



Article Polymorphisms of CYP7A1 and HADHB Genes and Their Effects on Milk Production Traits in Chinese Holstein Cows

Ao Chen¹, Qianyu Yang¹, Wen Ye¹, Lingna Xu¹, Yuzhan Wang¹, Dongxiao Sun^{1,2} and Bo Han^{1,*}

- ¹ Department of Animal Genetics and Breeding, College of Animal Science and Technology, Key Laboratory of Animal Genetics, National Engineering Laboratory for Animal Breeding, State Key Laboratory of Animal Biotech Breeding, China Agricultural University, Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, Beijing 100193, China; capudrooks@foxmail.com (A.C.); 2021304010308@cau.edu.cn (Q.Y.); s20233040757@cau.edu.cn (W.Y.); xulingna@caas.cn (L.X.); yuzhanwang@cau.edu.cn (Y.W.); sundx@cau.edu.cn (D.S.)
- ² Beijing Jingwa Agricultural Innovation Center, Beijing 100193, China
- Correspondence: bohan@cau.edu.cn

Simple Summary: This study aimed to identify genetic variants related to differences in milk production in Chinese Holstein cows. By analyzing specific genes involved in lipid metabolism, researchers identified variants linked to the milk, fat, and protein yield cows produce over time. They also found patterns in these genetic variations that could affect how the genes work together. The results of this study suggest that these genetic markers could be useful for selecting cows with better milk-production traits. Additionally, the findings could help scientists better understand how these genes influence milk production, potentially leading to improved strategies for breeding dairy cattle. Overall, this research provides valuable insights into enhancing milk production in dairy farming, which could ultimately benefit both farmers and consumers.

Abstract: Our preliminary research proposed the cytochrome P450 family 7 subfamily A member 1 (CYP7A1) and hydroxyacyl-coenzyme A dehydrogenase trifunctional multienzyme complex beta subunit (HADHB) genes as candidates for association with milk-production traits in dairy cattle because of their differential expression across different lactation stages in the liver tissues of Chinese Holstein cows and their potential roles in lipid metabolism. Hence, we identified single-nucleotide polymorphisms (SNPs) of the CYP7A1 and HADHB genes and validated their genetic effects on milk-production traits in a Chinese Holstein population with the goal of providing valuable genetic markers for genomic selection (GS) in dairy cattle, This study identified five SNPs, 14:g.24676921A>G, 14:g.24676224G>A, 14:g.24675708G>T, 14:g.24665961C>T, and 14:g.24664026A>G, in the CYP7A1 gene and three SNPs, 11:g.73256269T>C, 11:g.73256227A>C, and 11:g.73242290C>T, in HADHB. The single-SNP association analysis revealed significant associations (p value < 0.0461) between the eight SNPs of CYP7A1 and HADHB genes and 305-day milk, fat and protein yields. Additionally, using Haploview 4.2, we found that the five SNPs of CYP7A1 formed two haplotype blocks and that the two SNPs of HADHB formed one haplotype block; notably, all three haplotype blocks were also significantly associated with milk, fat and protein yields (*p* value ≤ 0.0315). Further prediction of transcription factor binding sites (TFBSs) based on Jaspar software (version 2023) showed that the 14:g.24676921A>G, 14:g.24675708G>T, 11:g.73256269T>C, and 11:g.73256227A>C SNPs could alter the 5' terminal TFBS of the CYP7A1 and HADHB genes. The 14:g.24665961C>T SNP caused changes in the structural stability of the mRNA for the CYP7A1 gene. These alterations have the potential to influence gene expression and, consequently, the phenotype associated with milk-production traits. In summary, we have confirmed the genetic effects of CYP7A1 and HADHB genes on milk-production traits in dairy cattle and identified potential functional mutations that we suggest could be used for GS of dairy cattle and in-depth mechanistic studies of animals.

Keywords: CYP7A1; HADHB; SNP; genetic effect; milk-production traits; dairy cattle



Citation: Chen, A.; Yang, Q.; Ye, W.; Xu, L.; Wang, Y.; Sun, D.; Han, B. Polymorphisms of *CYP7A1* and *HADHB* Genes and Their Effects on Milk Production Traits in Chinese Holstein Cows. *Animals* **2024**, *14*, 1276. https://doi.org/10.3390/ani14091276

Academic Editor: Bianca Castiglioni

Received: 8 April 2024 Revised: 21 April 2024 Accepted: 22 April 2024 Published: 24 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

Milk, often referred to as "white gold", plays a pivotal role as a rich source of essential nutrients supporting growth, bone-mass accrual, and immune function, containing proteins, fats, carbohydrates, minerals, and vitamins [1]. The surge in demand for high-quality dairy products is attributed to factors like global population growth, increasing income levels, and shifting dietary preferences [2]. This change in demand is particularly pronounced in developing regions, notably in Asia and Africa. Enhancing yield and quality traits represent paramount objectives for dairy breeding, necessitating the implementation of effective breeding strategies such as genomic selection (GS) to optimize these characteristics.

GS facilitates early and precise selection of young bulls and reserve heifers without relying on phenotypic data, thereby shortening the dairy breeding cycle from 5–6 years to about 2 years. By significantly reducing generation intervals, this approach lowers breeding costs and accelerates genetic progress. Therefore, improved accuracy in GS is crucial for the development of improved breeding strategies [3]. Single-nucleotide polymorphisms (SNPs), which are prevalent in genomic regions like promoters or enhancers, serve as primary genetic markers. They impact gene expression by altering how regulatory proteins bind or how chromatin is structured, thus affecting protein structure, RNA splicing [4], and miRNA binding [5]. They are crucial in influencing pathways associated with milk synthesis and lactation [6–9]. Integrating functional SNPs into genomic-selection chips can enhance the efficiency, accuracy, and sustainability of selection [10], leading to improved milk-production traits in cows.

The rapid development of -omics technology has facilitated the process of functionalgene/locus mining. Our preliminary research proposed that the cytochrome P450 family 7 subfamily A member 1 (CYP7A1) and hydroxyacyl-coenzyme A dehydrogenase trifunctional multienzyme complex beta subunit (HADHB) genes are differentially expressed in different lactation stages in dairy cows and are involved in lipid metabolism [11]. The CYP7A1 gene encodes cholesterol 7 alpha-hydroxylase, the rate-limiting enzyme in the classic pathway of bile-acid synthesis in the liver [12]. Bile acids are integral to processes such as lipid digestion, absorption, and excretion, as well as to the regulation of cholesterol homeostasis and metabolic pathways. The HADHB gene encodes hydroxyacyl-CoA dehydrogenase, a vital component of the mitochondrial trifunctional protein (MTP) complex, which is responsible for catalyzing the final three steps of mitochondrial beta-oxidation of long-chain fatty acids [13]. Furthermore, the CYP7A1 gene is positioned within the quantitative trait loci (QTL) QTL_ID: 2732 and QTL_ID: 3408, which have been demonstrated to affect milkfat yield and percentage and milk protein percentage [14–16]. HADHB is situated within the QTL region related to milk yield (QTL_ID: 1512) and is approximately 11.8 kb from the SNP rs110711742, which is associated with milk yield [17]. These findings suggest a potential association between CYP7A1 and HADHB and milk-production traits.

This study investigates the genetic associations between candidate genes *CYP7A1* and *HADHB* and milk-production traits in dairy cattle, including 305-day milk yield, fat yield, fat percentage, protein yield, and protein percentage. Additionally, predictive analyses were used to assess the regulatory effects of SNPs, particularly the effects on transcription factor binding sites (TFBSs) and mRNA stability, offering insights to support further exploration of causal mutations and their potential applications to developing GS chips for the breeding of dairy cattle.

2. Materials and Methods

2.1. Animal Selection and Phenotypic Data Collection

We selected 898 dairy cows from 45 Chinese Holstein sire families on 22 farms in Beijing Sunlon Livestock Development Co., Ltd. (Beijing, China) as the experimental population. The 45 bulls were used for SNP identification, and 898 cows were used for association analysis (898 cows in the first lactation and 611 in the second lactation). Each sire family had an average of 21 daughters, and each cow had three generations of pedigree information and Dairy Herd Improvement (DHI) records that were provided by the Beijing Dairy Cattle Centre (Beijing, China) (Table S1). The cows in each sire family were distributed across various dairy farms and were maintained with the same feeding conditions. Data consisted of the milk-production phenotype of each cow for the whole lactation period of the parity, which comprised 305-day milk yield, fat yield, fat percentage, protein yield, and protein percentage. The study was conducted in accordance with Guide for the Care and Use of Laboratory Animals and was approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University (Beijing, China).

