

## RESEARCH ARTICLE

# Genomic mapping identifies two genetic variants in the *MC1R* gene for coat colour variation in Chinese Tan sheep

Gebremedhin Gebreselassie<sup>1,2</sup> , Benmeng Liang<sup>1</sup> , Haile Berihulay<sup>1</sup>, Rabul Islam<sup>1</sup>, Adam Abied<sup>1</sup>, Lin Jiang<sup>1</sup>, Zhengwei Zhao<sup>3\*</sup>, Yuehui Ma<sup>1\*</sup>

**1** Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China, **2** Department of Agricultural Biotechnology, Biotechnology Center, Ethiopian Biotechnology Institute, Ministry of Innovation and Technology, Addis Ababa, Ethiopia, **3** Institute of animal science, Ningxia Academy of Agriculture and Forestry Sciences, Ningxia, Yinchuan, China

 These authors contributed equally to this work.

\* [mayuehui@caas.cn](mailto:mayuehui@caas.cn) (YM); [zhaoweibing2008@163.com](mailto:zhaoweibing2008@163.com) (ZZ)



## OPEN ACCESS

**Citation:** Gebreselassie G, Liang B, Berihulay H, Islam R, Abied A, Jiang L, et al. (2020) Genomic mapping identifies two genetic variants in the *MC1R* gene for coat colour variation in Chinese Tan sheep. PLoS ONE 15(8): e0235426. <https://doi.org/10.1371/journal.pone.0235426>

**Editor:** Raluca Mateescu, University of Florida, UNITED STATES

**Received:** January 12, 2020

**Accepted:** June 15, 2020

**Published:** August 20, 2020

**Copyright:** © 2020 Gebreselassie et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** The authors would like to thank the National Natural Science Foundation of China (31601910, U1603232) and the earmarked funds for the Modern Agro-Industrial Technology Research System (CARS-40-01) for their financial support. Z.Z. is supported by Ningxia Natural Science Foundation project with grant number (2019AAC03286).

## Abstract

Coat colour is one of the most important economic traits of sheep and is mainly used for breed identification and characterization. This trait is determined by the biochemical function, availability and distribution of pheomelanin and eumelanin pigments. In our study, we conducted a genome-wide association study to identify candidate genes and genetic variants associated with coat colour in 75 Chinese Tan sheep using the ovine 600K *SNP* BeadChip. Accordingly, we identified two significant SNPs (*rs409651063* at 14.232 Mb and *rs408511664* at 14.228 Mb) associated with coat colour in the *MC1R* gene on chromosome 14 with  $-\log_{10}(P) = 2.47E-14$  and  $1.00E-13$ , respectively. The consequence of *rs409651063* was a missense variant (g.14231948 G>A) that caused an amino acid change (Asp105Asn); however, the second SNP (*rs408511664*) was a synonymous substitution and is an upstream variant (g.14228343G>A). Moreover, our PCR analysis revealed that the genotype of white sheep was exclusively homozygous (GG), whereas the genotypes of black-head sheep were mainly heterozygous (GA). Interestingly, allele-specific expression analysis (using the missense variant for the skin cDNA samples from black-head sheep) revealed that only the G allele was expressed in the skin covered with white hair, while both the G and A alleles were expressed in the skin covered with black hair. This finding indicated that the missense mutation that we identified is probably responsible for white coat colour in Tan sheep. Furthermore, qPCR analysis of *MC1R* mRNA level in the skin samples was significantly higher in black-head than white sheep and very significantly higher in GA than GG individuals. Taken together, these results help to elucidate the genetic mechanism underlying coat colour variation in Chinese indigenous sheep.

