	RPL23A-9F	AGTCCAGCTGGCTAATTCCA
DDIT3-9R	TCCTAATGAA TGCCAGTGCT T	RPL23A-9R	CTCCTCCGCTCTGACTTCTG
DDIT3-10F	CCAGAATCAT CCACTGATGG	RPL23A-10F	GTCCTTGCCACAAATCCCTA
DDIT3-10R	TGTTCTCACA ACCCATTCCA	RPL23A-10R	GCCTGGTGCCTATACCTGAA
DDIT3-11F	TGGAATGGGT TGTGAGAACA	RPL23A-11F	AGCCCTCAGCACTGTAGCAT
DDIT3-11R	TATCCTGCCC ACAAACTTCC	RPL23A-11R	GATTCCGCTTTCCATCACTC
DDIT3-12F	CAGACCTACC GTGCCAATCT	RPL23A-12F	AAGCCGAGAA TCAGGTGAGA
DDIT3-12R	GGTTTTGTGC CTGAGATGGT	RPL23A-12R	CGAAACCGTG TGTGTGAAGT
DDIT3-13F	AGGGTGATAT GGGTCAGCAG	RPL23A-13F	GAACTTTCTCTCCGCCTTCA
DDIT3-13R	CACAGGGGTA GAGATTGGA	RPL23A-13R	CTGACACCTT GTGGTCGAGA
DDIT3-14F	TCCAATCTCC TACCCCTGTG	NR4A1-1F	GACATCCTGG AATGCGAAGT
DDIT3-14R	TTCTCTAAGA GCCGGGGTGT	NR4A1-1R	GGCAGGTCAA GTGGTCTGAT
SESN2-1F	CGCTGTGGAG GTGGTTTAGT	NR4A1-2F	GCCTTTGACT GTGTGGATCA
SESN2-1R	AGTTGCCCTC CTCTGTAGCA	NR4A1-2R	TGCAATCACC ACCTTCAGTG
SESN2-2F	TGCTACAGAG GAGGGCAACT	NR4A1-3F	TGAAGAATTT GGAGGGATGG
SESN2-2R	TCTCTGGCCA CCGAACTCTA	NR4A1-3R	GACACAGCCC CTCAACCTTA
SESN2-3F	GGCCAGAGAT TTGCAGAGAG	NR4A1-4F	GAGAGGGGTG TTCTGCTCTG
SESN2-3R	GTCGCTTCTA CCAGCACCTC	NR4A1-4R	CACCCCCACT ACCAAAGAAA
SESN2-4F	GCTCCTGCTG GGATTCTCTA	NR4A1-5F	CCCAAACACT GTTGCCTCAT
SESN2-4R	CTTCAGCTTT TCGGTTCCTG	NR4A1-5R	CAATCACCTC CCTCATCCAG
SESN2-5F	GCCAAGGAGA TTCCAATGAA	NR4A1-6F	TATTTTGGGCTCTCGCTGAC
SESN2-5R	CCAATCAGAT GCCTGTTTCC	NR4A1-6R	GGGCTGTCAA GCTTTCACTC
SESN2-6F	GAAGTCCCGT TTCATCATGC	NR4A1-7F	CGAGTCTGCG AGCTGCTATT
SESN2-6R	TATCCGCTTG TCTGGCTTCT	NR4A1-7R	CCTTCATGCT AACCCCAAAA
SESN2-7F	GAAGCCAGAC AAGCGGATAG	NR4A1-8F	CTAGGGGCTC TGTTGTCTGG
SESN2-7R	AGAAGGCTCC CCTCTCTTTG	NR4A1-8R	ACTGAGGTGG CTGTGTAGGG
SESN2-8F	TGTGCCTCAA CTGAGACCTG	NR4A1-9F	GATTCCTGGG TTCTGTGGTG
SESN2-8R	TCTCGGAGCT TGTTTCTTCC	NR4A1-9R	AGTAGTCAGA GCCGCTGGAG
SESN2-9F	CTCCGCTGAT CTCCTTTTTG	NR4A1-10F	GAGGACTTCC AGGTGTACGG
SESN2-9R	GGCTCTGCTC TAAGCCTCCT	NR4A1-10R	AGCCTGCCCT CTTCCTAAAG
SESN2-10F	GGGTCAGAAC CGAAGTTTCC	NR4A1-11F	AGAGCGCTTTTGTCTGCAAT
SESN2-10R	GTGAGTGGCA GTGTGTGAGC	NR4A1-11R	CCTACAGCGATCTCCACTCC
SESN2-11F	GGCAAAGCAA CTGAGCAAGT	NR4A1-12F	GGAGTGGAGA TCGCTGTAGG
SESN2-11R	CCCAGTGCAG CCAAAAATAG	NR4A1-12R	CTTGCCAACT CTCGCCTATT
SESN2-12F	GGGTAGCTGA CCCTCAGAGA	NR4A1-13F	GATCTCGGCT CCATTCTCAC
SESN2-12R	TAGCAGGCCA GGATTCAAAC	NR4A1-13R	GAGAAGGCCA AGATGCTGTC

