Food Science of Animal Products

Research Article

# Association of polymorphisms in *IGF2*, *CLU* and *STAT5A* genes with milk production characteristics in Chinese Holstein cattle

Shangchen Fu, Ting Ku, Linqiang Li, Yufang Liu ✉, Yongfeng Liu ✉

*School of Food Engineering and Nutritional Science, Shaanxi Normal University, Xi'an 710062, China*

✉Address correspondence to Yufang Liu, yufang.liu@snnu.edu.cn; Yongfeng Liu, yongfeng200@126.com

Received: February 10, 2023; Revised: March 17, 2023; Accepted: April 13, 2023

**Abstract:** Reflecting the quality of milk at the molecular level is a frontier technology. The aim of this study was to analyze the polymorphisms of bovine insulin-like growth factor 2 (*IGF2*), signal transducer and activator of transcription 5A (*STAT5A*) and clusterin (*CLU*) genes in the raw milk from 507 Chinese Holstein cow using polymerase chain reaction (PCR)-restriction fragment length polymorphism techniques and to evaluate their correlations with the milk protein content (MPC), milk fat content (MFC), milk lactose content (MLC) and milk total solids content (MTSC). In *IGF2* gene, genotype GG was the most frequent genotype (51.68%) followed by the genotype GT (38.03%) and TT (10.29%). And the genotype TT of *IGF2* gene was superior to the other genotypes in MPC. In *CLU* gene, genotype GG was the most common genotype (63.99%) followed by the genotype GA (34.45%) and AA (1.56%). And the genotype AA of *CLU* gene had greater MFC and MLC, but lower MTSC than GA genotype individuals. For *STAT5A* gene, the frequency of genotype CC and CT was similar (45.30% and 45.08%), while the genotype TT had lowest frequency (9.62%). And the genotype TT of *STA5A* gene had highest MPC and lowest MLC. Thus, screening for the *IGF2*, *CLU* and *STAT5A* genes were available for evaluating milk quality and raw milk samples were graded according to the different genotypes.

**Keywords:** milk quality; *STAT5A* gene; *CLU* gene; *IGF2* gene; mutation

**Citation:** S. C. Fu, T. Ku, L. Q. Li, et al., Association of polymorphisms in *IGF2*, *CLU* and *STAT5A* genes with milk production characteristics in Chinese Holstein cattle, Food Sci. Anim. Prod. 1 (2023) 9240011. https://doi.org/10.26599/FSAP.2023.9240011.

## 1 Introduction

Milk is regarded as a high-quality food in the worldwide and account for a large proportion in our daily diet, especially infants and children [1]. Monitoring raw milk quality traits is the key to produce high quality milk products. At present, phenomics is becoming an increasingly important research field, which is most frequently used to justify causal links between genotypes and environmental factors and phenotypes through large-scale data sets [2]. Recently, a growing number of studies on prediction and improvement of the composition and physical traits of food materials were performed using phenomics, which could be applied to justify causal correlations between genotypes and phenotypes through large-scale data sets [3,4]. Studies related to monitoring and improving milk protein, fat, lactose, minerals, vitamins and other traits through phenomics techniques have received extensive attention [3,5]. In correlated studies for milk composition, several bovine genes with multiple transcripts could be applied to select the targeted traits of milk, including insulin-like growth factor 2 (*IGF2*) gene, signal transducer and activator of transcription (*STAT*) gene and clusterin (*CLU*) gene [6–8].

IGF2 is a 67 amino acid monomeric protein and plays a vital role in somatotropic axis [9,10]. The *IGF2* gene locates on the telomeric end of the short arm of chromosome 29, and contains 10 exons and 3 promoters [11,12]. The *IGF2* gene is critical for the synthesis of milk. A number of mutant loci in the bovine *IGF2* gene showed significant correlations with milk quality indicators such as milk protein content (MPC), milk fat content (MFC), somatic cell count and milk total solids content (MTSC), and the *IGF2* gene has been identified as an important gene for total milk yield and quality traits [13,14]. Bagnicka et al. [15] found IGF2/Bsrl and IGF2/HaeIII polymorphisms in the Polish Holstein-Friesian cattle *IGF2* gene relative to milk traits, and Flisikowski et al. [6] further demonstrated that *IGF2* gene could be associated with cattle performance traits. Berkowicz et al. [16] examined four single nucleotide polymorphisms (SNPs) in the bovine *IGF2* gene located on chromosome 29, where *rs42196901* and *IGF2.g-3815A>G* SNPs were associated with milk fat and milk protein rates, respectively, and *rs42196909* was associated with milk yield in Irish Holstein cows. Hence, bovine *IGF2* gene is an important candidate gene for milk yield and quality evaluation, but how its mutations affect milk quality traits in Chinese Holstein cows remains uncertain.

CLU is a glycoprotein in physiological fluids including plasma, milk, urine, cerebrospinal fluid and semen, and is ubiquitously expressed in almost all mammalian tissues, and has been proposed to be a secreted mammalian chaperone [17,18]. Its predominant form is a secreted heterodimeric glycoprotein of 75–80 kDa [19]. CLU is also one of the most abundant proteins in buffalo milk [20]. In addition, CLU has a tendency to interact with a variety of molecules, including itself, lipids, amyloid proteins, components of the complement membrane attack complex, and immunoglobulins [21]. Wang et al. [22] analyzed the association between the genotype of the G15781A mutation locus on the cow *CLU* gene and milk quality indicators and found a significant correlation between it and MFC and somatic cell count. Besides, the expression of the bovine *CLU* gene is dramatically increased in cows with mastitis, and bovine CLU induces reactive

oxygen species-dependent apoptosis during mammary gland degeneration [23,24]. Therefore, *CLU* gene has become an important gene for milk quality evaluation.