2.2. Genomic DNA Extraction

The frozen semen of 45 bulls and blood samples from 898 cows were provided by Beijing Dairy Cattle Center (Beijing, China). We extracted genomic DNA from the semen using the salting-out procedure and used a TIANamp Blood DNA Kit (Tiangen, Beijing, China) to extract DNA from the blood samples. The quantity and quality of extracted DNA were determined by a NanoDrop2000 spectrophotometer (Thermo Science, Hudson, NH, USA) and gel electrophoresis (1.5%), respectively.

2.3. Identification and Genotyping of SNPs in Candidate Genes

According to the genomic sequences of the two genes, *CYP7A1* (NC_037341.1) and *HADHB* (NC_037338.1) of *Bos taurus* in Genbank, primers were designed using Primer3.0 (https://bioinfo.ut.ee/primer3-0.4.0/ (accessed on 8 September 2021)). They were then synthesized by Beijing Liuhe Bgi Co., Ltd. (Beijing, China). The primers covered the entire coding region and 2000 base pairs upstream and downstream of the regulatory regions of the genes. The DNA samples from semen were diluted to 50 ng/µL and used for PCR amplification (Table S2). PCR products were analyzed by 2% gel electrophoresis, and the qualified products were sequenced bidirectionally by Beijing Qingke Xinye Biotechnology Co., Ltd. (Beijing, China). Then, the sequences were aligned to the reference sequence (ARS-UCD1.2) using NCBI-blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi (accessed on 20 November 2023)) to identify SNPs. Subsequently, we genotyped the SNPs in the 898 cows using Genotyping by Target Sequencing (GBTS) technology by Boruidi Biotechnology Co., Ltd. (Shijiazhuang, China). The allelic and genotypic frequencies were calculated, and the Hardy–Weinberg equilibrium was tested by the chi-squared test.

2.4. Linkage Disequilibrium (LD) Estimation

The extent of LD between the identified SNPs was estimated using Haploview 4.2 (Broad Institute of MIT and Harvard, Cambridge, MA, USA), with the solid spine algorithm. The haplotype block with a frequency greater than 0.05 was retained. The extent of LD is measured by the D' value, to which it is proportional.

2.5. Association Analysis between SNP/Haplotype and Milk-Production Traits

The MIXED procedure in SAS 9.4 software (SAS Institute Inc., Cary, NC, USA) was used to perform association analysis between the SNP or haplotype block and the five milk-production traits (305-day milk yield, fat yield, fat percentage, protein yield and protein percentage) using the following animal model:

$$y = \mu + HYS + b \times M + G + a + e$$

where y is the phenotypic value of each trait of each cow; μ is the overall mean; HYS is the fixed effect of farm (1–22: 22 farms), calving year (1–4: 2012–2015) and calving season (1: April–May; 2: June–August; 3: September–November and 4: December–March); M is the age at calving as a covariant, b is the regression coefficient of covariant M; G is the genotype or haplotype combination effect; a is the individual random additive genetic effect with a distribution of $N(0, I\delta_e^2)$; A is a pedigree-based relationship matrix with an additive genetic variance of δ_a^2 ; and e is the random residual with a distribution of $N(0, I\delta_e^2)$, where I is the unit matrix and δ_e^2 is the residual variance. A Bonferroni correction was carried out by

multiple tests with a significance level equal to the original *p* value divided by the number of genotype or haplotype combinations.

Furthermore, the additive, dominant, and allelic substitution effects of the SNPs were calculated by the following formulas:

$$\mathbf{a} = (\mathbf{X}_{\mathrm{AA}} - \mathbf{X}_{\mathrm{BB}})/2;$$

$$d = X_{AB} - (X_{AA} + X_{BB})/2;$$
(1)

$$\alpha = a + d(q - p), \tag{2}$$

where a is the additive effect; d is the dominant effect; α is the allelic substitution effect; X_{AA} , X_{AB} and X_{BB} are the least-squares means of milk-production traits for the corresponding genotypes; p is the frequency of allele A; and q is the frequency of allele B.

2.6. Prediction of Transcription Factor Binding Sites

We used Jaspar (http://jaspar.genereg.net/ (version 2023)) software to predict whether SNPs in the 5' flanking region of *CYP7A1* and *HADHB* genes changed the transcription factor binding sites (TFBS) (relative score ≥ 0.80).

2.7. Prediction of mRNA Structure

We used the NCBI database (https://blast.ncbi.nlm.nih.gov/ (accessed on 20 November 2023)) to query for synonymy or missense mutations on exons. Subsequently, we used RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi (accessed on 20 November 2023)) to predict the changes to the RNA secondary structure caused by the mutation in the untranslated region, with the parameters of minimum free energy (MFE), partition function, and avoidance of isolated base pairs. MFE was used for a direct comparison of the folding stability of RNAs of the same sizes; the smaller the MFE, the greater the stability.

3. Results

3.1. SNPs Identification

We identified five SNPs in the *CYP7A1* gene and three in the *HADHB* gene in 898 dairy cows from 45 Chinese Holstein sire families on 22 farms. In *CYP7A1*, 14:g.24676921A>G (rs42765357), 14:g.24676224G>A (rs109454495) and 14:g.24675708G>T (rs42765359) were located in the 5' flanking region; 14:g.24665961C>T (rs108958186), was found in the 3' untranslated region (UTR); and 14:g.24664026A>G (rs109680813) was found in the 3' regulatory region. In *HADHB*, 11:g.73256269T>C (rs110033443) was located in the 5' flanking region, 11:g.73256227A>C (rs134856746) in the 5' UTR, and 11:g.73242290C>T (rs137211407) in the intron between exons 5 and 6. The genotypic and allelic frequencies of all the identified SNPs are summarized in Table 1.

3.2. Single-Marker Association Analysis

The results of the association analysis (Table 2) showed that 14:g.24665961C>T, in *CYP7A1*, was significantly associated with milk yield (*p* value = 0.0443) and fat yield (*p* value \leq 0.0001) in the first lactation period; four SNPs, 14:g.24676921A>G, 14:g.24676224G>A, 14:g.24675708G>T, and 14:g.24665961C>T, were significantly associated with milk, fat and protein yields (*p* value \leq 0.0059) and 14:g.24664026A>G was highly significantly associated with fat yield (*p* value = 0.0037) in the second lactation.

In *HADHB*, all three SNPs, 11:g.73256269T>C, 11:g.73256227A>C, and 11:g.73242290C>T, were significantly associated with milk and protein yields in the first lactation (*p* value ≤ 0.0461) and with milk, fat, and protein yields in the second lactation (*p* value ≤ 0.001). Additionally, the results of the additive, dominant, and allelic substitution effects were shown in Table 3.

Gene	SNP	rs ID	Gene Region	Genotype	Genotypic Frequency	Allele	Allelic Frequency
CYP7A1	14:g.24676921A>G	rs42765357	5' flanking region	AA AG GG	$0.1341 \\ 0.4340 \\ 0.4319$	A G	0.3511 0.6489
	14:g.24676224G>A	rs109454495	5' flanking region	AA AG GG	0.3073 0.4720 0.2207	A G	0.5433 0.4567
	14:g.24675708G>T	rs42765359	5' flanking region	GG GT TT	0.3083 0.4720 0.2196	G T	0.5444 0.4566
	14:g.24665961C>T	rs108958186	3′ UTR	CC CT TT	0.2218 0.4931 0.2851	C T	0.4683 0.5317
	14:g.24664026A>G	rs109680813	3' flanking region	AA AG GG	0.4298 0.4572 0.1130	A G	0.6584 0.3416
HADHB –	11:g.73256269T>C	rs110033443	5' flanking region	CC CT TT	0.6410 0.3242 0.0348	C T	0.8031 0.1969
	11:g.73256227A>C	rs134856746	5′ UTR	AA AC CC	0.0201 0.2809 0.6990	A C	0.1605 0.8395
	11:g.73242290C>T	rs137211407	Intron	CC CT TT	0.1373 0.2154 0.6473	C T	0.2450 0.7550

Table 1. Detailed information about the identified SNPs and their genotypic and allelic frequencies.