SESN2-13F	GACCACGGGT CTGATTTACC	NR4A1-14F	GGCTTTGCTG AACTGTCTCC
SESN2-13R	CACAGCATTG AAGACCCAGA	NR4A1-14R	CTCTCTGCAC CTCGTCACC
SESN2-14F	GATTTGTTGG GGTGGAGAGA	NR4A1-15F	AGTGCGTCTA ATCCCACACC
SESN2-14R	GTTATTCCTG GGGGTGGAAT	NR4A1-15R	GTAAGGGTCA CCCCTGTGAA
SESN2-15F	TTCTGCCTTT TCTTGCCACT	NR4A1-16F	TTGGTGTGCA CGCTCATAAT
SESN2-15R	CCCATTTCAT CCTTCCCTTT	NR4A1-16R	TATGGGTAGA GGGTGCCAGA
SESN2-16F	GCACGAGTAC TTTGGCCTTC	NR4A1-17F	TCTGGCACCC TCTACCCATA
SESN2-16R	CAGGGGAGAT GAGACAGCAT	NR4A1-17R	AGGCTTGTGG GAGGGATACT
SESN2-17F	GAGTGTCTCC CCAGAGCTGA	NR4A1-18F	AGGTGCCTGT GAGGCTAGAG
SESN2-17R	GATCAGTCCT TGGCTGCTTC	NR4A1-18R	GACAGGAAAC CAAGGCTCAG
SESN2-18F	GCCTCTGGGG ATTTTACACA	NR4A1-19F	TCAGGAGATC TGCCTCGTTT
SESN2-18R	GCTAGGTCAG CTCCCTGATG	NR4A1-19R	GCTGGGAGAT AGGGGAAGAC
SESN2-19F	AGCTGACCTA GCAGCCCATA	NR4A1-20F	CACTCCAGGC TCCTCTGTCT
SESN2-19R	TGTGACGGCA GCATTAAGAG	NR4A1-20R	AGGCACTGCT GGAAACAACT
SESN2-20F	GTGACCTCGG TCAATCGTCT	NR4A1-21F	CTCTGCCCCT TTCTCCTTTT
SESN2-20R	TTTCTCTCCC ACCCTCTCAA	NR4A1-21R	GAGGCCAGTT CTTCCAGTTG
SESN2-21F	CCGCTTCCCT CTAACCATCT		
SESN2-21R	AGATGCAGAC TGAGGGCCTA		
SESN2-22F	CTAGGCCCTC AGTCTGCATC		
SESN2-22R	AGCCGTGCCT GTCTTTTCTA		
SESN2-23F	CTGTTCTGAT GCCAAGGTCA		
SESN2-23R	GGTGGGACAT AACCGTGAAC		

USA). The bioinformatical software Chromas (Technelysium Pty Ltd., Gold Coast, Queensland, Australia), DNAMAN (Lynnon Biosoft, San Ramon, CA, USA), and NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) were used to analyze the sequences and sequencing map. The location and mutation type of the SNPs were identified through sequence alignment. Based on the types and sequence characteristics of the base mutation loci to be typed, the genotype identification of the above two base mutation loci was carried out using Bruker Daltonics Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS) (Bruker, Billerica, Massachusetts, USA) following manufacturer's instructions.

Calculation of population gene frequency and genotype frequency

Gene frequency is the ratio of an allele to other genes within a population, and the frequency *P* of allele A was calculated below.

$$p = \frac{2N_{AA} + N_{AB}}{2(N_{AA} + N_{AB} + N_{BB})} \quad (1)$$

The frequency q of allele B was calculated as:

$$q = \frac{2N_{BB} + N_{AB}}{2(N_{AA} + N_{AB} + N_{BB})} \quad (2)$$

The frequency of genotype AA was calculated in equation (3).

$$D = \frac{N_{AA}}{N_{AA} + N_{AB} + N_{BB}} \quad (3)$$

The frequency of genotype AB was calculated in equation (4).

$$H = \frac{N_{AB}}{N_{AA} + N_{AB} + N_{BB}} \quad (4)$$

The frequency of genotype BB was calculated in equation (5).

$$R = \frac{N_{BB}}{N_{AA} + N_{AB} + N_{BB}} \quad (5)$$

where N_{AA}, N_{AB}, and N_{BB} represented the number of genotypes AA, BB, and CC, respectively.

Haplotype analysis

Haploview 4.2 (Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA) was employed to calculate the degree of linkage disequilibrium (LD) and haplotype segmentation between SNPs. D' and R^2 values represented the degree of LD, and the larger the value, the stronger the degree of LD. Haplotype segmentation was obtained using the software's haplotypes solid spine algorithm. Taking a single gene as the research object, subsequent correlation analysis was conducted between haplotypes located in different segments and MPTs.

Single marker/haplotype association analysis

This study used SAS 9.13 software (SAS Institute Inc, Cary, North Carolina, USA) and the MIXED program to perform correlation analysis on the five main milk production indicators and genotype (haplotype) combinations of dairy cows through a 305-day period including milk production, milk fat amount, milk fat rate, milk protein amount, and milk protein rate. Animal models were utilized to perform correlation analysis, and the specific model was shown in formula (6).

$$Y = \mu + \mathrm{hys} + \mathrm{b} \times M + G + \mathrm{a} + \mathrm{c} \quad (6)$$

where Y was the observed value for the shape of milk production. μ was the overall mean value. hys was the annual and seasonal field effects. b was the regression coefficients for the covariates. M was the calving age effect. G was the genotype/haplotype combination effect. a was the individual random additive genetic effect. c was the stochastic residual effect. It was necessary to preprocess the raw data of the annual and quarterly fixed effects of the model field before analysis.