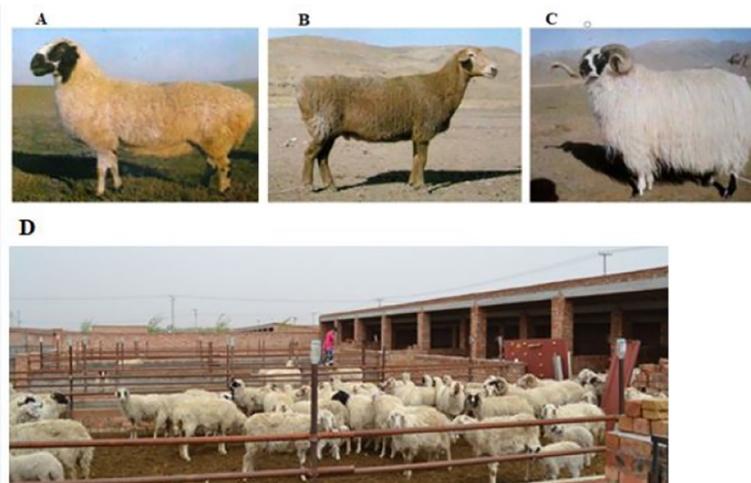
**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Animal coat colour is one of the most important economic traits, particularly for animal breeds that are kept for skin and wool production. Coat colour is a highly useful trait that affects the behaviour of animals and is an essential trait for survival (to cope with the environment and biological factors) in different environmental conditions [1–3]. Moreover, animal coat colour has played a crucial role in breed identification, characterization [4] and other applications, such as fibre production in goat, rabbit and sheep breeds [5]. In fact domesticated species are also characterized by a huge allelic variability of coat-colour-associated genes which leads to negative pleiotropic effects linked with coat-colour variants [6]. Animal coat colour variation is mainly regulated by genetic and environmental factors [6]. Classical genetics assumes that the presence of eumelanin or phaeomelanin is genetically controlled by the extension and agouti locus [7]. In mammals, coat colour variation is determined by the biochemical function, distribution and availability of phaeomelanin and eumelanin pigments [8]. Phaeomelanin produces a red and yellow colour, whereas eumelanin produces black and brown phenotypes [8]. The relative amount of eumelanin and phaeomelanin colours in melanocytes is regulated by the interactions of agouti signalling protein (*ASIP*) and melanocortin 1 receptor (*MC1R*) genes [9, 10]. In addition to *MC1R* and *ASIP*, there are also other genes involved in coat colour variation such as melanocyte inducing transcription factor (*MITF*), tyrosinase related protein-1 (*TYRP1*) and v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homologue (*KIT*) [6, 11–15].

*ASIP* is a small paracrine signalling molecule that interacts with the extension locus encoded by the agouti locus [16]. *ASIP* has an antagonist function with *MC1R* in the pigmentation process. *ASIP* blocks the  $\alpha$ -MSH receptor interaction, which causes pigment-type switching from eumelanin to phaeomelanin pigment [17, 18]. Thus, *ASIP* inhibits *MC1R* signalling and eumelanogenesis, and the antagonist function of *ASIP* promotes white and red colour against the black and brown phenotype [19–21]. The effect of *ASIP* on coat colour is determined by dominant or recessive agouti alleles. The dominant agouti alleles are responsible for phaeomelanin phenotypes and the recessive alleles for black coat colour [22]. Molecular variants have been reported in the *ASIP* gene associated with the coat colour phenotype in different sheep breeds, such as Finns sheep [23], Massese sheep [24], Soay sheep [25] and Australia merino sheep [4]. In contrast, the *MC1R* gene, also known as the  $\alpha$ -melanocyte stimulating hormone receptor ( $\alpha$ -MSHR), a seven-transmembrane G-protein coupled receptor, is encoded by an extension (E) locus [7]. The *MC1R* gene is reported as a potential candidate gene that plays a significant role in melanogenesis and wool pigmentation and is responsible for black coat colour in mammals [8, 26, 27]. For instance, several functional and non-functional molecular variants have been reported in the *MC1R* gene associated with coat colour phenotype in different sheep breeds such as Zandi, Baluchi and Zel sheep [28], Brazilian Crioula sheep [29], Manchega and Rasa Aragonesa sheep [30], Chinese sheep [15], Brazilian Creole sheep [8], Massese sheep [24], Xalda sheep [31], Norwegian Dala sheep [32].

Chinese indigenous sheep breeds are classified in 42 based on their geographical distribution and morphological characteristics. These sheep breeds are again categorized in three groups: Mongolian fat-tailed, Kazakh fat-rumped and Tibetan thin-tailed sheep (Fig 1A, 1B & 1C) [33]. Chinese Tan sheep are unique and typical sheep breeds that are employed for both fur and meat production and are categorized under the Mongolian fat-tailed sheep group. Tan sheep are widely distributed in northwestern China and famous for lamb pelt and lustrous white curly fleece production. These sheep are also renowned breeds for long-term adaptation to dry, cold and windy environments [34]. The coat colour of Chinese Tan sheep is solid white, and white with black and brown colours around the head, neck and face (Fig 1D). This



**Fig 1. The typical image of the three categories of indigenous Chinese sheep breeds and Chinese Tan sheep.** (A) Mongolian-fat tailed sheep group, (B) Kazakh-fat rumped sheep group, (C) Tibetan-thin-tailed sheep group and (D) shows solid white, white with black head and brown face phenotype of Tan sheep gathered in Ningxia conservation farm, Ningxia province, China.