UTR: untranslated region.

3.3. Haplotype Association Analysis

Using Haploview 4.2, five SNPs in *CYP7A1*, 14:g.24675708G>T, 14:g.24676224G>A, 14:g.24676921A>G, 14:g.24665961C>T and 14:g.24664026A>G, formed two haplotype blocks (Figure 1). In Block 1, the frequencies of haplotypes H1 (AC), H2 (GT) and H3 (AT) were 0.463, 0.342, and 0.194, respectively. Block 2 was composed of three haplotypes, H1 (TGA), H2 (GAG), and H3 (TGG), with frequencies of 0.349, 0.545, and 0.105, respectively. We found that the two blocks in *CYP7A1* were significantly associated with milk, fat, and protein yields in both lactations (*p* value \leq 0.0011; Table 4).



Figure 1. Linkage disequilibrium estimated between SNPs in *CYP7A1* and *HADHB* gene. The blocks indicate haplotype blocks, and the text above the horizontal numbers is the SNP names. The values in the red boxes are pair-wise SNP correlations (D', while bright red boxes without numbers indicate complete linkage disequilibrium (LD) (D' = 1)).

Gene	SNP	Lactation	Genotype (No.)	Milk Yield (kg)	Fat Yield (kg)	Fat Percentage (%)	Protein Yield (kg)	Protein Percentage (%)
- - - - - - - -	14:g.24676921A>G	1	AA (118) AG (391) GG (389) <i>p</i> value	$\begin{array}{c} 10,\!289\pm 64.59\\ 10,\!308\pm 43.99\\ 10,\!299\pm 43.98\\ 0.9512 \end{array}$	$\begin{array}{c} 345.06 \pm 2.27 \\ 344.94 \pm 1.55 \\ 345.75 \pm 1.55 \\ 0.8719 \end{array}$	$\begin{array}{c} 3.37 \pm 0.03 \\ 3.36 \pm 0.03 \\ 3.37 \pm 0.03 \\ 0.8957 \end{array}$	$\begin{array}{c} 306.85 \pm 2.73 \\ 308.93 \pm 2.17 \\ 308.37 \pm 2.18 \\ 0.6648 \end{array}$	$\begin{array}{c} 2.98 \pm 0.01 \\ 2.30 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.3686 \end{array}$
		2	AA (79) AG (272) GG (260) <i>p</i> value	$\begin{array}{c} 11,067 \ ^{\rm A} \pm 103.8 \\ 10,660 \ ^{\rm B} \pm 69.92 \\ 10,717 \ ^{\rm B} \pm 71.91 \\ 0.0003 \end{array}$	$\begin{array}{c} 396.65 \ ^{\rm A} \pm 4.38 \\ 378.52 \ ^{\rm B} \pm 3.10 \\ 383.4 \ ^{\rm B} \pm 3.17 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.59 \pm 0.04 \\ 3.56 \pm 0.03 \\ 3.59 \pm 0.03 \\ 0.5719 \end{array}$	$\begin{array}{c} 330.55 \\ 317.36 \\ ^{\mathrm{Bb}} \pm 3.10 \\ 319.70 \\ ^{\mathrm{ABb}} \pm 3.17 \\ 0.0059 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 2.98 \pm 0.01 \\ 2.99 \pm 0.01 \\ 0.5828 \end{array}$
	14:g.24676224G>A	1	AA (278) GA (422) GG (198) <i>p</i> value	$\begin{array}{c} 10,\!287\pm50.05\\ 10,\!277\pm43.34\\ 10,\!370\pm53.54\\ 0.1972 \end{array}$	$\begin{array}{c} 346.51 \pm 1.73 \\ 343.81 \pm 1.49 \\ 346.97 \pm 1.85 \\ 0.1421 \end{array}$	$\begin{array}{c} 3.38 \pm 0.03 \\ 3.36 \pm 0.03 \\ 3.37 \pm 0.03 \\ 0.574 \end{array}$	$\begin{array}{c} 308.51 \pm 2.31 \\ 307.49 \pm 2.15 \\ 310.12 \pm 2.41 \\ 0.3774 \end{array}$	$\begin{array}{c} 3.01 \pm 0.01 \\ 2.99 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.2884 \end{array}$
		2	AA (181) GA (293) GG (137) <i>p</i> value	$\begin{array}{c} 10,\!755 {}^{\rm B} \pm 78.19 \\ 10,\!622 {}^{\rm B} \pm 69.32 \\ 10,\!942 {}^{\rm A} \pm 84.75 \\ 0.0006 \end{array}$	$\begin{array}{c} 385.15 \\ 378.07 \\ ^{B} \pm 3.05 \\ 389.83 \\ ^{A} \pm 3.63 \\ 0.0011 \end{array}$	$\begin{array}{c} 3.59 \pm 0.03 \\ 3.58 \pm 0.03 \\ 3.57 \pm 0.03 \\ 0.8325 \end{array}$	$\begin{array}{c} 321.69 \\ ^{AB} \pm 3.38 \\ 315.65 \\ ^{B} \pm 3.05 \\ 326.73 \\ ^{A} \pm 3.63 \\ 0.0029 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 2.97 \pm 0.01 \\ 2.98 \pm 0.01 \\ 0.0677 \end{array}$
	14:g.24675708G>T	1	GG (279) GT (422) TT (197) <i>p</i> value	$\begin{array}{c} 10,\!271\pm50.23\\ 10,\!282\pm43.57\\ 10,\!378\pm53.8\\ 0.1313 \end{array}$	$\begin{array}{c} 346.01 \pm 1.73 \\ 343.92 \pm 1.49 \\ 347.33 \pm 1.86 \\ 0.1635 \end{array}$	$\begin{array}{c} 3.38 \pm 0.03 \\ 3.36 \pm 0.03 \\ 3.37 \pm 0.03 \\ 0.5444 \end{array}$	$\begin{array}{c} 308.12 \pm 2.31 \\ 307.55 \pm 2.15 \\ 310.48 \pm 2.41 \\ 0.2926 \end{array}$	$\begin{array}{c} 3.01 \pm 0.01 \\ 2.99 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.244 \end{array}$
		2	GG (181) GT (293) TT (137) <i>p</i> value	$\begin{array}{c} 10,755 \\ 10,622 \\ ^{B} \pm 69.32 \\ 10,942 \\ ^{A} \pm 84.76 \\ 0.0006 \end{array}$	$\begin{array}{c} 385.15 \\ 378.07 \\ ^{B} \pm 3.05 \\ 389.83 \\ ^{A} \pm 3.63 \\ 0.0011 \end{array}$	$\begin{array}{c} 3.59 \pm 0.03 \\ 3.58 \pm 0.03 \\ 3.57 \pm 0.03 \\ 0.8325 \end{array}$	$\begin{array}{c} 321.69 \\ ^{AB} \pm 3.38 \\ 315.65 \\ ^{B} \pm 3.05 \\ 326.73 \\ ^{A} \pm 3.63 \\ 0.0029 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 2.97 \pm 0.01 \\ 2.98 \pm 0.01 \\ 0.0677 \end{array}$
	14:g.24665961C>T	1	CC (198) CT (438) TT (262) <i>p</i> value	$\begin{array}{c} 10,\!355\pm53.37\\ 10,\!250\pm42.29\\ 10,\!354\pm50.21\\ 0.0443 \end{array}$	$\begin{array}{c} 348.72 \ {}^{\rm A} \pm 1.87 \\ 341.82 \ {}^{\rm B} \pm 1.48 \\ 348.93 \ {}^{\rm A} \pm 1.76 \\ <\!0.0001 \end{array}$	$\begin{array}{c} 3.39 \pm 0.03 \\ 3.35 \pm 0.03 \\ 3.38 \pm 0.03 \\ 0.2811 \end{array}$	$\begin{array}{c} 310.35 \ ^{A} \pm 2.42 \\ 306.39 \ ^{B} \pm 2.14 \\ 310.43 \ ^{A} \pm 2.33 \\ 0.0258 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 2.99 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.369 \end{array}$