Allelic effect analysis

For allele analysis, SAS 9.13 was applied to test the significance of SNPs, and the additive effect was calculated below.

$$\mathrm{a} = (X_{AA} - X_{BB})/2 \quad (7)$$

where a was the additive effect. X_{AA} was the least squares mean (LSM) of the AA genotype MPT. X_{BB} was the LSM of the BB genotype MPT. The dominant effect was calculated in equation (8).

$$\mathrm{d} = X_{AB} - (X_{AA} + X_{BB})/2 \quad (8)$$

where d was the dominant effect. X_{AB} was the LSM of the AB genotype's MPT. The allele substitution effect was calculated as:

$$\alpha = a + d(q - p) \quad (9)$$

where α was the allele substitution effect. *p* was allele A's frequency. *q* was allele B's frequency.

Statistical analysis

The numerical statistical analysis was carried out by SPSS 23.0 software (IBM, Armonk, New York, USA). The general data was described as mean ± standard deviation. The counting data was tested by χ^2. The econometric data conformed to normal distribution and homogeneity of variance were examined using t-test, otherwise using rank sum test. $P < 0.05$ indicated a significance difference, while $P < 0.01$ indicated a very significant difference.

Results

Analysis of the effect of genotyping technology

The study first used MALDI-TOF-MS genotyping technology to genotype a population of 1,109 Chinese Holstein cow born from 40 selected bull samples. The density distributions of the three genotypes T (18), TC (152), and C (212) were 0°, 45°, and 90°, and most of the other genotypes

Table 2. Analysis of the association between DDIT3 gene and milk production traits.

SNP	Lactation	Genotype (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
g.56283814C>T	1	CC (550)	10,3381±61.44	343.97±2.74	3.33±0.025^{Aa}	301.48±2.00	2.95±0.008
		CT (435)	10,406±62.63	343.46±2.8	3.29±0.026^{c}	304.01±2.05	2.95±0.008
		TT (86)	10,501±93.91	337.38±3.99	3.20±0.038^{Bb}	304.99±2.86	2.95±0.013
		P value	0.107	0.1645	0.0004**	0.1274	0.2073
	2	CC (377)	10,757±58.37	390.6±2.55^{Aa}	3.65±0.024^{A}	322.19±1.86	2.99±0.008^{Aa}
		CT (309)	10,734±60.26	385.39±2.64^{Ab}	3.58±0.025^{Bb}	321.74±1.93	2.97±0.008^{a}
		TT (57)	10,617±112.07	370.98±4.63^{B}	3.47±0.045^{Bc}	317.29±3.4	2.94±0.016^{Bb}
		P value	0.4537	<0.0001**	0.0002**	0.3225	0.0064**
g.56284880C<T	1	CC (553)	10,308±61.39	342.12±2.74	3.33±0.025^{Aa}	303.89±1.98	2.95±0.008
		CT (436)	10,350±62.66	341.01±2.79	3.29±0.026^{a}	305.45±2.04	2.95±0.008
		TT (82)	10,478±95.65	336.20±4.02	3.21±0.039^{Bb}	306.48±2.92	2.93±0.013
		P value	0.1441	0.2505	0.0016**	0.4052	0.4529
	2	CC (378)	10,859±59.31	389.59±2.51^{Aa}	3.64±0.024^{Aa}	318.75±1.84	2.99±0.008^{a}
		CT (309)	10,813±60.48	383.80±2.6^{Ab}	3.59±0.024^{b}	317.05±1.88	2.98±0.008^{a}
		TT (53)	10,708±114.44	370.21±4.73^{B}	3.51±0.046^{Bb}	313.42±3.47	2.94±0.016^{b}
		P value	0.3856	<0.0001**	0.008**	0.2453	0.0432*
c.*21A>G	1	AG (32)	10,306±139.99^{a}	343.00±5.82^{a}	3.39±0.056	303.60±4.17^{a}	2.95±0.020
		AG (287)	10,209±67.72^{Aa}	335.03±2.97^{Aa}	3.31±0.028	300.64±2.18^{Aa}	2.94±0.009
		GG (750)	10,383±58.72^{Ba}	341.38±2.65^{Ba}	3.32±0.024	304.74±1.93^{Ba}	2.95±0.008
		P value	0.0033**	0.0068**	0.3604	0.0258*	0.875
	2	AA (21)	10,615±171.81^{a}	379.35±7.00	3.63±0.069	314.98±5.09^{ab}	3.01±0.025
		AG (192)	10,693±69.63^{Aa}	384.38±2.95	3.65±0.028	316.59±2.15^{a}	2.99±0.01
		GG (528)	10,878±53.9^{Ba}	388.43±2.38	3.61±0.022	321.09±1.73^{b}	2.98±0.007
		P value	0.0084**	0.1568	0.3817	0.0419*	0.0838

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$). The different superscripted letters in the same column indicated significant differences with small case as significant difference and large case as very significant difference.

achieved the expected results except for a few deviations, ensuring the accuracy of the results (Figure 1).

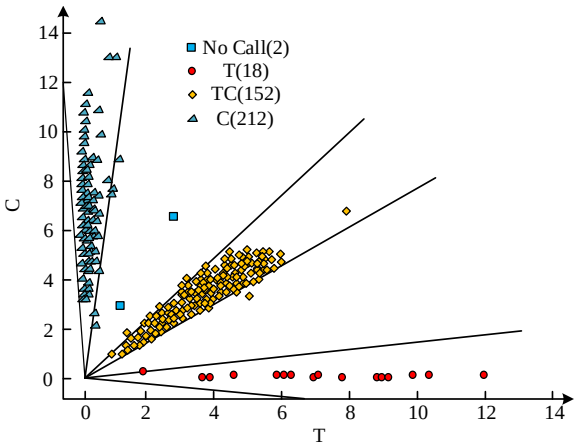

Figure 1. Scatter plot of Sequenom Mass Array classification.

