A 17-bp InDel (rs668420586) within goat *CHCHD7* gene located in growth-related QTL affecting body measurement traits

Haixia Li¹ · Xinyu Wang¹ · Hong Chen¹ · Lei Qu^{2,3} · Xianyong Lan¹

Received: 14 May 2020 / Accepted: 8 September 2020 / Published online: 18 September 2020
 © King Abdulaziz City for Science and Technology 2020

Abstract

The Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 7 (*CHCHD7*) gene was located in a growth-related major QTL that participated in the process of bone cells metabolism in animals by regulating mitochondrial copper homeostasis and cytochrome C oxidase assembly. Therefore, we speculated that *CHCHD7* gene might be involved in animal growth and body size. Herein, we discovered a 17-bp insertion/deletion (indel) within the goat *CHCHD7* gene. Then, we detected this variation in Shaanbei White Cashmere (SBWC, $n = 1055$) goats and Inner Mongolia White Cashmere (IMWC, $n = 743$) goats (Alathai type) using the mathematical expectation (ME) method. We then analyzed the correlation between these genotypes and goat body measurement traits. The results showed that the minor allelic frequency (MAF) was 0.011 in SBWC, and 0.048 in IMWC. In SBWC and IMWC, the reaction time by ME method was reduced by 36.78% and 27.59%, respectively, compared to the traditional method of screening samples one by one. Moreover, in SBWC goats, the 17-bp indel was significantly associated with body measurement traits (e.g. body height, and body length) in adults. In IMWC goats, the 17-bp indel was correlated with body measurement traits (e.g. body height) in weaners. In SBWC and IMWC goat populations, the body measurement traits of the individuals homozygous for 17-bp indel were higher than those in heterozygous individuals, except for the case of cannon circumference in IMWC weaners. These findings showed that the 17-bp insertion mutation within the goat *CHCHD7* gene significantly affected body morphometric traits, and could provide a basis for marker-assisted selection (MAS) breeding of cashmere goats.

Keywords Goat · *CHCHD7* · Insertion/deletion (indel) · Mathematical expectation (ME) · Body measurement traits

Abbreviations

BH Body height
 BL Body length
 bp Base pair

BW Body weight
 CC Cannon circumference
 CcO Cytochrome C oxidase
 CD Chest depth
CHCHD7 Coiled-coil-helix-coiled-coil-helix domain containing 7
 Cu Copper
 CW Chest width
COX23 Cytochrome C oxidase assembly homolog
ERR Estrogen-related receptor
ESRR Estrogen-related receptors
 HG Heart girth
 HHC Height at hip cross
 HW Hip width
 HWE Hardy–Weinberg equilibrium
 IMWC Inner Mongolia white cashmere
 InDel Insertion/deletion
 MAF Minor allele frequency
 MAS Marker-assisted selection
 ME Mathematical expectation

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-020-02434-8>) contains supplementary material, which is available to authorized users.

✉ Haixia Li
 lhx9529@163.com

✉ Xianyong Lan
 lanxianyong79@nwsuaf.edu.cn

¹ College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, People's Republic of China

² Shaanxi Provincial Engineering and Technology Research Center of Cashmere Goats, Yulin University, Yulin 719000, Shaanxi, People's Republic of China

³ Life Science Research Center, Yulin University, Yulin 719000, Shaanxi, People's Republic of China

NGn	The optimal number of individuals in one mixed group
NRn	Number of individuals in one reaction time
PCR	Polymerase chain reaction
PIC	Polymorphism information content
<i>PLAG1</i>	Pleiomorphic adenoma gene 1
pMAF	Predicting minor allelic frequencies
pRR	Predicting reduction rate
pRTn	Predicting reaction times by the formulate
<i>QTL</i>	Quantitative trait locus
RR	Reduction rate
RTn	Reaction times
SBWC	Shaanbei White Cashmere
<i>SF1</i>	Steroidogenic factor 1
<i>SF1F</i>	Vertebrate steroidogenic factor
<i>SORY</i>	SOX/SRY-sex/testis determining and HMG box factors
<i>SOX1</i>	SRY (sex determining region Y)-box 1
<i>SOX14</i>	SRY (sex determining region Y)-box 14
SPSS	Statistical product and service solutions
SNP	Single nucleotide polymorphisms
TD-PCR	Touch-down polymerase chain reaction

Introduction

Marker-assisted selection (MAS) breeding is a modern breeding method that depends on the functional molecular markers (genetic variations). Nowadays, MAS is often used in conjunction with traditional breeding methods to shorten the breeding cycle. The genetic variations often used in MAS include insertions/deletions (InDel), single nucleotide polymorphisms (SNP), and copy number variations (CNVs), etc. (Wang et al. 2019). Indel detection is easier and faster than SNP and CNV detections, so it is widely used in cashmere goat breeding (Cui et al. 2018; Kang et al. 2019a; Zhang et al. 2019a). As a very unstable quantitative traits, body measurement traits are regulated by many genes. Therefore, exploring gene functions is vital to understanding the mechanisms of growth (Erdenee et al. 2019; Wang et al. 2020b).