		2	CC (136) CT (305) TT (170) <i>p</i> value	$\begin{array}{c} 10,\!975 {}^{\rm Aa} \pm 85.38 \\ 10,\!604 {}^{\rm Bb} \pm 67.5 \\ 10,\!813 {}^{\rm ABa} \pm 80.72 \\ <\!0.0001 \end{array}$	$\begin{array}{c} 394.69 \ {}^{\rm Aa} \pm 3.66 \\ 377.09 \ {}^{\rm Bc} \pm 2.99 \\ 385.16 \ {}^{\rm ABb} \pm 3.49 \\ <\!\!0.0001 \end{array}$	$\begin{array}{c} 3.60 \pm 0.03 \\ 3.57 \pm 0.03 \\ 3.57 \pm 0.03 \\ 0.5643 \end{array}$	$\begin{array}{c} 328.62 \ {}^{\rm Aa} \pm 3.66 \\ 315.25 \ {}^{\rm Bb} \pm 2.99 \\ 322.85 \ {}^{\rm ABa} \pm 3.49 \\ 0.0001 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 2.98 \pm 0.01 \\ 2.99 \pm 0.01 \\ 0.1667 \end{array}$
	14:g.24664026A>G	1	AA (385) AG (410) GG (103) <i>p</i> value	$\begin{array}{c} 10,\!251\pm42.60\\ 10,\!358\pm42.67\\ 10,\!284\pm68.88\\ 0.0583 \end{array}$	$\begin{array}{c} 344.86 \pm 1.48 \\ 346.02 \pm 1.49 \\ 344.17 \pm 2.41 \\ 0.6425 \end{array}$	$\begin{array}{c} 3.38 \pm 0.03 \\ 3.36 \pm 0.03 \\ 3.36 \pm 0.04 \\ 0.5414 \end{array}$	$\begin{array}{c} 306.82 \pm 2.16 \\ 309.88 \pm 2.17 \\ 309.52 \pm 2.90 \\ 0.1429 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 2.99 \pm 0.01 \\ 3.01 \pm 0.01 \\ 0.5095 \end{array}$
		2	AA (261) AG (286) GG (64) <i>p</i> value	$\begin{array}{c} 10,\!798\pm70.29\\ 10,\!676\pm69.67\\ 10,\!758\pm113.28\\ 0.2012 \end{array}$	$\begin{array}{c} 386.32 \ ^{Aa} \pm 3.10 \\ 378.68 \ ^{Ab} \pm 3.09 \\ 389.85 \ ^{Aa} \pm 4.75 \\ 0.0037 \end{array}$	$\begin{array}{c} 3.59 \pm 0.03 \\ 3.57 \pm 0.03 \\ 3.62 \pm 0.05 \\ 0.4078 \end{array}$	$\begin{array}{c} 322.31 \pm 3.10 \\ 317.79 \pm 3.09 \\ 320.94 \pm 4.75 \\ 0.2597 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 2.98 \pm 0.01 \\ 2.99 \pm 0.02 \\ 0.7069 \end{array}$
HADHB	11:g.73256269T>C	1	CC (580) TC (287) TT (31) <i>p</i> value	$\begin{array}{c} 10,\!345\pm 40.5\\ 10,\!249\pm 49.36\\ 10,\!100\pm 118.79\\ 0.0236 \end{array}$	$\begin{array}{c} 345.52\pm1.40\\ 346.32\pm1.72\\ 338.41\pm4.16\\ 0.1926 \end{array}$	$\begin{array}{c} 3.36 \pm 0.03 \\ 3.39 \pm 0.03 \\ 3.35 \pm 0.06 \\ 0.2617 \end{array}$	$\begin{array}{c} 309.82 \pm 2.08 \\ 306.67 \pm 2.30 \\ 302.28 \pm 4.45 \\ 0.0461 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 3.00 \pm 0.01 \\ 2.99 \pm 0.02 \\ 0.8674 \end{array}$
		2	CC (370) TC (214) TT (27) <i>p</i> value	$\begin{array}{c} 10,\!873 \stackrel{A}{=} \pm 66.55 \\ 10,\!626 \stackrel{B}{=} \pm 74.67 \\ 10,\!091 \stackrel{C}{=} \pm 162.53 \\ <\!\!0.0001 \end{array}$	$\begin{array}{c} 387.93 \ ^{\rm Aa} \pm 2.95 \\ 380.21 \ ^{\rm Ab} \pm 3.26 \\ 351.42 \ ^{\rm Bc} \pm 6.68 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.58 \pm 0.03 \\ 3.60 \pm 0.03 \\ 3.50 \pm 0.07 \\ 0.3795 \end{array}$	$\begin{array}{c} 324.74 \ ^{A} \pm 2.95 \\ 315.93 \ ^{B} \pm 3.26 \\ 299.92 \ ^{B} \pm 6.68 \\ < 0.0001 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 2.98 \pm 0.01 \\ 2.98 \pm 0.02 \\ 0.585 \end{array}$
	11:g.73256227A>C	1	AA (18) AC (250) CC (630) <i>p</i> value	$\begin{array}{c} 10,\!532 {}^{AB} \pm 156.37 \\ 10,\!178 {}^{B} \pm 51.16 \\ 10,\!333 {}^{A} \pm 39.75 \\ 0.0015 \end{array}$	$\begin{array}{c} 349.98 \pm 5.49 \\ 344.16 \pm 1.81 \\ 345.45 \pm 1.42 \\ 0.499 \end{array}$	$\begin{array}{c} 3.31 \pm 0.08 \\ 3.40 \pm 0.03 \\ 3.36 \pm 0.03 \\ 0.2506 \end{array}$	$\begin{array}{c} 313.59 \\ 504.93 \\ a \\ \pm \\ 2.36 \\ 309.37 \\ b \\ \pm \\ 2.07 \\ 0.0184 \end{array}$	$\begin{array}{c} 3.00 \pm 0.03 \\ 3.00 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.9941 \end{array}$
		2	AA (19) AC (184) CC (408) <i>p</i> value	$\begin{array}{c} 9958.5 & ^{\rm Ac} \pm 193.86 \\ 10{,}540 & ^{\rm Bb} \pm 78.1 \\ 10{,}883 & ^{\rm Aa} \pm 65.13 \\ <\!0.0001 \end{array}$	$\begin{array}{c} 335.25 \ ^{\rm A} \pm 7.93 \\ 377.62 \ ^{\rm B} \pm 3.39 \\ 388.71 \ ^{\rm C} \pm 2.90 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.42 \pm 0.08 \\ 3.60 \pm 0.03 \\ 3.58 \pm 0.03 \\ 0.0665 \end{array}$	$\begin{array}{c} 294.18 \ ^{B} \pm 7.93 \\ 312.87 \ ^{B} \pm 3.39 \\ 325.20 \ ^{A} \pm 2.90 \\ < 0.0001 \end{array}$	$\begin{array}{c} 2.98 \pm 0.03 \\ 2.97 \pm 0.01 \\ 2.99 \pm 0.01 \\ 0.404 \end{array}$
	11:g.73242290C>T	1	CC (123) CT (189) TT (586) <i>p</i> value	$\begin{array}{c} 10,\!254 \overset{AB}{=} \pm 65.02 \\ 10,\!168 \overset{B}{=} \pm 54.74 \\ 10,\!358 \overset{A}{=} \pm 40.35 \\ 0.0012 \end{array}$	$\begin{array}{c} 344.89 \pm 2.30 \\ 344.54 \pm 1.95 \\ 345.66 \pm 1.45 \\ 0.8134 \end{array}$	$\begin{array}{c} 3.38 \pm 0.04 \\ 3.40 \pm 0.03 \\ 3.36 \pm 0.03 \\ 0.2182 \end{array}$	$\begin{array}{c} 306.77 \ ^{ab} \pm 2.73 \\ 304.97 \ ^{b} \pm 2.47 \\ 309.91 \ ^{a} \pm 2.08 \\ 0.0206 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 3.00 \pm 0.01 \\ 3.00 \pm 0.01 \\ 0.7748 \end{array}$
		2	CC (88) CT (147) TT (376) <i>p</i> value	$\begin{array}{c} 10,476 \ ^{B} \pm 99.09 \\ 10,545 \ ^{B} \pm 82.76 \\ 10,884 \ ^{A} \pm 66.08 \\ < 0.0001 \end{array}$	$\begin{array}{c} 368.81 \ ^{Bc} \pm 4.20 \\ 379.36 \ ^{ABb} \pm 3.58 \\ 387.96 \ ^{Aa} \pm 2.95 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.54 \pm 0.04 \\ 3.62 \pm 0.03 \\ 3.57 \pm 0.03 \\ 0.1681 \end{array}$	$\begin{array}{c} 308.42 \ ^{B} \pm 4.20 \\ 315.16 \ ^{B} \pm 3.58 \\ 325.03 \ ^{A} \pm 2.95 \\ < 0.0001 \end{array}$	$\begin{array}{c} 2.96 \pm 0.01 \\ 2.99 \pm 0.01 \\ 2.99 \pm 0.01 \\ 0.0674 \end{array}$