Analysis of the genetic effect of DDIT3 gene on MPT

The correlation analysis between the three SNPs loci and milk fat yield, milk fat amount, milk protein amount, and milk protein rate were analyzed by multivariate statistical method. The results showed that there was a very significant or extremely significant correlation between the locus g.56285028A > G and milk yield quality and milk protein content in lactation 1 and 2. The locus g.56283814 and g.56284880 reached extremely significant levels with milk fat content (fat content) and milk protein content (fat content) (Table 2). The addition and substitution of g.66218917C > T and g.56284880C > T with the milk fat content in the second lactation period were very obvious, and the replacement of the T allele by each C allele could increase the milk fat yield by 11.80 kg and 11.41 kg ($P < 0.01$) (Table 3). The milk protein content in the second lactation period was added and replaced, indicating that the replacement of T allele by C allele could increase by 0.0287% and 0.0260%, respectively ($P < 0.05$). Compared with single-labeled SNPs, haplotype analysis had advantages

Table 3. Results of allelic additive, dominant, and substitution effects test for DDIT3 gene.

SNP	Lactation	Genotype (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
g.56283814C>T	1	Additive effect (a)	-81.78	3.29	0.0652**	-1.76	0.0111
		Dominant effect (d)	-13.62	2.79	0.0222	0.78	0.0062
		Substitution effect (a)	-87.69	4.50	0.0748*t	-1.42	0.0138
	2	Additive effect (a)	70.05	9.81**	0.0864**	2.45	0.0246**
		Dominant effect (d)	46.50	4.60	0.0223	2.00	0.0093
		Substitution effect (a)	90.21	11.80**	0.0961**	3.32	0.0287**
g.56284880C>T	1	Additive effect (a)	-84.98	2.96	0.061**	-13.00	0.008
		Dominant effect (d)	-42.58	1.84	0.0268	0.26	0.0071
		Substitution effect (a)	-103.73	3.77	0.0728**	-1.18	0.0111
	2	Additive effect (a)	75.47	9.69**	0.0635**	2.66	0.0206*
		Dominant effect (d)	29.52	3.90	0.0137	0.96	0.0122
		Substitution effect (a)	88.47	11.41**	0.0695*	3.09	0.0260*
c.*21A>G	1	Additive effect (a)	-38.25	0.81	0.0350	-0.57	0.0021
		Dominant effect (d)	-135.11	-7.16*	-0.0424	-3.52	-0.0054
		Substitution effect (a)	52.74	5.63	0.0635	1.81	0.0057
	2	Additive effect (a)	-131.48	-4.54	0.0064	-3.05	0.0166
		Dominant effect (d)	-53.39	0.49	0.0298	-1.45	0.0015
		Substitution effect (a)	-95.53	-4.87	-0.0136	-2.08	0.0155

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$).

Table 4. DDIT3 haplotype analysis results.

Lactation	Haplotype combination (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
1	H1H1(547)	10,281±61.34	345.18±2.75	$3.35±0.025^{Aa}$	$302.88±1.99$	2.95±0.008
	H1H2(423)	10,321±62.65	343.231±2.79	$3.31±0.026^{b}$	304.08±2.05	2.95±0.008
	H2H2(81)	10,448±96.27	338.74±4.04	$3.23±0.039^{Ba}$	306.36±2.91	2.94±0.013
	P value	0.1572	0.1641	0.001**	0.3521	0.3798
2	H1H1(395)	10,801±57.83	$395.36±2.55^{A}$	$3.66±0.024^{Aa}$	$321.91±1.85^{a}$	2.99±0.008
	H1H2(317)	10,742±60.78	$386.58±2.61^{B}$	$3.66±0.024^{Aa}$	$318.23±1.89^{b}$	2.98±0.008
	H2H2(55)	10,635±114.22	$373.68±4.67^{C}$	$3.54±0.046^{Bb}$	$314.02±3.43^{b}$	2.95±0.016
	P value	0.2926	< 0.001**	0.0075**	0.0226*	0.0912

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$). The different superscripted letters in the same column indicated significant differences with small case as significant difference and large case as very significant difference.

in the detection of complex traits. The study performed whole-genome sequencing on 1,109 progenies in the early stage and found that two SNP loci g.66218917G > A and g.56284880C > T constituted one haplotype, among which HI(CC)(CC) and H2(TT) were composed of two haplotypes with frequencies of 71.9% and 28.1%, respectively. The results of haplotype and association analysis were shown in Table 4. The results showed that, in the first and second lactation stages, the content of each haplotype and milk fat reached extremely significant ($P < 0.01$). In lactation 2, the milk protein content and milk fat content reached significant levels ($P < 0.05$) and extremely significant ($P < 0.01$), respectively, indicating that H1 haplotype was the main haplotype.