STATs are a family of cytoplasmic proteins containing 7 structurally and functionally related proteins, which are activated in response to a large number of cytokines, growth factors and hormones, and are involved in cell proliferation, differentiation and apoptosis [25,26]. Among them, STAT5 exists in two isoforms (A and B), which differ by a few amino acids in the carboxylic end of the protein molecule. In cattle, *STAT5A* and *STAT5B* genes locate close to each other (within 40 kb) at chromosome 19 [27]. STAT5A was originally found in the mammary gland of goats that was involved in transducing prolactin (PRL) signalling in mammary epithelial cells and regulating promoter activity in the $\beta$-casein gene [28,29]. Miyoshi et al. [30] found that female mice with STAT5A deficiency fail to produce milk. Selvaggi et al. [29] assessed the association among *STAT5A/AvaI* polymorphism and some milk production traits in Jersey cows and found milk from TT animals had higher fat content than that of TC animals. Selvaggi et al. [31] also detected the effect of *STAT5A* gene variants on milk composition in Agerolese cattle, and found that TT genotype had a positive effect on milk fat and protein content. The *STAT5A* gene is at the core of the lactation related gene network in dairy cows, which can regulate the content of protein, fat, lactose and other nutritional indicators in milk. It can be concluded that the *STAT5A* gene is a vital candidate for regulating milk composition and is an important gene for assessing milk quality.

In this context, the present study extracted DNA from raw milk of Chinese Holstein cows, amplified *IGF2*, *STAT5A* and *CLU* genes with specific primers, obtained gene mutation site information by enzyme digestion and database search, and determined MPC, MFC, milk lactose content (MLC) and MTSC of raw milk. The relationship between *IGF2*, *STAT5A* and *CLU* genes and milk quality were investigated. The aim of this study was to reveal the association between *STAT5A*, *CLU* and *IGF2* genes and milk quality at the molecular level, thereby providing a theoretical basis for marker-assisted selection of dairy cows for milk production traits.

## 2 Materials and methods

### 2.1 Milk sampling and composition analysis

Milk samples were collected from 507 Chinese Holstein cows (all about 3 years old in a single herd) in Huayin Dairy Cattle Breeding Base of Modern Agricultural Comprehensive Development Corporation (Xi'an, Shaanxi, China). All Chinese Holstein cows were healthy and bovine milk was collected in the second or third month of lactation. Prior to the milk collection, the udder surface of each cow was cleaned with sterile water and rinsed with 70% ethanol. After the disposal of the first three drops, milk samples were put into a sterile tube, transported on ice packs to the laboratory, and stored at –80 °C until use. Milk composition, including milk protein, fat, lactose and total solids, was determined using a milk composition analyzer (Lactoscan LA, MCC, Bulgaria). The detection of main components was conducted in triplicates.

### 2.2 DNA extraction

The extraction of bovine genomic DNA from raw milk was performed according to our previous study [32]. Briefly, 50 microliters cryopreserved milk samples were centrifuged at 2 415 × *g* for 20 min at 4 °C and the upper and middle layers were removed carefully. The sediment was suspended in 600 μL of phosphate buffered saline (PBS, pH 7.4) and centrifuged at 9 660 × *g* for 10 min at 4 °C. The sediment was washed with PBS for three times. Then the sediment was mixed with 350 μL of extraction buffer (NaCl 1 mol/L, Tris-HCl 0.5 mol/L, EDTA-Na 0.5 mol/L, pH 8.0), 50 μL of 20 g/100 mL sodium dodecyl sulfate (SDS), and 10 μL of 20 mg/mL proteinase K, and was incubated at 56 °C for 4 h. The mixture was centrifuged at 9 660 × *g* for 10 min at 4 °C, and the supernatant was collected and mixed with an equal volume of phenol-chloroform-isoamyl alcohol mixture (volume ratio of 25:24:1). The mixture was inverted repeatedly to dissolve precipitation for 10 min and centrifuged at 9 660 × *g* for 10 min at 4 °C. The supernatant was transferred to a 1.5 mL plastic tube with an equal volume of chloroform-isoamyl alcohol mixture (volume ratio of 24:1), inverted repeatedly to dissolve the precipitate and centrifuged at 9 660 × *g* for 10 min at 4 °C. The supernatant was mixed with 0.8 mL ice-cold anhydrous ethanol, gently shaking, and the mixture was incubated for 30 min at –20 °C. Then, the mixture was centrifuged at 9 660 × *g* for 10 min at 4 °C. After the removal of upper ethanol layer, the sediment was washed with 70% ice-cold ethanol and centrifuged at 9 660 × *g* for 10 min at 4 °C. The upper ethanol layer was carefully removed and 25 μL Tris-EDTA (pH 8.0) was added to dissolve DNA. The mixture was stored at –20 °C before the use.