Table 2. Associations of eight SNPs in the *CYP7A1* and *HADHB* genes with milk yield and milk-composition traits in Chinese Holstein cattle during the first and second lactations.

The number in the table represents the mean \pm standard deviation; the number in the bracket represents the number of cows with the corresponding genotype; the *p* value shows the significance calculated for the genetic effects of SNPs; the superscript letters indicate the significance calculated for the effects different genotypes via the comparison of each pair of pairs; ^a, ^b, ^c within the same column with different superscripts means that the *p* value is < 0.05; and ^A, ^B, ^C within the same column with different superscripts means that the *p* value is < 0.05; and ^A, ^B, ^C within the same column with different superscripts means that the *p* value is < 0.01.

Gene	SNP Name	Lactation	Effect	Milk Yield (kg)	Fat Yield (kg)	Fat Percentage (%)	Protein Yield (kg)	Protein Percentage (%)
	14.~ 24676021 4 > C	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-5.0162 14.2908 -9.3289	$-0.3432 \\ -0.467 \\ -0.2023$	$-0.00126 \\ -0.00879 \\ 0.001392$	-0.7593 1.3145 -1.156	-0.00757 0.007425 -0.00981
	14.g.240/0921A>G	2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	174.61 ** -232.24 ** 243.41 **	6.6291 ** -11.5028 ** 10.0366 **	$\begin{array}{c} 0.00234 \\ -0.02803 \\ 0.01064 \end{array}$	5.4211 * -7.7643 ** 7.7212 **	$-0.00187 \\ -0.00847 \\ 0.000633$
	14:g.24676224G>A	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	$-41.8497 \\ -51.9072 \\ -46.4739$	$-0.2291 \\ -2.9335 * \\ -0.4904$	$\begin{array}{c} 0.006524 \\ -0.018 \\ 0.00492 \end{array}$	$-0.8039 \\ -1.8257 \\ -0.9666$	$0.004882 \\ -0.00847 \\ 0.004128$
		2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-93.7702 * -226.82 ** -110.1 *	-2.34 -9.4213 ** -3.0184	$0.009039 \\ -0.00727 \\ 0.008515$	-2.5205 -8.5594 ** -3.1369	$0.007237 \\ -0.01769 \\ 0.005963$
	14:g.24675708G>T	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-53.7569 -42.5283 -57.6403	$-0.6575 \\ -2.7457 \\ -0.9082$	$0.0065 \\ -0.01911 \\ 0.004755$	$-1.1824 \\ -1.7491 \\ -1.3421$	$0.004344 \\ -0.00997 \\ 0.003434$
CYP7A1 .		2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-93.7702 * -226.82 ** -110.1 *	-2.34 -9.4213 ** -3.0184	$0.009039 \\ -0.00727 \\ 0.008515$	-2.5205 -8.5594 ** -3.1369	0.007237 - 0.01769 0.005963
	14:g.24665961C>T	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	$0.911 \\ -104.99 * \\ 8.3937$	$-0.106 \\ -6.9981 \\ 0.3928$	$\begin{array}{c} 0.002497 \\ -0.03212 \\ 0.004787 \end{array}$	$-0.04286 \\ -3.9993 * \\ 0.2422$	$-0.00055 \\ -0.01024 \\ 0.000182$
		2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	80.7445 289.97 ** 96.8803 *	4.764 * -12.8359 ** 5.4783 **	$\begin{array}{c} 0.01721 \\ -0.01506 \\ 0.01805 \end{array}$	$2.8865 \\ -10.4784 ** \\ 3.4696$	$\begin{array}{c} 0.001474 \\ -0.0176 \\ 0.002453 \end{array}$
	14:g.24664026A>G	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	16.3957 89.7199 11.7791	0.3409 1.5028 0.8128	$\begin{array}{c} 0.009981 \\ -0.0135 \\ 0.005741 \end{array}$	-1.3534 1.7121 -0.8157	-0.00523 -0.00847 -0.00789
		2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	$20.2446 \\ -101.99 \\ -14.8297$	$-1.7674 \\ -9.407 ** \\ -5.0024$	$-0.01594 \\ -0.03741 \\ -0.0288$	$0.6858 \\ -3.8298 \\ -0.6313$	$-0.00119 \\ -0.00827 \\ -0.00403$
HADHB .	11:g.73256269T>C	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	122.03 * 26.4066 104.04 **	3.5596 4.3527 0.5932	$\begin{array}{c} 0.003985 \\ 0.03899 \\ -0.02259 \end{array}$	3.7702 0.6195 3.348 *	0.004206 0.001006 0.003521
		2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	390.89 ** 143.14 471.25 **	18.2568 ** 10.5408 ** 24.1742 **	0.03657 0.05419 0.067	12.4128 ** 3.5935 14.4302 **	$\begin{array}{c} 0.001967 \\ -0.00841 \\ -0.00275 \end{array}$
	11·σ 73256227Δ×C	1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	99.5398 254.17 ** 272.76 *	$2.2649 \\ -3.5579 \\ 4.6897$	-0.02704 0.05961 * -0.06766	$2.1125 \\ -6.5477 * \\ 6.5749$	-0.00149 0.001585 -0.00257
	11.g./ 020022/ A/C	2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-462.42 ** 118.71 -538 **	-26.7305 ** 15.6455 ** -36.6914 **	$-0.08246 * \\ 0.1017 * \\ -0.1472 *$	-15.5101 ** 3.184 -17.5372 **	-0.0063 ** -0.00734 -0.00162 **
		1	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-51.7106 -137.38 19.1215	-0.383 -0.7373 -0.00287	0.01289 0.02938 -0.00226	$-1.5714 \\ -3.3741 \\ 0.1682$	-0.00237 0.0067 -0.00583
	11:g./3242290C>1	2	Additive Effect (a) Dominance Effect (d) Allele Substitution Effect (α)	-203.9 ** -135.25 -140.15 *	-9.573 ** 0.9699 -10.0302 **	$\begin{array}{r} -0.01824 \\ 0.06063 \\ -0.04682 \end{array}$	-8.3062 ** -1.5693 -7.5665 **	$-0.01506 * 0.01771 \\ -0.02341 *$

Table 3. Additive, dominant and allele substitution effects of 8 SNPs in the *CYP7A1* and *HADHB* genes on milk yield and milk-composition traits in Chinese Holstein cattle during two lactations.

Note: The number in the table represents the mean; * indicates p < 0.05; ** indicates p < 0.01.

Similarly, two SNPs, 11:g.73256269T>C and 11:g.73256227A>C, in *HADHB* formed Block 3 (Figure 1). The frequencies of haplotypes H1 (AT), H2 (CC), and H3 (CT) were 0.159, 0.806, and 0.035, respectively. This block was found to be associated with milk, fat, and protein yields in the first lactation (p value \leq 0.0315), and milk, fat, and protein yields in the second lactation (p value \leq 0.0052; Table 4).