<https://doi.org/10.1371/journal.pone.0235426.g001>

coat colour variation among the Chinese Tan sheep was the main focus of our study. A genome-wide association study (GWAS) is a powerful and preferred approach to detect causal variants and define narrower genomic regions than Quantitative trait loci (QTLs). It is a hypothesis-free method to determine the associations between genetic regions and phenotypic traits in livestock and humans [35–37]. In sheep, GWAS was applied in Soay sheep for the first time to investigate horn types [35] and was subsequently applied to Corriedale sheep to examine the inheritance of rickets [36]. There are also GWAS reports for other traits in the Chinese sheep breed such as coat colour [13].

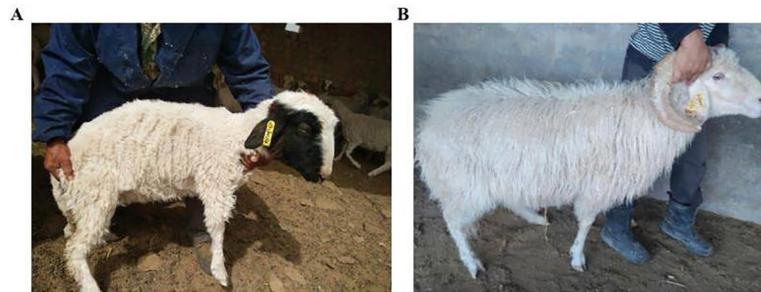
## Materials and methods

### Ethics approval

All the procedures involved in handling and collecting the samples from sheep were approved by the ethics committees of the Ministry of Agriculture of the People’s Republic of China (IAS-CAAS-AE-03).

### Animals and samples

For GWAS, a total of 10 ml of blood from the jugular vein into a tube and phenotype data were collected from 75 Chinese Tan sheep from the Ningxia sheep conservation farm, Ningxia province, China. Among the sampled sheep, 29 were coded as cases (black head/face coat colour), whereas 46 were coded as controls (white coat colour sheep) (Fig 2). For gene expression analysis, tissue samples ( $n = 14$ ) were collected from the skin of solid white sheep and black-head sheep. Similarly, for allele-specific expression (ASE) analysis, tissue samples ( $n = 14$ ) were collected from the skin under the black hair and under the white hair part of the black-head sheep (from the same sheep). The tissues were collected via skin punch biopsy under anaesthesia and placed in liquid nitrogen. Moreover, blood samples ( $n = 10$ ) and tissue samples ( $n = 14$ ) were collected from white and black-head Tan sheep to validate the gene.



**Fig 2. The coat colour phenotype of Chinese Tan sheep.** (A) Shows white with black coat colour phenotype, whereas (B) referred to solid white coat colour phenotype.

<https://doi.org/10.1371/journal.pone.0235426.g002>

### DNA and RNA extraction

Genomic DNA was extracted from 3-ml blood samples using Promega Wizard Genomic Purification according to the manufacturer's protocol Kits, whereas RNA was extracted from skin tissue samples using the RNeasy Pure Kit (For Tissue). After DNA extraction, genotyping was performed using Illumina's ovine 600K SNP HD BeadChip for GWAS work. For SNP identification, we performed polymerase chain reaction (PCR) using the reagent Phanta Max Super-Fidelity DNA Polymerase (Cat: P505-d1, Vazyme). Complementary DNA (cDNA) was also prepared from the extracted RNA samples using a PrimeScript™ RT reagent kit for quantitative polymerase chain reaction (qPCR) and nested PCR.

### Genome wide association study (GWAS)

A case-control model was applied to manipulate the genotype data (S1 File and S2 File) and phenotype data (S3 File). Both genotype data quality control and phenotype-genotype association data generation were performed using PLINK version 1.9 software [38]. Manhattan plots and QQ plots were generated using the qqman package in R v.3.3.2. Genomic regions and candidate genes were identified on the livestock genome browser (UCSC) using the sheep assembly Aug. 2012 (ISGC Oar\_v3.1/oviAri3). Finally, gene and SNP annotation was performed using Ensemble genome browser 95 and NCBI.