### 2.3 Assessment of DNA quality

The purity and concentration of extracted DNA were detected by determining the optical density at 160 nm and 280 nm using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA quality and size were evaluated by 2% agarose gel electrophoresis at a voltage of 80 V for 40 min. The genomic DNA was diluted in sterile distilled water to 1 ng/μL as the template of polymerase chain reaction (PCR) experiment.

### 2.4 PCR-restriction fragment length polymorphism (PCR-RFLP) amplification and genotype

PCR amplification for the *STAT5A* and *CLU* genes were performed in triplicate with a reaction mixture containing 1.5 mmol/L $MgCl_2$, 50 μmol/L dNTP mix, 0.3 μmol/L of each primer, 1 × PCR buffer, 2 U *Taq* polymerase and 50 ng of genomic DNA template. PCR reaction for *IGF2* gene was carried out in 10 μL reaction mixture, which contained 1 μL template DNA, 0.4 μmol/L of assay-specific primer, 3.4 μL 2 × *Taq*Man Universal PCR Master Mix (CWBIO, Beijing, China) and DNase-/DNA-free water (CWBIO). The primers for *STAT5A*, *CLU* and *IGF2* genes were shown in Table 1. The PCR protocol was 94 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 59 °C annealing for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 8 min.

**Table 1** Primers used for PCR amplification and sequencing of the bovine *IGF2*, *STAT5A* and *CLU* genes.

| Gene | Primer sequence (5′ to 3′) | Product size (bp) |
|---|---|---|
| *IGF2* | Forward: AATCCCTGTACCGTCCTGTC<br>Reverse: TTTGCTTTTCTGGTGTTGCT | 273 |
| *STAT5A* | Forward: CTGCAGGGCTGTTCTGAGAG<br>Reverse: TGGTACCAGGACTGTACACAT | 215 |
| *CLU* | Forward: CTGCCTCCCTCAAGGTGCCCCTC<br>Reverse: AGTTCTTGGCTAAATGTCTTAGGGGG | 260 |

Typing of mutant loci in genes using PCR-RFLP. The 215 bp PCR products of *STAT5A* gene were purified using a Wizard Prep PCR purification kit (Shanghai Bioasia Biotechnology, Shanghai, China) and then digested at 37 °C for 8 h by *Ava* I restriction endonuclease (10 U, MBI Fermentas, Lithuania). The amplified 260 bp PCR products of *CLU* gene were also purified with Wizard Prep PCR purification kit and then digested for 6 h at 37 °C with the *Hinf* I restriction endonuclease (10 U, MBI Fermentas, Lithuania). The PCR products of *IGF2* gene were purified with Wizard Prep PCR purification kit and then digested with 5 U *Hae* III (Promega) at 37 °C for 8 h following the manufacturer's manual. The concentration of digested products was detected by 3.0% agarose gel electrophoresis.

### 2.5 Statistical analyses

According to our previous report [33], the genotype frequency, allele frequency, gene heterozygosity and polymorphism information content (PIC) were calculated by formulae (1)–(7):

$$\text{TT (GG) genotype frequency (\%)} = \frac{M}{Z} \times 100 \quad (1)$$

$$\text{TC (GA) genotype frequency (\%)} = \frac{N}{Z} \times 100 \quad (2)$$

$$\text{CC (AA) genotype frequency (\%)} = \frac{K}{Z} \times 100 \quad (3)$$

$$\text{T (G) allele frequency (\%)} = \frac{M \times 2 + N}{Z \times 2} \times 100 \quad (4)$$

$$\text{C (A) allele frequency (\%)} = \frac{K \times 2 + N}{Z \times 2} \times 100 \quad (5)$$

$$\text{Gene heterozygosity (\%)} = 1 - X \times X - Y \times Y \quad (6)$$

$$\text{PIC (\%)} = 1 - X \times X - Y \times Y - 2X \times X \times Y \times Y \quad (7)$$

Where $M$ is the number of TT (GG) milk samples, $N$ is the number of TC (GA) milk samples, $K$ is the number of CC (AA) milk samples, $Z$ is the total number of milk samples, $X$ is the frequency of the T (G) allele, $Y$ is the frequency of the C (A) allele.

The general linear model program of SPSS 17.0 statistical software was used to analyze the relationship between *IGF2*, *STAT5A* and *CLU* genes polymorphisms and milk quality traits (MFC, MPC, MLC and MTSC) in Chinese Holstein cows using least-squares method.

A value of $P < 0.05$ was considered to be statistically significant. Results were expressed as mean ± standard error of the mean (SEM).

## 3 Results

### 3.1 Genotyping of *IGF2*, *STAT5A* and *CLU* genes

DNA samples from bovine milk were extracted and analyzed for polymorphisms of nucleotide sequence using the PCR technique. The PCR amplified a 233 bp fragment at 24 507 G>T on exon 10 of the *IGF2* gene. As shown in Figure 1, this fragment could be digested into three genotypes by the restriction endonucleases (*Ava* I enzyme): TT, GT and GG. The genotype TT had fragment size of 233 and 43 bp, GT had fragments of 233, 190 and 43 bp, and GG had fragments of 190 and 43 bp.