Block	Lactation	Haplotype Combination	Milk Yield (kg)	Fat Yield (kg)	Fat Percentage (%)	Protein Yield (kg)	Protein Percentage (%)
BLOCK1	1	H1H1 (197) H1H2 (282) H1H3 (155) H2H2 (102)	$\begin{array}{c} 10,\!410 \ ^{\rm A} \pm 53.81 \\ 10,\!329 \ ^{\rm A} \pm 46.52 \\ 10,\!125 \ ^{\rm B} \pm 57.38 \\ 10,\!314 \ ^{\rm AB} \pm 68.66 \\ 10,\!141 \ ^{\rm AB} \pm 68.66 \end{array}$	$\begin{array}{c} 352.81 \ ^{ABa} \pm 1.9 \\ 346.8 \ ^{Bb} \pm 1.64 \\ 343.65 \ ^{Bb} \pm 2.02 \\ 351.65 \ ^{ABab} \pm 2.42 \\ \end{array}$	3.41 ± 0.03 3.38 ± 0.03 3.41 ± 0.03 3.41 ± 0.04	$\begin{array}{c} 312.82 \ ^{A} \pm 2.41 \\ 310.03 \ ^{AB} \pm 2.23 \\ 304.38 \ ^{B} \pm 2.52 \\ 310.75 \ ^{AB} \pm 2.85 \end{array}$	3.01 ± 0.01 3.01 ± 0.01 3.01 ± 0.01 3.02 ± 0.01
		H2H3 (127) p value	$ \begin{array}{r} 10,466 \\ < 0.0001 \end{array} $	356.37 ^{Aa} ± 2.23 <0.0001	$3.42 \pm 0.04 \\ 0.6801$	$313.93 \times \pm 2.68$ 0.0011	3.00 ± 0.01 0.7621
	2	H1H1 (136) H1H2 (201) H1H3 (104) H2H2 (64) H2H3 (85) <i>p</i> value	$\begin{array}{c} 10,937 \ ^{Aa}\pm 86.81 \\ 10,539 \ ^{Bb}\pm 76.03 \\ 10,453 \ ^{Bb}\pm 91.26 \\ 10,560 \ ^{ABb}\pm 108.83 \\ 10,971 \ ^{Aa}\pm 98.06 \\ <\!\! 0.0001 \end{array}$	$\begin{array}{c} 392.63 \ ^{Aa} \pm 3.73 \\ 374.21 \ ^{Bb} \pm 3.31 \\ 372.18 \ ^{Bb} \pm 3.89 \\ 379.53 \ ^{ABb} \pm 4.56 \\ 385.25 \ ^{ABab} \pm 4.15 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.60 \pm 0.03 \\ 3.57 \pm 0.03 \\ 3.57 \pm 0.04 \\ 3.60 \pm 0.04 \\ 3.52 \pm 0.04 \\ 0.3899 \end{array}$	$\begin{array}{c} 327.99 \ ^{A} \pm 2.71 \\ 316.09 \ ^{B} \pm 2.40 \\ 311.78 \ ^{B} \pm 2.83 \\ 319.76 \ ^{AB} \pm 3.32 \\ 329.89 \ ^{A} \pm 3.02 \\ < 0.0001 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 3.00 \pm 0.01 \\ 2.99 \pm 0.01 \\ 3.04 \pm 0.02 \\ 3.01 \pm 0.014 \\ 0.1286 \end{array}$
BLOCK2	1	H1H1 (118) H1H3 (71) H2H2 (278) <i>p</i> value	$\begin{array}{c} 10,\!243 \ ^{\rm B} \pm 66.47 \\ 10,\!679 \ ^{\rm A} \pm 80.86 \\ 10,\!336 \ ^{\rm B} \pm 52.06 \\ <\!0.0001 \end{array}$	$\begin{array}{c} 349.59 \ ^{\rm b} \pm 2.33 \\ 359.13 \ ^{\rm a} \pm 2.84 \\ 353.69 \ ^{\rm ab} \pm 1.83 \\ 0.017 \end{array}$	$\begin{array}{c} 3.43 \pm 0.04 \\ 3.39 \pm 0.04 \\ 3.45 \pm 0.03 \\ 0.329 \end{array}$	$\begin{array}{c} 309.81 \ ^{\rm B} \pm 2.90 \\ 323.14 \ ^{\rm A} \pm 3.30 \\ 312.8 \ ^{\rm B} \pm 2.45 \\ 0.0004 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 3.02 \pm 0.02 \\ 3.02 \pm 0.01 \\ 0.3323 \end{array}$
	2	H1H1 (79) H1H3 (50) H2H2 (181) <i>p</i> value	$\begin{array}{c} 10,988 \ ^{\rm Aa} \pm 108.25 \\ 10,544 \ ^{\rm Bb} \pm 131.80 \\ 10,597 \ ^{\rm Bb} \pm 83.65 \\ 0.0007 \end{array}$	$\begin{array}{c} 391.81 \ ^{Aa} \pm 4.58 \\ 369.04 \ ^{Bb} \pm 5.51 \\ 379.16 \ ^{ABb} \pm 3.64 \\ 0.0006 \end{array}$	$\begin{array}{c} 3.58 \pm 0.04 \\ 3.51 \pm 0.05 \\ 3.58 \pm 0.03 \\ 0.4171 \end{array}$	$\begin{array}{c} 327.91\ ^{a}\pm 3.33\\ 316.52\ ^{b}\pm 4.02\\ 319.04\ ^{b}\pm 2.65\\ 0.0088 \end{array}$	$\begin{array}{c} 2.99 \pm 0.02 \\ 3.00 \pm 0.02 \\ 3.02 \pm 0.01 \\ 0.1168 \end{array}$
BLOCK3	1	H1H2 (239) H2H2 (580) H2H3 (48) <i>p</i> value	$\begin{array}{c} 10,\!194\pm73.90\\ 10,\!292\pm62.97\\ 10,\!439\pm118.76\\ 0.06 \end{array}$	$\begin{array}{c} 345.47 \ ^{b} \pm 3.22 \\ 346.25 \ ^{b} \pm 2.82 \\ 357.32 \ ^{a} \pm 4.93 \\ 0.0315 \end{array}$	$\begin{array}{c} 3.40 \pm 0.03 \\ 3.38 \pm 0.03 \\ 3.43 \pm 0.05 \\ 0.3858 \end{array}$	$\begin{array}{c} 305.64 \ ^{B} \pm 2.34 \\ 308.43 \ ^{AB} \pm 2.05 \\ 315.61 \ ^{A} \pm 3.59 \\ 0.0103 \end{array}$	$\begin{array}{c} 3.00 \pm 0.01 \\ 3.00 \pm 0.01 \\ 3.03 \pm 0.02 \\ 0.2938 \end{array}$
	2	H1H2 (176) H2H2 (370) H2H3 (37) <i>p</i> value	$\begin{array}{c} 10,\!456 {}^{\rm Ab} \pm 79.36 \\ 10,\!700 {}^{\rm Aa} \pm 67.79 \\ 10,\!860 {}^{\rm Aa} \pm 138.53 \\ 0.0005 \end{array}$	$\begin{array}{c} 312.66 \ ^{b} \pm 3.40 \\ 320.34 \ ^{a} \pm 2.97 \\ 327.69 \ ^{a} \pm 5.71 \\ 0.0052 \end{array}$	$\begin{array}{c} 3.58 \pm 0.03 \\ 3.55 \pm 0.03 \\ 3.60 \pm 0.06 \\ 0.3623 \end{array}$	$\begin{array}{c} 312.67 \ ^{C} \pm 2.50 \\ 320.34 \ ^{B} \pm 2.19 \\ 327.69 \ ^{A} \pm 4.18 \\ < 0.0001 \end{array}$	$\begin{array}{c} 2.99 \pm 0.01 \\ 3.00 \pm 0.01 \\ 3.02 \pm 0.02 \\ 0.496 \end{array}$

Table 4. Associations of haplotype blocks in the *CYP7A1* and *HADHB* genes with milk yield and milk-composition traits in Chinese Holstein cattle during first and second lactations.

The number in the table represents the mean \pm standard deviation; the number in the bracket represents the number of cows with the corresponding genotype; the *p* value shows the significance for the genetic effects of SNPs; the superscript letters indicate the significance calculated for the effects of different genotypes via the comparison of each pair of pairs; ^a, ^b within the same column with different superscripts indicates *p* < 0.05; and ^A, ^B, ^C within the same column with different superscripts indicates *p* < 0.01.