### Primer designing

Primers for qPCR, PCR and nested PCR were designed by Primer3Plus using the *MC1R* mRNA sequence (GenBank accession no. NM\_001282528.1) and *GAPDH* mRNA sequences (GenBank accession no. NM\_001190390.1). Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as a housekeeping gene for gene expression analysis. The primers were synthesized by Beijing Tianyihuiyuan Biotechnology Co. Ltd., Beijing-China (Table 1).

### Polymerase chain reaction (PCR)

PCR was performed in a total volume of 26  $\mu$ L containing 2xPhanta Max Buffer (12.5  $\mu$ L), dNTP Mix (0.5  $\mu$ L), Phanta Max Super-Fidelity DNA Polymerase (0.5  $\mu$ L), ddH<sub>2</sub>O (8.5  $\mu$ L), forward primer (1  $\mu$ L), reverse primer (1  $\mu$ L) and 2  $\mu$ L of DNA. The PCR analysis was performed on an Applied Biosystems® Gene Amp® PCR System 9700 and programmed as follows: at 95°C initial denaturation for 5 min, 35 amplification cycles of 30 s denaturing at 95°C, 30 s annealing at 54°C and extension at 72°C for 2 min. The final extension was run at 72°C for 5 min. Agarose gel electrophoresis was carried out to determine whether the PCR products properly amplified the target region using 1.5% agarose, 50 ml 1X TAE buffer and 5  $\mu$ L God

**Table 1. Primer information for qPCR, nested PCR and PCR applications.**

| Primers name           | Primer sequence                                 | Product size | Applications         | Locus/SNPs                |
|------------------------|-------------------------------------------------|--------------|----------------------|---------------------------|
| <i>MC1R-F MC1R-R</i>   | CTCTCCATCACCTACTACAACC<br>CAGCATGTGGACATAGAGGAC | 102 bp       | For qPCR             | 14:14231948 (rs409651063) |
| <i>GAPDH-F GAPDH-R</i> | GTCCGTTGTGGATCTGACCT<br>GGAGACAACCTGGTCCTCAG    | 130 bp       |                      |                           |
| <i>MC1R-F1 MC1R-R1</i> | GCTGGTGAGTCTTGTGGAGA<br>GCCAAAGCCCTGATGAATGG    | 562 bp       | For Nested PCR & PCR | 14:14231948 (rs409651063) |
| <i>MC1R-F2 MC1R-R2</i> | GTGAGCGTCAGCAACGTG<br>ACATAGAGGACGGCCATCAG      | 366 bp       |                      |                           |
| <i>MC1R-F3 MC1R-R3</i> | CCACCTGCCTCTGCTCTTCTA<br>GGAGGGTGCCTCAGTAGACAA  | 466 bp       | For PCR              | 14:14228343 (rs408511664) |

Abbreviations: F, forward primer; R, reverse primer.

<https://doi.org/10.1371/journal.pone.0235426.t001>

view. The gel was checked using a molecular imager (Gel Doc XR, BIO-RAD). Care was taken to avoid DNA damage due to extended exposure to UV light.

### cDNA preparation

The Complementary DNA (cDNA) was synthesized from 1 µg of DNase-treated RNA using a primeScript™ RT reagent kit. First, we prepared a total volume of 10 µL of DNA elimination reaction containing 5× gDNA Eraser Buffer (2 µL), gDNA Eraser (1 µL), RNA (1 µL) and RNase Free ddH<sub>2</sub>O (6 µL). Next, 10 µL of master mix was composed of 5X PrimeScript Buffer 2 (4 µL), PrimerScript RT Enzyme Mix I (1 µL), RT primer Mix (1 µL) and RNase Free ddH<sub>2</sub>O (4 µL). Then, the DNA elimination reaction and master mix were mixed to prepare 20 µL of reverse transcription reaction. Finally, the cDNA solution was heated on a Veriti 96-well thermal cycler at 37°C for 15 min and at 85°C for 5 s. This solution was diluted with ddH<sub>2</sub>O and stored at 4°C for PCR, nested PCR and qPCR use.