The goat Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 7 (*CHCHD7*) gene is located on chromosome 14 and contains six exons. *CHCHD7* gene is only 440 bp away from the Pleiomorphic Adenoma Gene 1 (*PLAG1*) gene, located in a growth-related major quantitative trait nucleotide (*QTN*) (Wu et al. 2019), and the two genes are transcribed in opposite directions. The major pleiotropic quantitative trait locus (*QTL*) in the *PLAG1-CHCHD7* intergenic region affects the growth and development traits of cattle, as observed by two functional variants in the bidirectional promoters of *PLAG1* and *CHCHD7* (Karim et al. 2011; Fink et al. 2017). *PLAG1* is a fusion

partner of *CHCHD7* and has an important effect on the reproduction and growth of mice (Karim et al. 2011; Matsuyama et al. 2011). Knocking out *PLAG1* gene in mice impairs spermatogenesis and sperm motility and results in growth retardation and dwarfism in mice (Juma et al. 2016). *PLAG1* can also could regulate the growth factor insulin like growth factor 2 (*IGF2*) in dog and human (Juma et al. 2016). Moreover, *CHCHD7*, also known as *COX23* cytochrome C oxidase assembly homolog (*COX23* or *MGC2217*), contains a conserved twin CX9C structural motif (Westerman et al. 2004; Banci et al. 2012). The products encoded by *CHCHD7* are essential conserved factors for the assembly of cytochrome C oxidase present in the mitochondrial membrane space (Barros et al. 2004; Cavallaro 2010; Banci et al. 2012). Cytochrome C oxidase, an enzyme at the end of the mitochondrial electron transport chain, is a typical copper-containing enzyme involved in bone cell metabolism, indicating that *CHCHD7* is indirectly involved in the growth and development of animals (Liu 2017; Murtha et al. 2019). Moreover, *CHCHD7* protein is potentially involved in copper (Cu) delivery during cytochrome C oxidase (CcO) biosynthesis, Cu homeostasis, or the Cox1 heme during heme transfer (Dela Cruz et al. 2016; Murtha et al. 2019).

The effects of *PLAG1* on animal reproduction and growth have been reported. A 19-bp indel mutation identified in the intron 1 of *PLAG1* gene is related to the growth and development of Chinese yellow cattle (Xu et al. 2018; Zhou et al. 2019). In the same position, two important indel mutations were identified in the sheep *PLAG1* gene, and they were all associated with the growth in these Chinese breeds (Wu et al. 2019). These results indicate that *PLAG1* gene is an important DNA marker for Chinese animals. Therefore, we speculated that some genetic variations in *CHCHD7*, a fusion partner of *PLAG1*, might also affect growth and development in goats. Genome-wide analysis revealed that *CHCHD7* gene polymorphisms affect growth traits, including body height (Lettre et al. 2008), body size and muscle formation (Jiao et al. 2014; Zhao et al. 2015), in cattle (Nishimura et al. 2012; Utsunomiya et al. 2013), pigs (Taye et al. 2018), and human (Randhawa et al. 2014). Moreover, *CHCHD7* gene is highly homologous in many species (Utsunomiya et al. 2013). However, no specific studies have revealed the relationship between body measurement traits of goats at different developmental stages and the *CHCHD7* gene indel polymorphism. Shaanbei White cashmere (SBWC) goats and Inner Mongolia White Cashmere (IMWC) goats are the two famous species raised locally in Shaanxi Province. SBWC goats are a cashmere and meat breed, who was bred from male Liaoning cashmere goats and female northern Shaanxi black goats (Zhang et al. 2019a). The IMWC goats (Alathai type) are a local fine breed of both cashmere and meat arising from long-term

selection and breeding of Mongolian goats (Xu and Zou 2017).

According to the polymorphism criteria of mutation sites (when polymorphism information content (PIC) < 0.25, it is low; when $0.25 < \text{PIC} < 0.5$, it is moderate; when $\text{PIC} > 0.5$, it is high) (Wang et al. 2020a), genetic variant loci can be classified as mutations of high, medium, low, and rare frequencies (Wang et al. 2020a). However, genetic variant loci with different frequencies have its most effective detection method. If there were genetic variants of goat *CHCHD7*, then the normal methods would be used, including directly sequencing (Zhang et al. 2019b), agarose gel electrophoresis with sequencing (Kang et al. 2019a), PCR-RFLP (Wang et al. 2019), and the Mathematical Expectation (ME) method (Yang et al. 2016). Then, which methods should be used in practical applications? First, we could randomly choose some (50–100) individuals to detect the frequency of the potential mutation. Then, in case where the mutation frequency was very low, the ME method would be the best choice (Li et al. 2017). Therefore, this study was designed to identify *CHCHD7* polymorphisms using the ME method (Yang et al. 2016) and explore the effects of identified polymorphisms on body measurement traits in SBWC and IMWC goats (Alathai type) at different developmental stages. The results presented in this study will lay the foundation for the fattening of cashmere goats by quick detecting relevant markers with low mutation frequency.

Materials and methods

Ethics statement

All experiments on all animals in this study were approved through the Institutional Animal Care and Use Committee of Northwest Agriculture and Forestry University (IACUC-NWAFU) in China. All experiments conformed to the guidelines of the Regional Animal Welfare Policy and Laws in Shaanxi province.

Animal samples and data collection

A total of 1798 female goat samples at different developmental stages (the lamb, weaner, and adult periods) were used in this study. The goats were of two breeds, Shaanbei white cashmere goats (SBWC, $n = 1055$) and Inner Mongolia white cashmere goats (IMWC, $n = 743$) (Alathai type). Weaner goats were 1 year old, and adult goats were between one and a half years old and four years old. The samples of these two breeds included in the study were reared randomly in the provinces of Shaanxi. All healthy individuals were allowed free access to feed and water in normal conditions. Body measurement traits of 1045 SBWC goats (367 weaners

and 678 adults) and 741 IMWC goats (507 weaners and 234 adults) were collected, including body weight (BW), body height (BH), body length (BL), heart girth (HG), chest depth (CD), chest width (CW), hip width (HW), cannon circumference (CC), and height at hip cross (HHC) (Wang et al. 2017).

Genomic DNA isolation and DNA pool construction

Genomic DNA samples were isolated from ear tissues (saved in 70% alcohol at $-80\text{ }^{\circ}\text{C}$) using the high-salt extraction method (Zhang et al. 2019a). The quality and purity of each DNA sample was assessed using a Nanodrop 1000 (Thermo Scientific, Waltham, MA, USA). Then, each sample was diluted with ddH₂O to a standard concentration of 50 ng/μl, respectively, and stored at $-20\text{ }^{\circ}\text{C}$ for subsequent genotyping. A genomic DNA pool was constructed by randomly selecting 36 DNA samples and mixing them in one tube. As PCR templates, the genomic DNA pools were used to amplify the DNA spanned by two primers and determine if an indel was present in the goat *CHCHD7* gene.