3.4. Changes of Transcription Factor Binding Sites Caused by SNPs in 5' Region

Using the JASPAR, changes in TFBS were predicted for four SNPs in the 5' regulatory region of the *CYP7A1* and *HADHB* genes. As a result, it was found that, in *CYP7A1*, the allele G of 14:g.24676921A>G created a binding site (BS) for transcription factors (TFs) EBF1 and SP1, while the mutation from G to T of 14:g.24675708G>T eliminated the BS for ELK1 but created a BS for FOXC1 and ZNF354C. For *HADHB*, before and after mutation, 11:g.73256269T>C resulted in the disappearance of the BS for PAX2 and NR2F1 and the appearance of that for TFAP2A; in addition, the allele A of 11:g.73256227A>C created a BS for ZNF354C and allele C created a BS for HINFP (Table 5).

3.5. mRNA Structural Stability Altered by the Mutation in Untranslated Region

Using RNAfold, changes in the mRNA minimum free energy (MFE) were predicted. The mutation of C to T in SNP 14:g.24665961C>T causes the MFE of mRNA to change from -578.10 kcal/mol to -539.20 kcal/mol, resulting in a decrease in the structural stability of *CYP7A1* mRNA.

Gene	SNP	Allele	Transcription Factor	Relative Score (≥0.80)	Predicted Binding Site Sequence
CYP7A1	14:g.24676921A>G	А	-	-	-
		G	EBF1 SP1	0.89 0.88	ATTCCA <u>G</u> GGA <u>G</u> GGACGGGG
	14:g.24675708G>T	G	ELK1	0.80	<u>G</u> GCACTGAAA
		14:g.24675708G>T	Т	FOXC1 ZNF354C	0.82 0.82
HADHB	11:g.73256269T>C	Т	PAX2 NR2F1	0.84 0.82	GG <u>T</u> CGTGC GAAGGG <u>T</u> CGTGCG
	~ <u>-</u>	0	С	TFAP2A	0.92

Table 5. Transcription factor binding sites (TFBSs) prediction for CYP7A1 and HADHB genes.

Relative scores represent the correlation score for each predicted binding site, with higher values indicating a stronger possibility of binding. The SNP in the predicted binding-site sequence is underlined.

4. Discussion

The CYP7A1 and HADHB genes have been implicated in milk-production traits in dairy cows, as indicated by recent research studies. CYP7A1 is known to influence triglyceride and cholesterol metabolism in the liver tissue of dairy cows and is a rate-limiting enzyme for bile-acid synthesis [18]; in chickens, HADHB plays a crucial role in liver lipid metabolism, particularly in inducing peroxisomal and mitochondrial β -oxidation activity [19]. These findings affirm the possible beneficial influence of these genes on milk production. Here, further identification of SNPs in the CYP7A1 and HADHB genes and analysis of their association with milk-production traits in dairy cows provide valuable information for understanding the mechanism of inheritance of these traits and linking the metabolic functions of these genes to their phenotypic effects. Furthermore, the integration of significant genetic sites from the CYP7A1 and HADHB genes into GS models can facilitate more precise selection of individuals with favorable milk-production traits. The purpose of these models is to predict the genetic advantage of the trait of interest based on the SNP profile of the individual, such that breeders can screen individuals for specific genetic variants associated with particular milk-production phenotypes, select them for breeding, and thus improve the efficiency of breeding.

In this study, we identified SNPs in the *CYP7A1* and *HADHB* genes and confirmed their significant associations with milk-production traits in dairy cows. SNPs can cause phenotypic variation by affecting the function or expression of genes involved in various biological processes [20–23]. This study did find significant differences in the milk-production phenotypes of cows with different SNP sites in the *CYP7A1* and *HADHB* genes. The number of individual cows with lower phenotypic values was also smaller, possibly because individuals of genotypes with higher-production phenotypes had been selected for while those with lower-production phenotypes had been eliminated during artificial breeding over the long term [24]. In addition, the phenotypic values of the cows and the significance of SNP/haplotype block associations were lower in the first lactation than in the second lactation. This difference could be explained by either of two possible causes. First, the varying numbers of cows in the two lactation periods may impact the statistical significance of the corresponding results. Second, there are physiological differences between the two lactation periods, as dairy cows tend to produce more milk in their second lactation [25].

Transcription factors are proteins that bind to specific DNA sequences and modulate the expression of target genes involved in various biological processes [26]. In this study, we found changes in TFBSs caused by SNPs in the 5' flanking regions of the *CYP7A1* and *HADHB* genes. These changes may result in altered regulation of gene expression by TFs, affecting the participation of the corresponding expression products in normal physiological and metabolic activities and thus, in turn, contributes to changes in milk production. For instance, in *HADHB*, TFAP2A is a specific nuclear transcription factor that may promote cell

proliferation by positively regulating certain cell-proliferation-associated proteins such as EIA, SV40, and c-myc [27]. This TF binding with 11:g.73256269C may promote expression of the *CYP7A1* gene and explain the significantly higher milk, fat and protein yields in the cows with genotype CC compared to those with genotype TT in the second lactation. Conversely, ZNF354C binds to form a number of transcriptional-repression complexes that inhibit the transcription of the intended target genes [28]. Binding of the *HDAHB* gene to this TF at 11:g.73256227A may result in the repression of its expression, which may have contributed to the significantly lower production of milk, fat, and milk proteins by the individuals with genotype AA than by those with CC in the second lactation.

Although the region is situated within gene exons, the term "untranslated region" (UTR) refers to the non-coding segment located at the mRNA molecule's terminus. Consequently, SNPs found within the UTR do not exert a direct influence on transcription and translation processes and are typically regarded as neutral variants. However, despite their lack of direct impact on encoded proteins, such SNPs in UTRs may change the mRNA's minimum free energy, leading to reduced mRNA stability. The synonymous mutation 14:g.24665961C>T, located in the untranslated region of the *CYP7A1* gene, reduces mRNA stability, resulting in reduced efficiency of transcription and translation of this gene, possibly affecting triglyceride and cholesterol metabolism [29,30], which may explain why the amount of milkfat yield in second lactation was significantly higher in the individuals with the genotype CC than in those with the genotype TT at the SNP site 14:g.24665961C>T.

5. Conclusions

This study identified polymorphisms in the *CYP7A1* and *HADHB* genes and their significant genetic effects on milk-production traits in Chinese Holstein cows, offering genetic markers for GS of dairy cattle; thus, four SNPs that altered TFBSs, 14:g.24676921A>G, 14:g.24675708G>T, 11:g.73256269T>C, and 11:g.73256227A>C, were proposed as possible causative mutations for the formation of milk-production traits. Further in-depth verification is needed before these SNPs can be used to provide genetic information to support the breeding of dairy cattle.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/ani14091276/s1, Table S1: Descriptive statistics of phenotypic values for dairy production traits in two lactations; Table S2: Primers and procedures for PCR used in SNPs identification of *CYP7A1* and *HADHB* genes.

Author Contributions: B.H. and D.S. conceived and designed the experiments, L.X. prepared the DNA samples for SNP identification and genotyping, A.C., Q.Y., W.Y. and Y.W. analyzed the data, and the manuscript was prepared by A.C. and B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the STI 2030-Major Projects (2023ZD04069); the National Key R&D Program of China (2021YFF1000700); the National Natural Science Foundation of China (32072716); Beijing Digital Agriculture Innovation Consortium Project (BAIC10-2023); the Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62); Youth Elite Development Program of College of Animal Science and Technology, China Agricultural University, China Agricultural University; China Agriculture Research System of MOF and MARA (CARS-36).

Institutional Review Board Statement: The animal study protocol was approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University (protocol code DK996).

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available in the article and its Supplementary Materials.

Acknowledgments: We appreciate Beijing Dairy Cattle Center for providing the samples (semen and blood) and the corresponding phenotypic data.

Conflicts of Interest: The authors declare that they have no competing interest.