Analysis of SESN2 haplotype and correlation

The correlation between the haplotype of Block 1 and the milk fat content reached a very significant degree ($P < 0.01$), while haplotypes H1 and H3 had great advantages, and H2 was the weaker haplotype (Table 5). There was a strong correlation between the haplotype in Block 2 and the milk protein content during lactation (*P < 0.01*) with haplotype H4 being the dominant haplotype (Table 6). The haplotype in Block 3 was significantly or very significantly correlated with milk yield and milk protein content at 305 days. Haplotypes H1 and H3 were the main

Table 5. Results of haplotype analysis of SESN2 gene Block1.

Lactation	Haplotype combination (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
1	H1H1(682)	10,352±59.63	341.74±2.70^{Aa}	3.33±0.025^{Aa}	303.70±1.98	2.96±0.008^{Aa}
	H1H2(335)	10,324±68.16	335.17±3.00^{Bb}	3.27±0.028^{Bb}	300.82±2.19	2.93±0.009^{Bb}
	H1H3 (60)	10,398±114.02	339.59±4.72^{Aa}	3.30±0.046^{ab}	302.98±3.50	2.97±0.016^{a}
	P value	0.7465	0.0046**	0.0142*	0.1432	0.0005**
2	H1H1(481)	10,827±55.62^{Aa}	391.64±2.42^{Aa}	3.61±0.023	322.18±1.77^{Aa}	2.98±0.008
	H1H2(223)	10,647±68.96^{Bb}	378.96±2.92^{Bb}	3.58±0.028	316.19±2.12^{Bb}	2.98±0.009
	H1H3 (45)	10,879±131.70^{ab}	399.07±5.42^{Aa}	3.63±0.053	329.75±3.92^{Aa}	3.01±0.019
	P value	0.0143*	< 0.0001**	0.4578	0.0004**	0.1681

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$). The different superscripted letters in the same column indicated significant differences with small case as significant difference and large case as very significant difference.

Table 6. Results of haplotype analysis of SESN2 gene Block2.

Lactation	Haplotype combination (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
1	H1H1(394)	10349±64.72	341.4±2.87	3.32±0.027	307.49±2.12^{Aa}	2.96±0.009^{Aa}
	H1H2(270)	10410±71.24	342.58±3.13	3.32±0.029	307.05±2.26^{a}	2.95±0.010^{Aa}
	H1H3(177)	10316±77.67	335.09±3.4	3.29±0.032	301.65±2.47^{Bb}	2.92±0.011^{Bb}
	H1H4(82)	10343±98.18	338.81±4.15	3.31±0.040	307.62±3.03^{a}	2.98±00014^{Aa}
	H2H3(70)	10346±103.54	337.13±4.37	3.28±0.042	305.61±3.17^{ab}	2.95±0.015^{a}
	P value	0.751	0.0861	0.6324	0.0461*	<0.0001**
2	H1H1(273)	10901±63.40	392.01±2.76^{Aa}	3.65±0.026	320.23±2.00	2.98±0.009^{a}
	H1H2(193)	10772±71.48	390.35±3.00^{Aa}	3.65±0.029	317.68±2.21	2.97±0.01A^{ab}
	H1H3(116)	10775±84.40	385.73±3.55^{ab}	3.62±0.34	315.03±2.58	2.98±0.012^{ab}
	H1H4(62)	10692±109.18	376.28±4.57^{Bb}	3.63±0.044	316.94±3.32	3.02±0.016^{Bb}
	H2H3(47)	10801±122.08	383.44±5.02^{ab}	3.58±0.049	316.62±3.68	2.94±0.017^{Ab}
	P value	0.2295	0.0053**	0.6239	0.3423	0.0069**

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$). The different superscripted letters in the same column indicated significant differences with small case as significant difference and large case as very significant difference.

Table 7. Results of haplotype analysis of SESN2 gene Block3.

Lactation	Haplotype combination (No.)	Milk yield (kg)	Fat yield (kg)	Fat percentage (%)	Protein yield (kg)	Protein percentage (%)
1	H1H1(91)	10,582±96.00^{Aa}	346.24±4.08	3.30±0.039	313.32±2.94^{Aa}	2.96±0.013
	H1H2(247)	10,501±74.03^{Aa}	344.43±3.24	3.31±0.030	310.98±2.37^{ACa}	2.96±0.010
	H1H3(603)	10,327±60.36^{Bb}	340.05±2.73	3.31±0.025	305.59±1.97^{Bb}	2.95±0.008
	H2H3(125)	10,373±90.61^{ab}	338.74±3.85	3.30±0.037	304.65±2.80^{BCb}	2.94±0.013
	P value	0.0024**	0.1196	0.9435	0.0011**	0.4517
2	H1H1(61)	10,498±112.17^{Aa}	381.59±4.63	3.66±0.045	312.59±3.37^{Aa}	2.98±0.016
	H1H2(181)	10,723±75.10^{ab}	392.45±3.18	3.65±0.030	317.69±2.30^{ab}	2.97±0.011
	H1H3(418)	10,816±57.81^{Bb}	389.21±2.52	3.63±0.023	321.26±1.91^{Bb}	2.99±0.008
	H2H3(77)	10,912±111.15^{Bb}	393.09±4.65	3.62±0.045	324.00±3.39^{b}	2.98±0.016
	P value	0.0203*	0.1142	0.7685	0.0259*	0.3121

Notes: * indicated significant difference ($P < 0.05$). ** indicated very significant difference ($P < 0.01$). The different superscripted letters in the same column indicated significant differences with small case as significant difference and large case as very significant difference.

haplotypes, and their effects on the lactation process were different. Haplotype H1 had a significant advantage in the first lactation period, while H3 had a significant advantage in the second lactation period, which might be related to the difference in physiological conditions during lactation (Table 7). These excellent haplotypes could be used as molecular markers for the genetic improvement of Holstein cow in China.