Primer design, PCR amplification, and indel genotyping using the ME method

Based on the goat *CHCHD7* gene reference sequences (GenBank accession no. NC_030821.1) and the indel loci information from the Ensembl database (<https://asia.ensembl.org/index.html>), indel loci primers were designed to produce amplicons equal to or greater than 6 bp using NCBI primer-blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (Table 1). Each PCR reaction mixture consisted of 12.5 μl 2×Eco Taq PCR Super mix (+ dye), 1.0 μl of each primer (forward and reverse primer), 1.0 μl genomic DNA, and 9.5 μl ddH₂O. A touch-down PCR cycling was run (Zhang et al. 2019a). Specific homozygous insertion and deletion individuals were detected by sequencing (Sangon Biotech, Shanghai, China). PCR products were detected by 3.5% agarose gel electrophoresis using a constant voltage of 120 V for 50 min.

The mutation frequency of the two indel loci was identified by randomly selecting and testing 48 individuals one by one. However, only one individual carried the ID genotype, and the other 47 individuals were of the II genotype. The estimated mutation frequency was approximately 0.01. Then, based on the sample sizes and indel loci mutation frequencies, the genotype of all individual samples was assessed by designing the most efficient pooling strategy using the mathematical expectation (ME) method (Yang et al. 2016; Li et al. 2017). The ME method is shown in Fig. 1. The formula of the ME method was (Yang et al. 2016) as follows:

Table 1 Used PCR primer sequences of goat *CHCHD7*

Locus	Primer sequence	T _m (°C)	Production sizes (bp)	Region
125 bp-17 bp	F: 5' GCAAACCTTC CCAGAGGGTGA3' R: 5' CTTGAGGC CACACACAATG3'	TD-PCR	125	Downstream gene variant
98 bp-6 bp	F: 5'AGGACAGGC CACAGAGTAGA3' R: 5'AGCACACAA AATAAGGGT GGAA3'	TD-PCR	98	Upstream gene variant

Fig. 1 The mathematical expectation (ME) method flow chart (Yang et al. 2016)

$$n = N \times (1/a \times (1 - p)^a + (1 + 1/a) \times (1 - (1 - p)^a)),$$

where n is the reaction times; N is the total sample; a is the sample size of a group, and p is the estimated mutation frequency (Yang et al. 2016).

CHCHD7 sequence analysis

Using the NCBI database reference sequences, nucleotide sequence similarities between cattle, sheep, and goat were calculated through counting the number and percent of “A”, “T”, “C”, “G”. Then, the homology of the reference sequences of cattle (NC_037341.1), goat (NC_030821.1), and sheep (NC_040260.1) registered in GenBank was analyzed by DNAMAN (version 6.0). In addition, the similarity of CHCHD7 amino acid sequences of cattle, sheep, and goat was calculated using NCBI-BLAST, and the ten amino acid sequences were aligned using the MEGA-X software.

Statistical analysis population

The SHEsis platform (<https://analysis.bio-x.cn>) (Shi et al. 2016) was used to analyze the genotypic frequencies, the allelic frequencies, and the Hardy–Weinberg equilibrium (HWE). The population parameters, including gene homozygosity (Ho), heterozygosity (He), effective allele numbers (Ne), and polymorphism information content (PIC), were analyzed using PopGene (version 1.3.1) (Chen et al. 2013). The distribution of genotypic and allelic frequencies in SBWC and IMWC goats were calculated using the chi-square test. The relationship between the genotypes and the different body measurement traits was confirmed by the unspecific linear model and the reduced linear model. Moreover, the model was as follows:

$$Y_{ijk} = u + A_i + B_j + E_{ijk},$$

where Y_{ijk} is the observation of the body measurement trait evaluated on the i th level of the fixed factor age (A_i) and

the j th level of the fixed factor genotype (B_j), μ is the overall mean for each trait, and E_{ijk} is the random error for the ijk th individual (Wang et al. 2017). Association analyses between *CHCHD7* gene and body measurement traits in the two goat breeds were assessed using the independent sample t test in statistical product and service solutions (*SPSS*) (version 25.0). Additionally, the goat *CHCHD7* pattern diagram of position and gene structure was generated using MapChart (version 2.32) and R (version 4.0.0). The results are expressed as mean $\pm$ standard error. Genomatix MatInspector (<https://www.genomatix.de>) online software was used to predict the transcription factor binding site on the mutations (Kang et al. 2019b).

Results

CHCHD7 gene sequence analysis

Using the sequences in the NCBI database, stacked bar graphs were generated to calculate the nucleotide sequence similarity between cattle, sheep, and goat *CHCHD7*. The three species had similar composition of “A”, “T”, “C”, and “G” (Fig. 2), and goat *CHCHD7* was 67.76% and 72.50% homologous to that of cattle and sheep, respectively (Table 5). Comparison of the amino acid sequences of cattle,

sheep, and goat *CHCHD7* revealed that the ten different transcripts contained a common motif (Fig. 3).