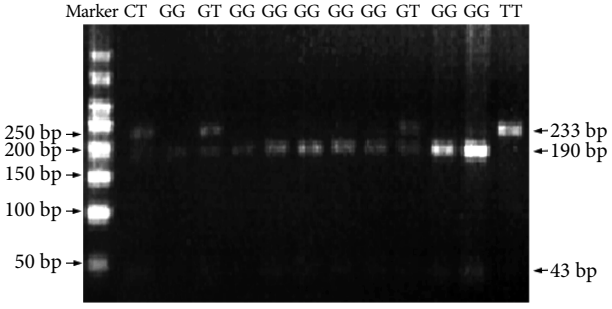


**Figure 1** Electrophoretic patterns of polymorphic loci in the bovine *IGF2* gene at 24 507 G>T mutation site. GG genotype demonstrates two fragments (190 and 43 bp), GT genotype shows three fragments (233, 190 and 43 bp) and TT genotype shows two fragments (233 and 43 bp).

For *CLU* gene, the 260 bp fragment at mutant site 15 781 G>A on exon 7 was amplified by PCR and digested using the restriction endonucleases (*Hinf* I enzyme) into three genotypes: AA, GG and GA. The genotype AA had fragment size of 260 bp, GG had fragments of 149 and 111 bp, and GA had fragments of 260, 149 and 111 bp (Figure 2).

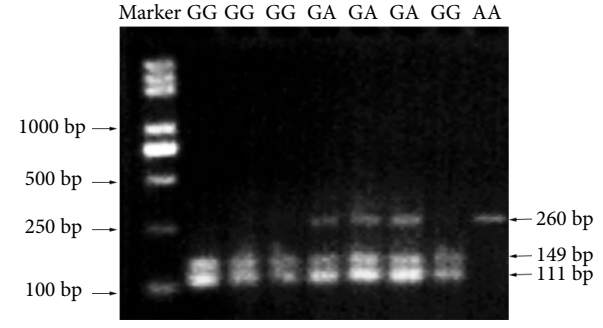


**Figure 2** Electrophoretic patterns of polymorphic loci in the bovine *CLU* gene at 15 781 G>A mutation site. GG genotype demonstrates two fragment (149 and 111 bp), GA genotype shows three fragments (260, 149 and 111 bp) and AA genotype shows one fragments (260 bp).

For *STAT5A* gene, the amplified fragment with 215 bp located on mutant site 6 853 C>T of exon 7. The fragments after *Ava* I enzyme digestion of 215 bp PCR product were 215 bp for TT genotype, 181 and 34 bp for CC genotype, and 215, 181 and 34 bp for CT genotype (Figure 3).

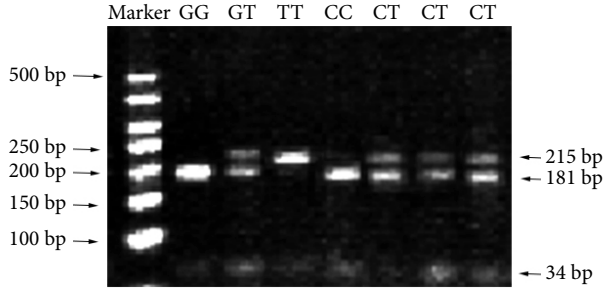


**Figure 3** Electrophoretic patterns of polymorphic loci in the bovine *STAT5A* gene at 6 853 C>T mutation site. CC genotype demonstrates two fragment (181 and 34 bp), CT genotype shows three fragments (215, 181 and 34 bp) and TT genotype shows one fragments (215 bp).

### 3.2 Gene frequency

The data set in this study included 507 Chinese Holstein cows. In *IGF2* gene of the Holstein cows, genotype GG was the most

**Table 2** Proportion of genotypes, allele frequencies and PIC of *IGF2* gene for the 24 507 G>T mutation, *STAT5A* gene for the 6 853 C>T mutation, and *CLU* gene for the 15 781 G>A mutation in milk samples of Chinese Holstein cows.

| Gene | Mutation sites | Proportion of genotypes (%) (number) | | | Total | Allelic frequencies (%) | | PIC |
|---|---|---|---|---|---|---|---|---|
| *IGF2* | 24 507 G>T | GG: 51.68 (231) | GT: 38.03 (170) | TT: 10.29 (46) | 507 | G: 71 | T: 29 | 0.33 |
| *STAT5A* | 6 853 C>T | CC: 45.30 (212) | CT: 45.08 (211) | TT: 9.62 (24) | 457 | C: 68 | T: 32 | 0.34 |
| *CLU* | 15 781 G>A | GG: 63.99 (247) | GA: 34.45 (133) | AA: 1.56 (6) | 386 | G: 81 | A: 19 | 0.29 |