Table 8. Results of haplotype analysis of SESN2 gene Block3.

Gene	SNPs	Mutation	TFBSs	Relative score
DDIT3	5:g.56283814 C>T	C	IRF7	0.87
		T	-	-
	5:g.56284880 C>T	C	-	-
		T	NFYA	0.86
RPL23A	19:g20702212 C>G	C	-	-
		G	EGR1, SP8	0.94, 0.88
	19:g20702782 ->G	-	GLIS2, GLIS3	0.92, 0.89
		G	SP8, KLF16	0.88, 0.91
SESN2	2:g.125716884 A>G	A	-	-
		G	PAX1	0.88
	2:g.125716735 G>T	G	ZBTB33	0.85
		T	-	-
	2:g.125716120 C>T	C	E2F4, E2F6	0.91, 0.93
		T	SMAD2::SMAD3::SMAD4	0.89
	2:g.125714860-125714860del	AGCGGGGTGGGGG	SP1, Klf4, KLF5	0.93, 0.93, 0.98
NR4A1	2:g.125714850 A>G	-	ZNF740	0.92
		A	E2F6	0.95
		G	KLF16	0.90

Analysis of genetic effects of NR4A1 gene and milk production traits

The study conducted a correlation analysis on the 7 SNPs sites of the NR4A1 gene and 5 MPTs including milk production, milk fat amount, milk fat rate, milk protein amount and milk protein rate in the first and second lactation periods of dairy cows in 305 days. Through multiple comparison analysis of the effects of different genotypes on MPTs, the results found that Loci g.27993737 A>G, g.27992897 C>T, g.27980964 C>T, and g.27975652 A>G all reached a significant or very significant correlation level with MPTs in the second lactation period ($P < 0.05$ or $P < 0.01$). In the first and second lactation periods, significant SNP sites were mainly concentrated on three traits including 305-day milk production, milk fat volume, and milk protein volume. Among them, locus g.27993737 A>G, g.27980964 C>T, and g.27975652 A>G reached a significant or very significant correlation level with 305-day milk production, milk fat amount, and milk protein amount in both lactation periods ($P < 0.05$ or $P < 0.01$). The site g.27992897C>T reached significant and very significant levels with milk fat rate in the first and second lactation periods ($P < 0.05$ or $P < 0.01$).

Analysis of genetic effects of RPL23A gene and milk production traits

The study compared 10 SNPs sites and 2 insertion/deletion sites in the RPL23A gene with milk production, milk fat volume, milk fat rate, milk protein amount, and milk protein rate in the 305 days of the first and second lactation periods of dairy cows. Correlation analysis was performed on milk traits, and the effects of different genotypes on MPTs were analyzed through multiple comparisons. Correlation analysis showed that twelve (12) mutation sites were significantly correlated with milk fat mass traits, and 8 SNPs sites reached very significant or significant correlation levels with milk fat mass in both lactation periods ($P < 0.01$ or $P < 0.05$). The other four SNPs sites reached a highly significant correlation level with the milk fat amount in the first lactation period ($P < 0.01$). In the first and second lactation periods, the loci g.20702122C>G, g.20704141 C>T, and g.20705000 C>T reached a significant correlation level with the milk fat rate ($P < 0.05$), and the sites g.20702088 A>G, g.20707028 C>T, and g.20707919C>G all reached a significant correlation with the milk protein rate ($P < 0.05$). All 12 mutation sites reached a significant correlation level with 305-day milk production and milk protein amount in at least one lactation period ($P < 0.05$).

Analysis of transcription factor binding sites

By using JASPAR (https://jaspar.elixir.no/), the study analyzed 2, 3, 12, and 3 SNP sites, and analyzed the binding sites of 3, 12, 3, 4, and 5 'upstream regulatory regions for 2, 3, 12, and 3 SNPs. 2, 2, 9, and 1 SNP sites were found, respectively (Table 8). The C→T mutation at the G.56283814C>T locus resulted in the loss of IRF7 transcription factor binding sites, while the conversion of the C > T locus from the C→T locus at g.56284880 CT locus led to an increase in the nuclear factor-Y (NFYA) transcription factor binding locus (relative score: 0.86). The C > G mutation of g20702122C > G led to the increase of transcription factor binding sites such as EGR1 and SP8, and the binding sites of GLIS2 and GLIS3 were deleted after the insertion of g20702782- > G locus, while the binding sites of SP8 and KLF1 were increased. In our previous studies, results showed that C > T mutation at C > T at NR4A1 led to an increase in SP4 transcription factor binding sites and a deletion in HIC2 transcription factor binding sites.

Discussion

The dairy industry is a major pillar of China's animal husbandry development and a sign of a country's agricultural development level [15]. Dairy industry plays a very important role in optimizing the agricultural industry structure, promoting the transformation and upgrading of animal husbandry industry, increasing farmers' income, increasing the production and consumption of dairy products in China [16, 17]. The results of this study showed that DDIT3, SESN2, RPL234, and NR4A1 genes were significantly or very significantly correlated with milk fat percentage or milk protein percentage, which indicated that there was a certain correlation between those genes and milk fat percentage and milk protein percentage of dairy cows. Previous studies have found that the SNPs of two milk fat metabolism key genes, DDIT3 and SESN2, were related to five important traits including milk fat content, milk fat percentage, milk protein content, and milk protein percentage, and milk protein content with significant or very significant correlations with milk protein content [16, 18, 19].