Indel genotyping using the ME method

Using the information in the NCBI and ensemble databases, two pairs of primers were designed to detect the insertion/deletion variant in the goat *CHCHD7* gene (Table 1). The results showed that the 6-bp indel locus (rs667044267) did not exist in SBWC goats and IMWC goats, but the 17-bp indel locus (rs668420586) was identified in the two goat breeds. Therefore, we focused on the 17-bp indel locus. At the 17-bp indel locus, samples with the II genotype produced one band (125 bp), those with the ID genotype produced three bands (the non-target product “A”, 125 bp, and 108 bp), and DD genotype samples produced one band (108 bp) after PCR and electrophoresis (Fig. 4). Two genotypes were observed in SBWC goats, and three in IMWC goats (Alathai type). Then, software based sequence analysis (Bioedit UK) revealed that the 17-bp indel sequence was “CTTGACAGACATCCT” (Fig. 4). Finally, comparing *CHCHD7* gene sequences on the Ensembl website with that on the NCBI website revealed the 17-bp indel 3756 bp downstream of the *CHCHD7* gene. This indel was described as NC 030821.1:g58695432-58695448delCTTGACAGA

Fig. 2 Comparison of the *CHCHD7* gene base content in different species. **a** The number of “A”, “T”, “C”, and “G” bases in different species; **b** the percentages of “A”, “T”, “C”, and “G” in different species

Fig. 3 The *CHCHD7* amino acid sequence similarity in cattle, sheep, and goat

Fig. 4 The location, electrophoresis pattern, and the sequence diagram of the 17-bp insertion mutation in the goat *CHCHD7* gene. The “M” represents Marker I. The “A” represents the non-target fragment, which is possibly heteroduplex

CACATCCT and was not the wild-type *CHCHD7* entry in the NCBI database.

According to the estimated mutation frequency and every breed sample size, the ideal number of individuals in one mixed group (NGn) was all 11 (SBWC and IMWC) (Fig. 5, Table 2). Then, the predicting reaction times (pRTn) were separately 207 (SBWC) and 146 (IMWC). However, in actual experiments, the actual reaction times (RT) were counted as 667 (SBWC) and 538 (IMWC), respectively. Compared with the traditional detection method (one by one), the ME method was 36.78% (SBWC) and 27.59% (IMWC) more efficient.

The genetic parameter analysis of the 17-bp goat *CHCHD7* gene

The genotypic and the allelic frequencies of the 17-bp indel within goat *CHCHD7* gene were summarized in Table 3. The results showed that the frequency of the ID genotype was the lowest in the SBWC goats and the frequency of the DD genotype was the lowest in the IMWC goats. The II genotype was the most frequent in both SBWC goats and IMWC goats. The deletion (D) allele frequencies were, respectively, 0.011 and 0.048 in the SBWC goats and IMWC goats, respectively. The polymorphism information content

(PIC) in the two breeds were separately 0.022 and 0.087. In addition, the gene homozygosity (Ho), Heterozygote (He), and the number of effective alleles (Ne) were separately calculated for both breeds. χ^2 test results showed that this locus was at Hardy–Weinberg equilibrium (HWE) in both breeds ($P > 0.05$).

The association between the 17-bp *CHCHD7* indel and body measurement traits in different goat breeds

The associations between the 17-bp indel with *CHCHD7* gene and body measurement traits were assessed in different developmental stages of two goat breeds. Significant relationship was observed between genotypes at the 17-bp indel locus and numerous body measurement traits in the 678 adult SBWC ewes and 507 weaner IMWC ewes. However, there was no correlation between this locus and weaner SBWC ewes ($n = 367$) and adult IMWC ewes ($n = 234$).

This 17-bp mutation locus was significantly associated with most body measurement traits in adult SBWC ewes. For example, body height (cm) ($P = 0.049$), body length (cm) ($P = 0.002$), heart girth (cm) ($P = 2.62 \times 10^{-4}$), hip width (cm) ($P = 0.004$), height at hip cross (cm) ($P = 2.00 \times 10^{-6}$), cannon circumference (cm) ($P = 0.004$),

Fig. 5 The reaction times at different sizes of one mixed group (different breeds according to their total samples). The least reaction time will be obtained when the sample size in one group was 11

Table 2 The reaction times of different groups in two goat breeds

Goat breeds	Shaanbei White Cashmere goats	Inner Mongolia White Cashmere goats
Sample sizes	1055	743
Estimated mutation frequency	0.01	0.01
NR ₁ (number of individuals in one reaction time)	1	1
RT ₁ (reaction times)	1055	743
NG _n (the optimal number of individuals in one mixed group)	11	11
pRT _n (predicting reaction times by the formulate)	207	146
pRR (predicting reduction rate)	80.38%	80.35%
RT _n (reaction times)	667	538
RR (reduction rate)	36.78%	27.59%

Table 3 Allelic frequency, genotypic frequency, and genetic diversity of the 17-bp indel of *CHCHD7* gene in different goat breeds

Breeds	Genotypic frequencies			Allelic frequencies		Population parameters				HWE
	II	ID	DD	I	D	Ho	He	Ne	PIC	P values
SBWC (n = 1055)	0.977 (n = 1031)	0.023 (n = 24)	–	0.989	0.011	0.978	0.023	1.023	0.022	P > 0.05
IMWC (n = 743)	0.907 (n = 674)	0.090 (n = 67)	0.003 (n = 2)	0.952	0.048	0.909	0.091	1.100	0.087	P > 0.05

HWE Hardy–Weinberg equilibrium, Ho homozygosity, He heterozygosity, Ne effective allele numbers, PIC polymorphism information content, SBWC Shaanbei White Cashmere goats, IMWC Inner Mongolia White Cashmere goats

and chest depth (cm) ($P = 1.73 \times 10^{-4}$) (Table 4). Individuals with the II genotype were significantly larger than those with the ID genotype (Table 4). Three traits were

significantly associated with the 17-bp mutation locus in IMWC weaner ewes (Table 4), including BH (cm) ($P = 0.030$), CC (cm) ($P = 1.95 \times 10^{-4}$), and HHC (cm)

Table 4 Association between the 17-bp indel of *CHCHD7* gene and body measurement traits in cashmere goats