**Table 3** Association of gene mutation with milk quality traits in the population of milk.

| Gene | Mutation sites | Position | Milk quality | Proportion of genotypes (%) | | | *P* value |
|---|---|---|---|---|---|---|---|
| | | | | GG | GT | TT | |
| *IGF2* | 24 507 G>T | Exon 10 | MFC | 4.06 ± 0.08 | 4.14 ± 0.09 | 4.17 ± 0.15 | 0.83 |
| | | | MPC | 2.98 ± 0.02$^{b}$ | 2.99 ± 0.03$^{b}$ | 3.29 ± 0.05$^{a}$ | 0.04 |
| | | | MLC | 5.07 ± 0.02 | 5.03 ± 0.03 | 5.07 ± 0.08 | 0.69 |
| | | | MTSC | 12.12 ± 0.16 | 12.32 ± 0.15 | 12.48 ± 0.38 | 0.65 |
| *STAT5A* | 6 853 C>T | Exon 7 | MFC | 4.11 ± 0.08 | 4.08 ± 0.08 | 4.03 ± 0.17 | 2.70 |
| | | | MPC | 2.99 ± 0.02$^{b}$ | 3.00 ± 0.02$^{b}$ | 3.21 ± 0.04$^{a}$ | 0.04 |
| | | | MLC | 5.06 ± 0.02$^{a}$ | 5.07 ± 0.03$^{a}$ | 4.55 ± 0.06$^{b}$ | 0.03 |
| | | | MTSC | 12.32 ± 0.16 | 12.03 ± 0.14 | 12.49 ± 0.29 | 0.65 |
| *CLU* | 15 781 G>A | Exon 7 | MFC | 4.05 ± 0.09$^{b}$ | 4.09 ± 0.18$^{b}$ | 4.39 ± 0.18$^{a}$ | 0.03 |
| | | | MPC | 2.96 ± 0.03 | 3.01 ± 0.06 | 3.09 ± 0.07 | 2.04 |
| | | | MLC | 5.10 ± 0.02$^{ab}$ | 4.96 ± 0.07$^{b}$ | 5.12 ± 0.05$^{a}$ | 0.04 |
| | | | MTSC | 12.22 ± 0.12$^{ab}$ | 12.67 ± 0.42$^{a}$ | 11.57 ± 0.30$^{b}$ | 0.04 |

Note: a, b within a same row with different superscript letters are significantly different ($P < 0.05$).

frequent genotype (51.68%), followed by the genotype GT (38.03%), while the genotype TT had the lowest frequency (10.29%). The frequency of the G and T alleles was 71% and 29%, respectively. In *CLU* gene, genotype GG was the most common genotype (63.99%), followed by the genotype GA (34.45%) and AA (1.56%). The frequency of the G and A alleles in *CLU* gene was 81% and 19%, respectively. For *STAT5A* gene, the frequency of genotype CC and CT was similar with the value of 45.30% and 45.08%, respectively, while the genotype TT had lowest frequency (9.62%) (Table 2).

### 3.3 Association of the gene polymorphisms with milk production traits in Chinese Holstein cows

Data reported in Table 3 displayed the effect of three gene polymorphisms (*IGF2*, *CLU* and *STAT5A* genes) on milk production traits. The result from association analysis showed that MPC was significantly correlated with the gene polymorphism at the mutant site 24 507 G>A on exon 10 of *IGF2* gene. Compared with GG (2.96%–3.00%) and GT (2.97%–3.01%) cows, TT (3.24%–3.34%) individuals produced the milk with a higher content of protein significantly ($P < 0.05$). Milk from TT cows also had higher fat, lactose and solid content than those from GG and GT cows, but not statistically significant. No correlation was displayed between the genotypes at 24 507 G>A and MFC, MLC and MTSC.

Three genotypes (GG, GA, and AA) were detected for *CLU* gene and GG was the most frequent genotype (63.99%). The relationships among *CLU* gene polymorphism and MFC, MLC and MTSC were found to be significant ($P < 0.05$). The AA genotyped cows had greater MFC and MLC, but lower MTSC than GA genotyped individuals. Relationships of *CLU* gene polymorphism with MPC were not detected to be significant.

For *STAT5A* gene, gene polymorphism had significant relationships with MPC, MLC and MTSC. The TT genotyped individuals produced bovine milk with highest MPC and lowest MLC and MTSC. Significant relationship was not found between *STAT5A* gene polymorphism and MFC.

## 4 Discussion

In this study, we identified the polymorphisms and genotype frequency of three genes (*IGF2*, *CLU*, and *STAT5A* genes) and analyzed the relationship among the polymorphism of *IGF2*, *CLU*, and *STAT5A* genes and milk production traits in Chinese Holstein cows. Milk production traits highly depend on the genotype and environment, such as feeding conditions and feedstuff [31]. It was reported that these milk traits were controlled by quantitative trait loci, multiple genotypes and scattered throughout the genome [3]. The selection of superior genotypes could be expected to provide offspring with a greater frequency of the relevant alleles and as a result, increased production traits as compared with the average population. Hence, identification of the genes that determine these traits and the polymorphism in these genes that influent the phenotypic amplification of these traits play an important role on the improvement in the production traits of nutritional importance during the selection of animals.

IGF2, a 67-amino-acid monomeric protein, has been reported as a local regulator of lactogenesis and mammogenesis [15]. In dairy goat, local infusion of IGF2 could enhance milk synthesis. In a previous study, Hovey et al. [34] demonstrated that the influence of prolactin on the mammary gland was regulated by endogenously secreted IGF2 and IGF2 also could stimulate alveolar growth in