References

- 1. Pereira, P.C. Milk nutritional composition and its role in human health. *Nutrition* **2014**, *30*, 619–627. [CrossRef]
- Khaw, K.-T.; Friesen, M.D.; Riboli, E.; Luben, R.; Wareham, N. Plasma phospholipid fatty acid concentration and incident coronary heart disease in men and women: The EPIC-Norfolk prospective study. *PLoS Med.* 2012, *9*, e1001255. [CrossRef]
- Vignal, A.; Milan, D.; SanCristobal, M.; Eggen, A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* 2002, *34*, 275–305. [CrossRef] [PubMed]
- Lalonde, E.; Ha, K.C.; Wang, Z.; Bemmo, A.; Kleinman, C.L.; Kwan, T.; Pastinen, T.; Majewski, J. RNA sequencing reveals the role of splicing polymorphisms in regulating human gene expression. *Genome Res.* 2011, 21, 545–554. [CrossRef] [PubMed]
- 5. Gong, J.; Tong, Y.; Zhang, H.M.; Wang, K.; Hu, T.; Shan, G.; Sun, J.; Guo, A.Y. Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Hum. Mutat.* **2012**, *33*, 254–263. [CrossRef]
- 6. Raven, L.-A.; Cocks, B.G.; Pryce, J.E.; Cottrell, J.J.; Hayes, B.J. Genes of the RNASE5 pathway contain SNP associated with milk production traits in dairy cattle. *Genet. Sel. Evol.* **2013**, *45*, 25. [CrossRef]
- 7. Fang, M.; Fu, W.; Jiang, D.; Zhang, Q.; Sun, D.; Ding, X.; Liu, J. A multiple-SNP approach for genome-wide association study of milk production traits in Chinese Holstein cattle. *PLoS ONE* **2014**, *9*, e99544. [CrossRef]
- 8. Ibeagha-Awemu, E.M.; Peters, S.O.; Akwanji, K.A.; Imumorin, I.G.; Zhao, X. High density genome wide genotyping-bysequencing and association identifies common and low frequency SNPs, and novel candidate genes influencing cow milk traits. *Sci. Rep.* **2016**, *6*, 31109. [CrossRef]
- 9. Pasandideh, M.; Mohammadabadi, M.; Esmailizadeh, A.; Tarang, A. Association of bovine PPARGC1A and OPN genes with milk production and composition in Holstein cattle. *Czech J. Anim. Sci.* **2015**, *60*, 97–104. [CrossRef]
- MacLeod, I.; Bowman, P.; Vander Jagt, C.; Haile-Mariam, M.; Kemper, K.; Chamberlain, A.; Schrooten, C.; Hayes, B.J.; Goddard, M.E. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genom.* 2016, 17, 144. [CrossRef]
- 11. Xu, L.; Shi, L.; Liu, L.; Liang, R.; Li, Q.; Li, J.; Han, B.; Sun, D. Analysis of liver proteome and identification of critical proteins affecting milk fat, protein, and lactose metabolism in dariy cattle with Itraq. *Proteomics* **2019**, *19*, 1800387. [CrossRef]
- Pullinger, C.R.; Eng, C.; Salen, G.; Shefer, S.; Batta, A.K.; Erickson, S.K.; Verhagen, A.; Rivera, C.R.; Mulvihill, S.J.; Malloy, M.J.; et al. Human cholesterol 7α-hydroxylase (CYP7A1) deficiency has a hypercholesterolemic phenotype. *J. Clin. Investig.* 2002, 110, 109–117. [CrossRef]
- Das, A.M.; Illsinger, S.; Lucke, T.; Hartmann, H.; Ruiter, J.P.; Steuerwald, U.; Waterham, H.R.; Duran, M.; Wanders, R.J.J.C.c. Isolated mitochondrial long-chain ketoacyl-CoA thiolase deficiency resulting from mutations in the HADHB gene. *Clin. Chem.* 2006, 52, 530–534. [CrossRef]
- Zhang, Q.; Boichard, D.; Hoeschele, I.; Ernst, C.; Eggen, A.; Murkve, B.; Pfister-Genskow, M.; Witte, L.A.; Grignola, F.E.; Uimari, P. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* 1998, 149, 1959–1973. [CrossRef]
- 15. Boichard, D.; Grohs, C.; Bourgeois, F.; Cerqueira, F.; Faugeras, R.; Neau, A.; Rupp, R.; Amigues, Y.; Boscher, M.Y.; Levéziel, H. Detection of genes influencing economic traits in three French dairy cattle breeds. *Genet. Sel. Evol.* **2003**, *35*, 77–101. [CrossRef]
- 16. Kučerová, J.; Lund, M.; Sørensen, P.; Sahana, G.; Guldbrandtsen, B.; Nielsen, V.; Thomsen, B.; Bendixen, C. Multitrait quantitative trait loci mapping for milk production traits in Danish Holstein cattle. *J. Dairy Sci.* 2006, *89*, 2245–2256. [CrossRef]
- Pedrosa, V.B.; Schenkel, F.S.; Chen, S.-Y.; Oliveira, H.R.; Casey, T.M.; Melka, M.G.; Brito, L.F. Genomewide association analyses of lactation persistency and milk production traits in Holstein cattle based on imputed whole-genome sequence data. *Genes* 2021, 12, 1830. [CrossRef]
- 18. Wu, D.; Gu, M.; Wei, Z.; Bai, C.; Su, G.; Liu, X.; Zhao, Y.; Yang, L.; Li, G. Myostatin knockout regulates bile acid metabolism by promoting bile acid synthesis in cattle. *Animals* **2022**, *12*, 205. [CrossRef]
- 19. Shao, F.; Wang, X.; Yu, J.; Shen, K.; Qi, C.; Gu, Z. Expression of miR-33 from an SREBP2 intron inhibits the expression of the fatty acid oxidation-regulatory genes CROT and HADHB in chicken liver. *Br. Poult. Sci.* **2019**, *60*, 115–124. [CrossRef]
- Cargill, M.; Altshuler, D.; Ireland, J.; Sklar, P.; Ardlie, K.; Patil, N.; Lane, C.R.; Lim, E.P.; Kalyanaraman, N.; Nemesh, J.; et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* 1999, 22, 231–238. [CrossRef]
- Kumar, V.; Westra, H.-J.; Karjalainen, J.; Zhernakova, D.V.; Esko, T.; Hrdlickova, B.; Almeida, R.; Zhernakova, A.; Reinmaa, E.; Võsa, U.; et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet.* 2013, 9, e1003201. [CrossRef]
- 22. Ryan, B.M.; Robles, A.I.; Harris, C.C. Genetic variation in microRNA networks: The implications for cancer research. *Nat. Rev. Cancer* 2010, *10*, 389–402. [CrossRef]
- Lango Allen, H.; Estrada, K.; Lettre, G.; Berndt, S.I.; Weedon, M.N.; Rivadeneira, F.; Willer, C.J.; Jackson, A.U.; Vedantam, S.; Raychaudhuri, S.; et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010, 467, 832–838. [CrossRef]
- 24. 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]
- 25. Gengler, N. Persistency of lactation yields: A review. Interbull Bull. 1996, 12, 87-96.
- 26. Latchman, D.S. Transcription factors: An overview. Int. J. Biochem. Cell Biol. 1997, 29, 1305–1312. [CrossRef]

- 27. Orso, F.; Corà, D.; Ubezio, B.; Provero, P.; Caselle, M.; Taverna, D. Identification of functional TFAP2A and SP1 binding sites in new TFAP2A-modulated genes. *BMC Genom.* 2010, *11*, 355. [CrossRef]
- Oo, J.A.; Irmer, B.; Günther, S.; Warwick, T.; Pálfi, K.; Izquierdo Ponce, J.; Teichmann, T.; Pflüger-Müller, B.; Gilsbach, R.; Brandes, R.P.; et al. ZNF354C is a transcriptional repressor that inhibits endothelial angiogenic sprouting. *Sci. Rep.* 2020, 10, 19079. [CrossRef]
- 29. Sabi, R.; Tuller, T. Modelling the efficiency of codon–tRNA interactions based on codon usage bias. *DNA Res.* **2014**, *21*, 511–526. [CrossRef]
- Hanson, G.; Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* 2018, 19, 20–30. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.