Genetic analysis of DDIT3 gene

As a transcription factor-binding protein, DDIT3 can regulate a variety of genes that are widely involved in physiological processes such as immune function, cell differentiation, and cell proliferation [20, 21]. Studies found that DDIT3 could inhibit the differentiation of adipocytes by regulating the expression of DATA2. Therefore, DDIT3 also certainly affected milk fat metabolism in dairy cows. Previous studies in our group found that the C > T (C > T) of DDIT3 gene g.56283814 and g.56284880 were related to milk fat content.

Results of SESN2 haplotype analysis and correlation analysis

SESN2 is an important receptor in the TORCI signaling pathway, which regulates cell growth by regulating protein and fat synthesis [22]. The analysis of the SESN2 locus revealed that the SESN2 locus had significant additive and substitution effects related to the milk protein rate of dairy cows, which was basically consistent with the results of genome-wide association analysis in previous studies [23]. Therefore, it was likely that these variant sites influenced the traits of milk protein rate to some extent. There were 13 SNPs in SESN2 gene and 3 haplotypes, among which the second haplotype was highly correlated with milk protein rate ($P < 0.01$), and haplotype H4 was the main haplotype. The results were inconsistent with the association analysis of single-molecule markers, as these loci were concentrated on the third gene, which might be biased by the use of different combined assays [24].

Prediction of transcription factor binding sites of candidate genes

DDIT3 is a class of CCAAT transcription factors that cause growth arrest in a variety of cells under stresses such as DNA damage [25]. The results showed that there was an NFYA binding site in the promoter region of this gene. Mutations in the regulatory region at the 5' end

of the gene led to the binding of transcription factors to target genes, which in turn affected the expression of target genes. NFYA is a highly conserved transcription factor that can bind enhancers on promoters of multiple genes. Previous study found that DDIT3 was a transcription factor of NFYA. It was found that the milk fat content and fat content of TT type dairy cows at the locus g.56284880 C > T were lower than those of TT type, so it was very likely that NFYA could reduce milk fat synthesis by regulating the expression of DDIT3. In addition, in previous studies, it was also found that the IRF7 transcription factor binding site was absent at the G.56283814 C > T locus, and the milk fat content and fat content of TT cows were lower than those of CC genotype, but the difference was statistically significant. Therefore, it was speculated that IRF7 was likely to promote milk fat synthesis by regulating the expression of DDIT3. The SESN2 gene was found to be 12 SNP sites located in the 5' regulatory region, of which 9 base mutations led to changes in transcription factor binding sites, and 16 transcription factors might affect their expression by regulating promoter region activity. The results showed that the mutation of SESN2 gene 125716884 A > G increased the binding sites of Paxl transcription factor, which decreased the milk yield and milk protein content ($P < 0.01$). The milk production, milk fat, and milk protein content of dairy cows increased significantly in the second lactation period ($P = 0.05$), which were mainly due to the regulation of SESN2 expression by Paxl transcription factors.

Specific proteins belong to zinc finger transcription factors, both of which contain 81 amino acids and contain 3 zinc finger regions. Specific proteins bind to GC-rich regions of multiple gene promoters. Transcription factors represented by SP4 were involved in the coupling of neuronal energy synthesis, neural activity, and energy metabolism. Through bioinformatics analysis, it was found that the 5' regulatory region of NR4AI gene might be the binding site of transcription factor, and the site change from C→T to the increase of SP4 transcription factor binding sites and binding sites deletion of HIC2. Previous studies have found that specific proteins could bind to GC-rich regions on proximal promoters of multiple genes, and SP4 could be directly transcriptionally activated.

Conclusion

The genetic effects of DDIT3 candidate genes were further verified, and DDIT3, SESN2, RPL234, and NR4A1 genes were significantly or extremely significantly correlated with MPT such as milk fat percentage or milk protein rate. The haplotype analysis of SESN2 and its mutation sites indicated that the mutation sites had the potential to change the alignment factor binding sites. The transcription factor binding sites of candidate genes were predicted, and it was speculated that IRF7 was likely to promote milk fat synthesis by regulating the expression of DDIT3. The regulation of Paxl transcription factor in SESN2 gene might affect the expression of SESN2 gene. The specific protein could bind to GC-rich regions on the proximal promoter of multiple genes.

Acknowledgements

The research was supported by Bingtuan Science and Technology Program (Grant No. 2021DB016), Innovation and Development Project of Shihezi University (Grant No. CXFZ202111), High-level Talent Project of Shihezi University (Grant No. RCZK201904), Autonomous Region Agricultural Backbone Personnel Training Project (Grant No. 2022SNGGGCC003), and Science and Technology Research Project of the Twelve Division (Grant No. SR202101).

References

1. Richardson CM, Crowley JJ, Amer PR. 2023. Defining breeding objectives for sustainability in cattle: challenges and opportunities. Anim Prod Sci. 63(10):931-946.