mammary gland of mice and prolactin could induce the expression of IGF2 in mammary gland explants. Hence, *IGF2* was considered as a candidate gene for the determination of molecular marker correlated with milk characteristics. Ramadan et al. [35] reported that the AA genotype of *IGF2* gene displayed the superiority over the other genotype for milk yield in Sinai Gabali rabbits. Flisikowski et al. [6] detected a single nucleotide C deletion/insertion polymorphism in *Bos taurus* and *Bos indicus* cattle and this mutant site was significantly associated with MPC. In the present study, there was a significant relationship between 24 507 G>T mutation of *IGF2* gene and MPC according to association analysis, and MPC in milk samples with genotype TT ((3.29 ± 0.05)%) was higher than that with genotype GG and GT ((2.98 ± 0.02)% and (2.99 ± 0.03)%). This result was consistent with the findings of Bagnicka et al. [15]. However, their study also found that the genotype of the 24 507 G>T mutation locus also exhibited some correlation with MFC [15]. In addition, the frequencies of the alleles G and T of the 24 507 G>T mutation in cow in this study were 71% and 29%, respectively. And the frequency of the GG genotype was 51.68%, the GT type was 38.03%, and the TT type was 10.29%. The above trend was generally consistent with the study of Bagnicka et al. [15] in which the allele frequencies for G and T were 84% and 16%, respectively, with 69% for the GG genotype, 30% for the GT genotype and 1% for the TT genotype. The differences between the results of this study and those of Bagnicka et al. [15] might be related to factors such as the source of the milk samples and the season in which the milk samples were collected [36]. Overall, the results of this study indicated the potential role of the 24 507 G>T polymorphic locus of the *IGF2* gene in predicting MPC and that the TT genotype had a positive effect on MPC in Chinese Holstein cows.

CLU is a heterodimeric 80 kDa glycoprotein and secreted by ram Sertoli cells. French et al. [37] analyzed the expression of clusterin in mammary gland and the result showed that clusterin mRNA was increased during pregnancy and significantly down-regulated during lactation. CLU has a tendency to interact with a variety of molecules, including itself, lipids, amyloid proteins, components of the complement membrane attack complex, and immunoglobulins. In recent years, *CLU* gene has become an important gene for milk quality evaluation because its single nucleotide polymorphisms associate with milk production traits including milk yield, fat, and protein contents in Chinese Holstein cows. Thus, it is of significance to evaluate the influence of *CLU* gene on milk characteristics for improving the milk quality in Chinese Holstein cows. It was found in this study that the 15 781 G>A mutation in the *CLU* gene could be used for assessment of MFC, MLC and MTSC. The MFC and MLC of the AA genotype milk reach the highest levels, which are (4.39 ± 0.18)% and (5.12 ± 0.05)%, respectively. Wang et al. [22] also found that the MFC of AA genotype cows was higher than that of GG genotype in the CLU G+15781A polymorphism. Therefore, A allele at 15 781 G>A mutation tended to have increased MFC and MLC. Meanwhile, it reduced MTSC, which in AA genotype milk is the lowest among the three genotypes. Thus, milk with AA genotype would seem particularly suited for producing high-fat and high-lactose dairy products.

Transcription factors STAT5 have been considered as members of the somatotropic axis, which may contribute to progress in genetic selection of farm animals. In fact, the *STAT5A* gene is the first found to be associated with both milk production and fertility [38]. We found that the 6 853 C>T mutation in *STAT5A* gene could be applied to evaluate milk quality. The sequence of MPC in the three genotypes of milk was TT ((3.21 ± 0.04)%), CT ((3.00 ± 0.02)%), CC ((2.99 ± 0.02)%), while MLC with genotype TT ((4.55 ± 0.06)%) was lower than the other genotypes (CT for (5.07 ± 0.03)%, and CC for (5.06 ± 0.02)%). Sadeghi et al. [39] found a significantly higher quantity of protein yielded by CT animals when compared with CC individuals. Therefore, we inferred that the T allele had a positive impact on milk protein but a negative influence on lactose in milk. Although the TT genotype produced lower lactose, the protein content was the highest in three genotypes. The 15 781 G>A transition was a synonymous mutation located at coding regions (1 036 bp) in exon 7, which may affect gene mRNA splicing or transcription process.

In the present study, we detected the polymorphisms of *IGF2*, *CLU* and *STAT5A* genes in raw milk from Chinese Holstein cows and assessed their correlations with milk characteristics. The TT genotype of *IGF2* gene displayed the superiority over the other genotypes in MPC. The AA genotype of *CLU* gene had greater MFC and MLC, but lower MTSC than GA genotyped individuals. For *STA5A* gene, the TT genotype individuals yielded bovine milk with highest MPC and lowest MLC. Although further research should be performed to confirm these results, the correlation of genetic markers with better milk characteristics is a significant finding that could be applied to improve milk quality.

## Conflict of interest

The authors declare that they have no conflict of interest.

## Acknowledgements

This work was supported by Shaanxi science and technology plan projects of China (2022KXJ-010, 2022ZDLNY04-09); Science and technology plan projects in Xianyang city of Shaanxi Province (2021ZDZX-NY-0014); and the Xi'an city science and technology plan projects of China (22NYGG0012).

## References

[1] D. Porcellato, M. Aspholm, S. B. Skeie, et al., Microbial diversity of consumption milk during processing and storage, Int. J. Food Microbiol. 266 (2018) 21–30. https://doi.org/10.1016/j.ijfoodmicro.2017.11.004.

[2] D. Houle, D. R. Govindaraju, S. Omholt, Phenomics: the next challenge, Nat. Rev. Genet. 11(12) (2010) 855–866. https://doi.org/10.1038/nrg2897.

[3] N. S. Yudin, M. I. Voevoda, Molecular genetic markers of economically important traits in dairy cattle, Russ. J. Genet. 51(5) (2015) 600–612. https://doi.org/10.1134/S1022795415050087.