Breeds	Traits (cm)	Observed genotypes (LSM ^a ± SE)		P values
Shaanbei White Cashmere goats (adults)		II (n = 659)	ID (n = 16)	
	Body height	54.60 ^a ± 0.16 (n = 657)	52.53 ^b ± 0.72 (n = 16)	0.049
	Body length	66.22 ^A ± 0.24 (n = 655)	63.49 ^B ± 0.71 (n = 16)	0.002
	Heart girth	80.37 ^A ± 0.38 (n = 657)	71.88 ^B ± 1.80 (n = 16)	2.62 × 10 ⁻⁴
	Hip width	14.52 ^A ± 0.09 (n = 658)	12.88 ^B ± 0.47 (n = 16)	0.004
	Height at hip cross	57.52 ^A ± 0.17 (n = 656)	54.49 ^B ± 0.42 (n = 16)	2.00 × 10 ⁻⁶
	Cannon circumference	8.02 ^A ± 0.03 (n = 659)	7.38 ^B ± 0.23 (n = 16)	0.004
Inner Mon- golia White Cashmere goats (weaner)		II (n = 449)	ID (n = 53)	
	Body height	51.27 ^a ± 0.18	50.09 ^b ± 0.49	0.030
	Height at hip cross	53.09 ^a ± 0.20	51.86 ^b ± 0.48	0.038
	Cannon circumference	8.33 ^B ± 0.03	8.65 ^A ± 0.09	1.95 × 10 ⁻⁴

Cells with different letters (a, b; A, B) differed significantly ($P < 0.05/P < 0.01$)

The “a” represents a significant difference between the same traits in different genotypes groups, the number with “a” is significantly larger than that with “b” ($P < 0.05$). The “A” represents a significance difference between the same traits in different genotypes groups, the number with “A” is significantly larger than that with B ($P < 0.01$)

LSM least squared means, SE standard error

Fig. 6 Structural map of bioinformatics predicted transcription factor binding sites on the 17-bp indel sequences. *ESRR* estrogen-related receptors, *SF1F* vertebrate steroidogenic factor, *SORY* SOX/SRY-sex/testis determining, *HMG* box factor

($P = 0.038$). In IMWC weaner ewes, the homozygous insertion (wild) genotype was the larger one in the BH and HHC traits (Table 4), and the heterozygous genotype had the largest CC (Table 4).

Predicted results of the transcription factor binding sites in the 17-bp indel locus

Altering in transcription factor-binding sites might affect transcription in the corresponding genes (Bharathavikru et al. 2017; Yang et al. 2019). We used the online software Genomatix MatInspector (<https://www.genomatix.de>) to predict the transcription factor-binding sites in the 17-bp indel locus (Kang et al. 2019a, b). The results showed that the *ESRR* (*ERR.01*), *SF1F* (*SF1.01*), and *SORY* (*SOX1.03*, *SOX14.01*) transcription factors could bind, at different sites, to the 17-bp insertion sequence (Fig. 6).

Table 5 Comparison of *CHCHD7* gene homology in different species

Species	Cattle	Goat	Sheep
Cattle	100%		
Goat	67.76%	100%	
Sheep	90.87%	72.50%	100%

Discussion

CHCHD7 (*COX23*) plays an important role in mitochondrial copper homeostasis and is involved in animal bone cell metabolism by participating in cytochrome oxidase assembly (Modjtahedi et al. 2016). In addition, the *CHCHD7* gene is highly homologous in many species (Figs. 2, 3, Table 5) (Utsunomiya et al. 2013). Genome-wide analysis showed that *CHCHD7* gene polymorphism is related to the growth traits, including body height

and size, in cattle, human, and pig (Utsunomiya et al. 2013; Jiao et al. 2014; Randhawa et al. 2014; Zhao et al. 2015; Taye et al. 2018). Moreover, SNPs of *CHCHD2*, *CHCHD3*, and *CHCHD10*, which belong to the same family as *CHCHD7*, caused dyskinesia and growth metabolism (Modjtahedi et al. 2016). For example, Parkinson disease, Huntington disease, human autosomal dominant mitochondrial myopathy, and the backfat thickness of pig (Fan et al. 2011; Feyeux et al. 2012; Ajroud-Driss et al. 2015; Funayama et al. 2015). Previous studies have concluded that *CHCHD7* has great potential in influencing animal growth traits. In breeding work, indel technology in MAS is widely used in goat breeding due to its simplicity and high accuracy. Hence, we aimed to confirm the relationship between the indel mutations in *CHCHD7* and the various body measurement traits in cashmere goats.

We found that the genotype of the 17-bp indel mutation within goat *CHCHD7* gene differed in different goat breeds. There were only two genotypes (II and ID) in SBWC, and three genotypes were found in IMWC (Alathai type). This may be because the frequency of the DD genotype was too low to be detected in SBWC goats (Cui et al. 2019). Additionally, it is also possible that the 17-bp indel locus was linked to other gene loci that affected growth traits, leading to the death of DD individuals at the embryonic stage in SBWC varieties. Through comparing the average value of every growth trait in the two breeds at the same stage, we found that SBWC goats had larger data than IMWC goats (Table S1, Table S2). Indirectly, these results showed that DD is not conducive to the growth of cashmere goats. Moreover, the frequency of ID genotypes in SBWC goats was much lower than that in IMWC goats. This might be due to different breeding and genetic backgrounds. SBWC goats are a cashmere and meat breed, bred from male Liaoning cashmere goats and female northern Shaanxi black goats (Zhang et al. 2019a). The IMWC goats (Alathai type) are a local fine breed that has been naturally and artificially selected for a long time, and its genetic performance is stable (Xu and Zou 2017). In addition to detecting the three genotypes of the *CHCHD7* gene, non-target products in the ID genotype were also found (Fig. 4). Previous studies had reported that non-target products were heteroduplexes (Ren et al. 2017) and were only present in individuals heterozygous for the indel gene (Nagamine et al. 1989). Therefore, our results were consistent with those of previous studies.