2. Song H, Lee K, Subburaj S, McGregor C, Lee GJ. 2023. ClPSY1 gene-based SNP markers identified from whole-genome resequencing for the determination of orange flesh color and carotenoid content in watermelon. Sci Hortic. 318(8):112120-112129.
3. Abbas HMK, Zhou YY, Huang HX. 2022. QTL mapping, whole genome resequencing, and marker-assisted selection provide basics of early flowering in pumpkin. Plant Breed. 141(2):266-276.
4. Guo L, Li Y, Lei Y, Gao J, Song C, Guo D, *et al*. 2023. Genome-scale investigation and identification of variations associated with early flowering based on whole genome resequencing and transcriptome integrated analysis in tree peony. Sci Hortic. 310(6):11695-11702.
5. Oh C. 2023. Exploring the way to harmonize sustainable development assessment methods in article 6.2 cooperative approaches of the Paris Agreement. Green Low-Carbon Econ. 1(3):121-129.
6. Rani G, Banu JR, Kumar G. 2022. Statistical optimization of operating parameters of microbial electrolysis cell treating dairy industry wastewater using quadratic model to enhance energy generation. Int J Hydrogen Energy. 47(88):27401-37414.
7. Lam S, Miglior F, Fonseca PAS, Gómez-Redondo I, Zeidan J, Suárez-Vega A, *et al*. 2020. Identification of functional candidate variants and genes for feed efficiency in Holstein and Jersey cattle breeds using RNA-sequencing. J Dairy Sci. 104(2):1928-1950.
8. Mion B, Madureira G, Spricigo JFW, King K, Van Winters B, LaMarre J, *et al*. 2023. Effects of source of supplementary trace minerals in pre- and postpartum diets on reproductive biology and performance in dairy cows. J Dairy Sci. 103(7):5074-5095.
9. Raza SHA, Khan R, Pant SD, Shah MA, Quan G, Feng L, *et al*. 2023. Genetic variation in the OPN gene affects milk composition in Chinese Holstein cows. Anim Biotechnol. 34(4):893-899.
10. Yang Z, Lian Z, Liu G, Deng M, Sun B, Guo Y, *et al*. 2021. Identification of genetic markers associated with milk production traits in Chinese Holstein cattle based on post genome-wide association studies. Anim Biotechnol. 32(1):67-76.
11. Wang M, Do DN, Peignier C, Dudemaine PL, Schenkel FS, Miglior F, *et al*. 2020. Cholesterol deficiency haplotype frequency and its impact on milk production and milk cholesterol content in Canadian Holstein cows. Can J Anim Sci. 100(4):786-791.
12. Liu S, Deng T, Hua L, Zhao X, Wu H, Sun P, *et al*. 2022. Novel functional mutation of the PDIA3 gene affects milk composition traits in Chinese Holstein cattle. J Dairy Sci. 105(6):5153-5166.
13. Wang T, Li J, Gao X. 2020. Genome-wide association study of milk components in Chinese Holstein cows using single nucleotide polymorphism. Livest Sci. 233(9):103951-103957.
14. Ben Braiek M, Fabre S, Hozé C, Astruc JM, Moreno-Romieux C. 2021. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. Genet Sel Evol. 53(1):1-13.
15. Chen B, Craiu RV, Strug LJ. 2021. The X factor: A robust and powerful approach to X-chromosome-inclusive whole-genome association studies. Genet Epidemiol. 45(7):694-709.
16. Zhang DY, Zhang XX, Li FD, Liu T, Hu Z, Gao N, *et al*. 2021. Whole-genome resequencing identified candidate genes associated with the number of ribs in Hu sheep. Genomics. 113(4):2077-2084.
17. Yang C, He X, Wang H. 2023. Single-molecule monitoring of membrane association of the necroptosis executioner MLKL with discernible anchoring and insertion dynamics. Nano Lett. 23(11):4770-4777.
18. Kuang Y, Ye N, Kyani A. 2022. Induction of genes implicated in stress response and autophagy by a novel quinolin-8-yl-nicotinamide QN523 in pancreatic cancer. J Med Chem. 65(8):6133-6156.
19. Xing ZY, Zhang ML, Wang YY. 2020. Short communication: A decrease in diameter of milk fat globules accompanies milk fat depression induced by conjugated linoleic acid supplementation in lactating dairy cows. J Dairy Sci. 103(6):5143-5147.
20. Neto JMDS, De Souza J, Lock AL. 2020. Predicting the yield of milk fat and milk fatty acid sources from fatty acid intakes in lactating dairy cows: A meta-analysis. J Dairy Sci. 103(1):82-83.
21. Naranjo A, Johnson A, Rossow H. 2020. Greenhouse gas, water, and land footprint per unit of production of the California dairy industry over 50 years. J Dairy Sci. 103(4):3760-3773.
22. Lin MW, Shen CC, Lin YJ. 2021. Enhancing the yield and activity of defucosylated antibody produced by CHO-K1 cells using Cas13d-mediated multiplex gene targeting. J Taiwan Inst Chem Eng. 121(1):38-47.
23. Hu Y, Stilp AM, Mchugh CP. 2021. Whole-genome sequencing association analysis of quantitative red blood cell phenotypes: The NHLBI TOPMed program. Am J Hum Genet. 108(6):874-893.
24. Padilha CG, Ribeiro CVDM, Oliveira DE. 2022. Modeling the effect of trans-10 fatty acids associated with milk fat depression in dairy goats and ewes supplemented with *trans*-10, *cis*-12 conjugated linoleic acid. Livest Sci. 258(2):1871-1878.
25. Luo Y, Yang B, Dong W. 2023. DNA damage-inducible transcript 3 deficiency promotes bone resorption in murine periodontitis models. J Periodont Res. 58(4):841-851.