[4] Z. Wang, T. Li, W. Yu, et al., Determination of content of camel milk in adulterated milk samples by normalized real-time polymerase chain reaction system based on single-copy nuclear genes, J. Sci. Food Agr. 100(8) (2020) 3465–3470. https://doi.org/10.1002/jsfa.10382.

[5] M. G. Pizarro Inostroza, F. J. Navas González, J. M. León Jurado, et al., Bayesian evaluation of the effect of non-genetic factors on the phenomics for quality-related milk nutrients and yield in Murciano-Granadina goats, Trop. Anim. Health Pro. 54(6) (2022) 388. https://doi.org/10.1007/s11250-022-03385-3.

[6] K. Flisikowski, A. Maj, L. Zwierzchowski, et al., Nucleotide sequence and variation of *IGF2* gene exon 6 in *Bos taurus* and *Bos indicus* cattle, Anim. Biotechnol. 16(2) (2005) 203–208. https://doi.org/10.1080/10495390500278060.

[7] X. He, M. X. Chu, L. Qiao, et al., Polymorphisms of *STAT5A* gene and their association with milk production traits in Holstein cows, Mol. Biol. Rep. 39(3) (2012) 2901–2907. https://doi.org/10.1007/s11033-011-1051-4.

[8] C. Li, W. Cai, C. Zhou, et al., RNA-Seq reveals 10 novel promising candidate genes affecting milk protein concentration in the Chinese Holstein population, Sci. Rep. 6(1) (2016) 26813. https://doi.org/10.1038/srep26813.

[9] L. Yan, X. Fang, Y. Liu, et al., Exploring the genetic variants of insulin-like growth factor II gene and their associations with two production traits in Langshan chicken, J. Appl. Anim. Res. 45(1) (2017) 60–63. https://doi.org/10.1080/09712119.2015.1124328.

[10] L. A. Vaccaro, T. E. Porter, L. E. Ellestad, The effect of commercial genetic selection on somatotropic gene expression in broilers: a potential role for insulin-like growth factor binding proteins in regulating broiler growth and body composition, Front Physiol. 13 (2022) 935311. https://doi.org/10.3389/fphys.2022.935311.

[11] N. Ma, X. Wang, Y. Qiao, et al., Coexpression of an intronic microRNA and its host gene reveals a potential role for miR-483-5p as an IGF2 partner, Mol. Cell Endocrinol. 333(1) (2011) 96–101. https://doi.org/10.1016/j.mce.2010.11.027.

[12] Z. Huang, S. Murphy, Increased intragenic IGF2 methylation is associated with repression of insulator activity and elevated expression in serous ovarian carcinoma, Front Oncol. 3 (2013) 131. https://doi.org/10.3389/fonc.2013.00131.

[13] G. A. Rohrer, D. J. Nonneman, R. K. Miller, et al., Association of single nucleotide polymorphism (SNP) markers in candidate genes and QTL regions with pork quality traits in commercial pigs, Meat Sci. 92(4) (2012) 511–518. https://doi.org/10.1016/j.meatsci.2012.05.020.

[14] Y. Gao, J. Jiang, S. Yang, et al., CNV discovery for milk composition traits in dairy cattle using whole genome resequencing, BMC Genomics 18(1) (2017) 265. https://doi.org/10.1186/s12864-017-3636-3.

[15] E. Bagnicka, E. Siadkowska, N. Strzałkowska, et al., Association of polymorphisms in exons 2 and 10 of the insulin-like growth factor 2 (*IGF2*) gene with milk production traits in Polish Holstein-Friesian cattle, J. Dairy Res. 77(1) (2010) 37–42. https://doi.org/10.1017/s0022029909990197.

[16] E. W. Berkowicz, D. A. Magee, K. M. Sikora, et al., Single nucleotide polymorphisms at the imprinted bovine insulin-like growth factor 2 (*IGF2*) locus are associated with dairy performance in Irish Holstein-Friesian cattle, J. Dairy Res. 78(1) (2011) 1–8. https://doi.org/10.1017/S0022029910000567.

[17] D. C. Park, S. Geun Yeo, E. Young Shin, et al., Clusterin confers paclitaxel resistance in cervical cancer, Gynecol. Oncol. 103(3) (2006) 996–1000. https://doi.org/10.1016/j.ygyno.2006.06.037.

[18] J. K. Shin, K. A. Han, M. Y. Kang, et al., Expression of clusterin in normal and preeclamptic placentas, J. Obstet. Gynaecol. Res. 34(4) (2008) 473–479. https://doi.org/10.1111/j.1447-0756.2008.00723.x.

[19] D. Bradley, A. Blaszczak, Z. Yin, et al., Clusterin impairs hepatic insulin sensitivity and adipocyte clusterin associates with cardiometabolic risk, Diabetes Care 42(3) (2019) 466–475. https://doi.org/10.2337/dc18-0870.

[20] Y. Yang, D. Bu, X. Zhao, et al., Proteomic analysis of cow, yak, buffalo, goat and camel milk whey proteins: quantitative differential expression patterns, J. Phys. Chem. Lett. 12(4) (2013) 1660–1667. https://doi.org/10.1021/pr301001m.

[21] C. Rodríguez-Rivera, M. M. Garcia, M. Molina-Álvarez, et al., Clusterin: always protecting. Synthesis, function and potential issues, Biomed. Pharmacother. 134 (2021) 111174. https://doi.org/10.1016/j.biopha.2020.111174.