The traditional method of scanning numerous samples one by one would certainly require more time and effort than the method employed here. Compared with the traditional method, the mathematical expectation (ME) method, which is widely used in medical censuses for rare diseases, was not only saved time and money, but was also easy to operate (Zhang et al. 2003; Guo and Zhang 2006; Yang et al. 2016). The first time this efficient and feasible ME method was

used in our laboratory was to screen various bovine breeding samples (Yang et al. 2016). Based on our previous studies, the ME method was successfully used for a 20-bp indel of prion-related protein doppel (*PRND*) gene in sheep breeding (Li et al. 2017). In this study, we first randomly selected 48 individuals to test one by one. The deletion mutation was found to be present in low frequency. Genes with low-frequency indels are often important, so the ME strategy was used (Yang et al. 2016). Compared with the traditional method, the number of responses required for accuracy of the ME strategy was 667 times (SBWC, $n=1055$) and 538 times (IMWC, $n=743$), making the ME strategy simpler and more effective. Furthermore, our results revealed that the 17-bp indel of the goat *CHCHD7* gene was present in the HWE in SBWC and IMWC goats ($P>0.05$).

Of the studied goats, whether they were SBWC breeds or IMWC breeds or whether they were weaners or adults, the females with the II genotype had better morphometric traits. This might be because the 17-bp indel could affect the transcriptional efficiency and mRNA expression levels of other genes, as an enhancer in combination with different transcription factors, and affect the phenotype (Bharathavikru et al. 2017; Chen et al. 2019; Kang et al. 2019a). Therefore, we predicted a transcription factor binding site at the 17-bp indel site. Bioinformatics analysis revealed that this insertion site could be bound by the transcription factor estrogen-related receptors (*ESRR*), vertebrate steroidogenic factor (*SFIF*), SOX/SRY-sex/testis determining and HMG box factors (*SORY*) (Fig. 6). Moreover, the 17-bp indel was associated with body height in the two goat breeds, which is consistent with the role of the human *CHCHD7* gene (Lettre et al. 2008). Taken together, these results indicate that the mechanism of goat growth resembles that of human in respect to the *CHCHD7* gene.

These results showed that the individuals with the II genotype had the better morphometric traits. However, the largest cannon circumference (CC) was observed in IMWC weaners with the heterozygous ID genotype. The thicker CC could account for larger overall body weight, and the goat weaners have incomplete bone growth and are at a rapid development stage, so there are various and uncertain factors at play at this stage. Moreover, as far as *CHCHD7* itself is concerned, its products could be involved in Cu delivery, Cu homeostasis, and heme transfer processes in mitochondria (Barros et al. 2004; Cavallaro 2010; Banci et al. 2012). Appropriate amounts of copper have important biological significances for skeletal metabolism, growth, reproduction, and development of the central nervous system of ruminants (Tang 2006). Compared to other ruminants, goats have a higher sensitivity to copper, and lambs are the most sensitive to copper (Zhang et al. 2005). Correspondingly, *CHCHD7*-encoded soluble proteins are ubiquitously expressed in mammals, with the highest expression levels of expression in

adult tissues, particularly in the skeletal muscle and pancreas, and lower expression levels were lower in fetal tissues (Asp et al. 2006). These data indicate that *CHCHD7* might play a role at different developmental stages in goats, or that the different copper utilization rates in goats of different developmental stages result in different *CHCHD7* expression level (Zhang et al. 2005).

Conclusions

In summary, this experiment detected low-frequency mutations in *CHCHD7* gene in the large sample goats using the inexpensive and effective ME method. Moreover, the 17-bp indel examined significantly the numerous body measurement traits in the cashmere goats, and the homozygous insertion genotype (II) was the dominant genotype. This study might provide potentially useful DNA markers for the breeding of cashmere goats.

Acknowledgements This work was funded by the Provincial Key Projects of Shaanxi (2014KTDZ02-01). We were very grateful to the Qu Lei's team membership in Shaanxi Provincial Engineering and Technology Research Center of Cashmere Goats, Yulin University and Life Science Research Center, Yulin University, for the samples collection of SBWC goats. We sincerely thank Ph.D. ZWX, in the College of Animal Science and Technology, Northwest A&F University, for providing experimental instruments and equipment. We greatly thank the staffs of SBWC and IMWC breeding farm, Shaanxi province, P.R. China for their collecting samples.

Author contributions HL, XL, and LQ designed the study as well as wrote the paper. HL and XW performed the experiments. HL did the statistical analysis. LQ and HC collected the samples. XL and LQ edited and reviewed the paper.

Compliance with ethical standards

Conflict of interest We confirm that this manuscript has not been published in whole or in part and is not being considered for publication elsewhere. There are no any ethical conflicts of interest for any of the authors.

References

- Ajrroud-Driss S, Fecto F, Ajrroud K, Lalani I, Calvo SE, Mootha VK, Deng HX, Siddique N, Tahmouh AJ, Heiman-Patterson TD, Siddique T (2015) Mutation in the novel nuclear-encoded mitochondrial protein CHCHD10 in a family with autosomal dominant mitochondrial myopathy. *Neurogenetics* 16(1):1–9
- Asp J, Persson F, Kost-Alimova M, Stenman G (2006) CHCHD7-PLAG1 and TCEA1-PLAG1 gene fusions resulting from cryptic, intrachromosomal 8q rearrangements in pleomorphic salivary gland adenomas. *Gene Chromosomes Cancer* 45(9):820–828
- Banci L, Bertini I, Ciofi-Baffoni S, Jaiswal D, Neri S, Peruzzini R, Winkelmann J (2012) Structural characterization of CHCHD5 and CHCHD7: two atypical human twin CX9C proteins. *J Struct Biol* 180(1):190–200
- Barros MH, Johnson A, Tzagoloff A (2004) COX23, a homologue of COX17, is required for cytochrome oxidase assembly. *J Biol Chem* 279(30):31943–31947
- Bharathavikru R, Dudnakova T, Aitken S, Slight J, Artibani M, Hohenstein P, Tollervey D, Hastie N (2017) Transcription factor Wilms' tumor 1 regulates developmental RNAs through 3' UTR interaction. *Gene Dev* 31(4):347–352
- Cavallaro G (2010) Genome-wide analysis of eukaryotic twin CX9C proteins. *Mol Biosyst* 6(12):2459–2470
- Chen F, Shi J, Luo YQ, Sun SY, Pu M (2013) Genetic characterization of the gypsy moth from China (Lepidoptera, Lymantriidae) using inter simple sequence repeats markers. *PLoS ONE* 8(8):e73017
- Chen MY, Wang J, Liu N, Cui WB, Dong WZ, Xing BS, Pan CY (2019) Pig SOX9: expression profiles of sertoli cell (SCs) and a functional 18 bp indel affecting testis weight. *Theriogenology* 138:94–101
- Cui Y, Yan HL, Wang K, Xu H, Zhang XL, Zhu HJ, Liu JW, Qu L, Lan XY, Pan CY (2018) Insertion/deletion within the KDM6A gene is significantly associated with litter size in goat. *Front Genet* 9:91
- Cui W, Liu N, Zhang X, Zhang Y, Qu L, Yan H, Lan X, Dong W, Pan C (2019) A 20-bp insertion/deletion (indel) polymorphism within the CDC25A gene and its associations with growth traits in goat. *Arch Anim Breed* 62(1):353–360
- Dela Cruz R, Jeong MY, Winge DR (2016) Cox1 mutation abrogates need for Cox23 in cytochrome c oxidase biogenesis. *Microb Cell* 3(7):275–284
- Erdene S, Li J, Kang ZH, Xu HW, Zang RX, Cao X, Yang JT, Cai Y, Lan XY (2019) Sheep zinc finger proteins 395 (ZNF395): insertion/deletion variations, associations with growth traits, and mRNA expression. *Anim Biotechnol* 31(3):237–244
- Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF (2011) Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. *PLoS ONE* 6(2):e14726
- Feyoux M, Bourgois-Rocha F, Redfern A, Giles P, Lefort N, Aubert S, Bonnefond C, Bugi A, Ruiz M, Deglon N, Jones L, Peschanski M, Allen ND, Perrier AL (2012) Early transcriptional changes linked to naturally occurring Huntington's disease mutations in neural derivatives of human embryonic stem cells. *Hum Mol Genet* 21(17):3883–3895
- Fink T, Tiplady K, Lopdell T, Johnson T, Snell RG, Spelman RJ, Davis SR, Littlejohn MD (2017) Functional confirmation of PLAG1 as the candidate causative gene underlying major pleiotropic effects on body weight and milk characteristics. *Sci Rep* 7:44793
- Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, Li Y, Ogaki K, Ando M, Yoshino H, Tomiyama H, Nishioka K, Hasegawa K, Saiki H, Satake W, Mogushi K, Sasaki R, Kokubo Y, Kuzuhara S, Toda T, Mizuno Y, Uchiyama Y, Ohno K, Hattori N (2015) CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. *Lancet Neurol* 14(3):274–282
- Guo LJ, Zhang Y (2006) As application examples of the mathematical expectation. *Public Sci Technol* 7:169
- Jiao S, Maltecca C, Gray KA, Cassady JP (2014) Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genomewide association. *J Anim Sci* 92(7):2846–2860
- Juma AR, Damdimopoulou PE, Grommen SVH, Van de Ven WJM, De Groef B (2016) Emerging role of PLAG1 as a regulator of growth and reproduction. *J Endocrinol* 228(2):R45–R56
- Kang ZH, Jiang EH, Wang K, Pan CY, Chen H, Yan HL, Zhu HJ, Liu JW, Qu L, Lan XY (2019a) Goat membrane associated