[22] Z. Wang, J. Huang, J. Zhong, et al., Molecular cloning, promoter analysis, SNP detection of clusterin gene and their associations with mastitis in Chinese Holstein cows, Mol. Biol. Rep. 39(3) (2012) 2439–2445. https://doi.org/10.1007/s11033-011-0994-9.

[23] M. Nakamura, A. Tomita, H. Nakatani, et al., Antioxidant and antibacterial genes are upregulated in early involution of the mouse mammary gland: sharp increase of ceruloplasmin and lactoferrin in accumulating breast milk, DNA Cell Biol. 25(9) (2006) 491–500. https://doi.org/10.1089/dna.2006.25.491.

[24] E. Khalil, M. R. Digby, P. C. Thomson, et al., Acute involution in the tammar wallaby: identification of genes and putative novel milk proteins implicated in mammary gland function, Genomics 97(6) (2011) 372–378. https://doi.org/10.1016/j.ygeno.2011.03.003.

[25] A. Subramaniam, M. K. Shanmugam, E. Perumal, et al., Potential role of signal transducer and activator of transcription (STAT) 3 signaling pathway in inflammation, survival, proliferation and invasion of hepatocellular carcinoma, BBA-Reviews on Cancer 1835(1) (2013) 46–60. https://doi.org/10.1016/j.bbcan.2012.10.002.

[26] K. S. Siveen, S. Sikka, R. Surana, et al., Targeting the STAT3 signaling pathway in cancer: role of synthetic and natural inhibitors, BBA-Reviews on Cancer 1845(2) (2014) 136–154. https://doi.org/10.1016/j.bbcan.2013.12.005.

[27] Y. Verhoeven, S. Tilborghs, J. Jacobs, et al., The potential and controversy of targeting STAT family members in cancer, Semin. Cancer Biol. 60 (2020) 41–56. https://doi.org/10.1016/j.semcancer.2019.10.002.

[28] P. Brym, S. Kamiński, A. Ruść, New SSCP polymorphism within bovine *STAT5A* gene and its associations with milk performance traits in black-and-white and jersey cattle, J. Appl. Genet. 45(4) (2004) 445–452.

[29] M. Selvaggi, C. Dario, G. Normanno, et al., Genetic polymorphism of STAT5A protein: relationships with production traits and milk composition in Italian Brown cattle, J. Diary Res. 76(4) (2009) 441–445. https://doi.org/10.1017/S0022029909990070.

[30] K. Miyoshi, J. M. Shillingford, G. H. Smith, et al., Signal transducer and activator of transcription (Stat) 5 controls the proliferation and differentiation of mammary alveolar epithelium, J. Cell Biol. 155(4) (2001) 531–542. https://doi.org/10.1083/jcb.200107065.

[31] M. Selvaggi, S. Albarella, C. Dario, et al., Association of *STAT5A* gene variants with milk production traits in agerolese cattle, Biochem. Genet. 55(2) (2017) 158–167. https://doi.org/10.1007/s10528-016-9781-6.

[32] Y. F. Liu, J. L. Gao, Y. F. Yang, et al., Novel extraction method of genomic DNA suitable for long-fragment amplification from small amounts of milk, J. Dairy Sci. 97(11) (2014) 6804–6809. https://doi.org/10.3168/jds.2014-8066.

[33] Y. Liu, L. Zan, Y. Xin, et al., *ZBTB38* gene polymorphism associated with body measurement traits in native Chinese cattle breeds, Gene 513 (2012) 272–277. https://doi.org/10.1016/j.gene.2012.10.026.

[34] R. Hovey, J. Harris, D. Hadsell, et al., Local insulin-like growth factor-II mediates prolactin-induced mammary gland development, Mol. Endocrinol. 17 (2003) 460–471. https://doi.org/10.1210/me.2002-0214.

[35] S. Ramadan, E. Manaa, M. Elatrouny, et al., Association of growth hormone (*GH*), insulin-like growth factor 2 (*IGF2*) and progesterone receptor (*PGR*) genes with some productive traits in Gabali rabbits, World Rabbit. Sci. 28 (2020) 135–144. https://doi.org/10.4995/wrs.2020.12610.

[36] G. M. Vacca, G. Stocco, M. L. Dettori, et al., Milk yield, quality, and coagulation properties of 6 breeds of goats: environmental and individual variability, J. Diary Sci. 101(8) (2018) 7236–7247. https://doi.org/10.3168/jds.2017-14111.

[37] L. French, J. Soriano, R. Montesano, et al., Modulation of clusterin gene expression in the rat mammary gland during pregnancy, lactation, and involution, Biol. Reprod. 55 (1997) 1213–1220. https://doi.org/10.1095/biolreprod55.6.1213.

[38] H. Khatib, R. L. Monson, V. Schutzkus, et al., Mutations in the *STAT5A* gene are associated with embryonic survival and milk composition in cattle, J. Diary Sci. 91(2) (2008) 784–793. https://doi.org/10.3168/jds.2007-0669.

[39] M. Sadeghi, M. M. Shahrbabak, G. R. Mianji, et al., Polymorphism at locus of *STAT5A* and its association with breeding values of milk production traits in Iranian Holstein bulls, Livest. Sci. 123(1) (2009) 97–100. https://doi.org/10.1016/j.livsci.2008.10.010.