- ring-CH-type finger 1 (MARCH1) mRNA expression and association with litter size. *Theriogenology* 128:8–16
- Kang ZH, Zhang SH, He LB, Zhu HJ, Wang Z, Yan HL, Huang YZ, Dang RH, Lei CZ, Chen H, Qu L, Lan XY, Pan CY (2019b) A 14-bp functional deletion within the CMTM2 gene is significantly associated with litter size in goat. *Theriogenology* 139:49–57
- Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B, Harris BL, Keehan MD, Littlejohn MD, Spelman RJ, Georges M, Coppieters W (2011) Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. *Nat Genet* 43(5):405–413
- Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, Illig T, Hackett R, Heid IM, Jacobs KB, Lyssenko V, Uda M, Boehnke M, Chanock SJ, Groop LC, Hu FB, Isomaa B, Kraft P, Peltonen L, Salomaa V, Schlessinger D, Hunter DJ, Hayes RB, Abecasis GR, Wichmann HE, Mohlke KL, Hirschhorn JN, Initiative DG, FUSION, KORA, C PLCO, Study NH, SardinIA (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* 40(5):584–591
- Li J, Zhu XC, Ma L, Xu HW, Cao X, Luo RY, Chen H, Sun XZ, Cai Y, Lan XY (2017) Detection of a new 20-bp insertion/deletion (indel) within sheep PRND gene using mathematical expectation (ME) method. *Prion* 11(2):143–150
- Liu B (2017) Comparative study of copper intestinal injury and transport in goats and sheep. Jilin Agricultural University
- Matsuyama A, Hisaoka M, Nagao Y, Hashimoto H (2011) Aberrant PLAG1 expression in pleomorphic adenomas of the salivary gland: a molecular genetic and immunohistochemical study. *Virchows Arch* 458(5):583–592
- Modjtahedi N, Tokatlidis K, Dessen P, Kroemer G (2016) Mitochondrial proteins containing coiled-coil-helix-coiled-coil-helix (CHCH) domains in health and disease. *Trends Biochem Sci* 41(3):245–260
- Murtha K, Hwang M, Peccarelli MC, Scott TD, Kebaara BW (2019) The nonsense-mediated mRNA decay (NMD) pathway differentially regulates COX17, COX19 and COX23 mRNAs. *Curr Genet* 65(2):507–521
- Nagamine CM, Chan K, Lau YF (1989) A PCR artifact: generation of heteroduplexes. *Am J Hum Genet* 45(2):337–339
- Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, Sugimoto Y, Takasuga A (2012) Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. *BMC Genet* 13:40
- Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW (2014) Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genet* 15:34
- Ren F, Yu S, Chen R, Lv XY, Pan CY (2017) Identification of a novel 12-bp insertion/deletion (indel) of iPS-related Oct4 gene and its association with reproductive traits in male piglets. *Anim Reprod Sci* 178:59–64
- Shi T, Peng WW, Yan JY, Cai HF, Lan XY, Lei CZ, Bai YY, Chen H (2016) A novel 17 bp indel in the SMAD3 gene alters transcription level, contributing to phenotypic traits in Chinese cattle. *Arch Anim Breed* 59(1):151–157
- Tang H (2006) The mechanism of haemocytolysis of goat erythrocytic membrane caused by high copper. Northeast Agricultural University
- Taye M, Yoon J, Dessie T, Cho S, Oh SJ, Lee HK, Kim H (2018) Deciphering signature of selection affecting beef quality traits in Angus cattle. *Genes Genom* 40(1):63–75
- Utsunomiya YT, do Carmo AS, Carvalheiro R, Neves HHR, Matos MC, Zavarez LB, O'Brien AMP, Solkner J, McEwan JC, Cole JB, Van Tassell CP, Schenkel FS, da Silva MVGB, Neto LRP, Sonstegard TS, Garcia JF (2013) Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. *BMC Genet* 14:52
- Wang XY, Yang Q, Wang K, Zhang SH, Pan CY, Chen H, Qu L, Yan HL, Lan XY (2017) A novel 12-bp indel polymorphism within the GDF9 gene is significantly associated with litter size and growth traits in goats. *Anim Genet* 48(6):735–736
- Wang XY, Yang Q, Wang K, Yan HL, Pan CY, Chen H, Liu JW, Zhu HJ, Qu L, Lan XY (2019) Two strongly linked single nucleotide polymorphisms (Q320P and V397I) in GDF9 gene are associated with litter size in cashmere goats. *Theriogenology* 125:115–121
- Wang K, Kang ZH, Jiang EH, Yan HL, Zhu HJ, Liu JW, Qu L, Lan XY, Pan CY (2020a) Genetic effects of DSCAML1 identified in genome-wide association study revealing strong associations with litter size and semen quality in goat (*Capra hircus*). *Theriogenology* 146:20–25
- Wang Z, Zhang X, Jiang E, Yan H, Zhu H, Chen H, Liu J, Qu L, Pan C, Lan X (2020b) InDels within caprine IGF2BP1 intron 2 and the 3'-untranslated regions are associated with goat growth traits. *Anim Genet* 51(1):117–121
- Westerman BA, Poutsma A, Steegers EAP, Oudejans CBM (2004) C2360, a nuclear protein expressed in human proliferative cytotrophoblasts, is a representative member of a novel protein family with a conserved coiled coil-helix-coiled coil-helix domain. *Genomics* 83(6):1094–1104
- Wu H, Pan Y, Zhang QF, Cao Y, Li J, Chen H, Cai Y, Sun XZ, Lan XY (2019) Insertion/deletion (InDel) variations in sheep PLAG1 gene locating in growth-related major QTL are associated with adult body weight and morphometric traits. *Small Rumin Res* 178:63–69
- Xu DX, Zou J (2017) The resource distribution of cashmere goats in Inner Mongolia autonomous region. *China Fiber Insp* 07:126–128
- Xu W, He H, Zheng L, Xu JW, Lei CZ, Zhang GM, Dang RH, Niu H, Qi XL, Chen H, Huang YZ (2018) Detection of 19-bp deletion within PLAG1 gene and its effect on growth traits in cattle. *Gene* 675:144–149
- Yang Q, Zhang S, Liu L, Cao X, Lei C, Qi X, Lin F, Qu W, Qi X, Liu J, Wang R, Chen H, Lan X (2016) Application of mathematical expectation (ME) strategy for detecting low frequency mutations: an example for evaluating 14-bp insertion/deletion (indel) within the bovine PRNP gene. *Prion* 10(5):409–419
- Yang Q, Zhang S, Li J, Wang X, Peng K, Lan X, Pan C (2019) Development of a touch-down multiplex PCR method for simultaneously rapidly detecting three novel insertion/deletions (indels) within one gene: an example for goat GHR gene. *Anim Biotechnol* 30(4):366–371
- Zhang YE, Liu GY, Ji AB, Sun JP (2003) The application of the mathematical expectation in the disease census. *J Math Med* 16(1):83–84
- Zhang W, Jia ZH, Lu DX (2005) Copper requirement and influencing factors of ruminants. *Grass Feed Livest* 04:50–55
- Zhang SH, Jiang EH, Wang K, Zhang Y, Yan HL, Qu L, Chen H, Lan XY, Pan CY (2019a) Two insertion/deletion variants within SPAG17 gene are associated with goat body measurement traits. *Animals (Basel)* 9(6):379

- Zhang SH, Peng K, Zhang GL, Cao Y, Zhang M, Chen H, Lei CZ, Lan XY, Zhao YM (2019b) Detection of bovine TMEM95 p.Cys161X mutation in 13 Chinese indigenous cattle breeds. *Animals (Basel)* 9(7):444
- Zhao FP, McParland S, Kearney F, Du LX, Berry DP (2015) Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet Select Evol* 47(1):49
- Zhou ZH, Huang BZ, Lai ZY, Li SP, Wu F, Qu KX, Jia YT, Hou JW, Liu JY, Lei CZ, Dang RH (2019) The distribution characteristics of a 19-bp indel of the PLAG1 gene in Chinese cattle. *Animals (Basel)* 9(12):1082