qPCR was performed in a total volume of 20 µL containing 10 µL of TB Green Premix Ex Taq, 0.4 µL of forward primer, 0.4 µL of reverse primer, 0.4 µL of ROX Reference Dye II, 2 µL of diluted cDNA and 6.8 µL of ddH<sub>2</sub>O. The cycling parameters were 95°C for 30 s followed by 40 cycles (95°C for 5 s, 60°C for 34 s and 95°C for 15 s) and a cycle (60°C for 1 min and 95°C for 15 s). Nested PCR was also prepared using two sets of primers (562 bp and 366 bp) (Table 1). The first PCR was prepared using the first set of primers in a total volume of 20 µL composed of 1 µL of forward primer, 1 µL of reverse primer, 10 µL of Master mix, 6 µL of ddH<sub>2</sub>O and 2 µL of cDNA solution. The cycling parameters were 95°C for 5 min, 34 cycles (95°C for 30 s, 60°C for 30 s and 72°C for 1 min) and 72°C for 2 min and held at 4°C. The second reaction was performed using the second set of primers and 1 µL of PCR product from the first PCR (instead of cDNA template). All other reaction components and cycling parameters remained the same as those of the first reaction. Agarose gel electrophoresis was carried out using 1.5% agarose, 50 ml 1X TAE buffer and 5 µL God view to amplify PCR products.

### DNA genotyping and gene expression analysis

The gene (*MC1R*) and the genetic variants (rs409651063 and rs408511664) identified to be associated with the coat colour phenotype were confirmed using PCR approaches. For this, PCR products (from DNA and cDNA) were sequenced by Beijing Liuhe BGI Technology Co., Ltd. The DNA and cDNA sequence results were assembled, aligned and edited using MEGA 7.0.14 (7160202-x86\_64) to identify the genetic variants and for allele-specific expression analysis. SNP positions were cross-checked and verified by Chromas version 2.6.4 software.

**Table 2. GWAS results: genetic variants and their location, P-value, amino acid change and consequences.**

| SNPs        | Chr:bp      | Alleles | P-value  | Consequence Type      | AA  | Gene        |
|-------------|-------------|---------|----------|-----------------------|-----|-------------|
| rs408511664 | 14:14228343 | G/A     | 2.47E-14 | upstream gene variant |     | <i>MC1R</i> |
| rs409651063 | 14:14231948 | G/A     | 1.00E-13 | missense variant      | D/N |             |
| rs160909987 | 14:14227752 | C/T     | 6.62E-10 | upstream variant      |     | <i>MC1R</i> |
| rs409417796 | 14:14213367 | G/A     | 3.74E-09 | intron variant        |     | <i>MC1R</i> |
| rs412176113 | 14:14216978 | C/T     | 2.96E-07 | intron variant        |     | <i>MC1R</i> |

Abbreviations: SNP, single nucleotide polymorphism; Chr, chromosome; BP, base pair position; AA, amino acid change.

<https://doi.org/10.1371/journal.pone.0235426.t002>

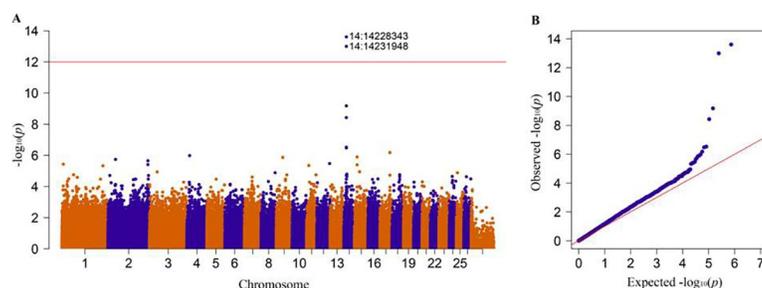
Moreover, qPCR data were analysed using GraphPad Prism v6.04 to determine whether the *MC1R* gene is differentially expressed in white and black-head Chinese Tan sheep.

## Results and discussion

### Results

GWAS: after quality control was performed using missing genotype data (0.1), Hardy-Weinberg equilibrium ( $1e^{-5}$ ) and minor allele frequency (0.05), a total of 366,991 SNPs remained for the genome-wide association study. Then, we perform the GWAS using phenotype data and found a candidate genomic region spanning 14.228Mb-14.232Mb on chromosome 14. The genomic location (SNPs) was then ranked using their p-value (smallest to largest) (Table 2). The Manhattan plot result indicated the significant SNPs as top signals for the black-head phenotype at chromosome 14 (Fig 3A). Moreover, the quantile-quantile plot (QQ plot) was also examined to determine the validity of the P-value of the significant SNPs. The corresponding QQ plot showed a reasonable distribution of the SNPs on GWAS. Most of the SNPs followed the normal curve distribution; however, only a few significant SNPs deviated from the normal curve for the black-head coat colour phenotype (Fig 3B).

We were also identified two significant SNPs (*rs409651063* at 14.232 Mb position and *rs408511664* at 14.228 Mb position) on chromosome 14, with  $-\log_{10}(P) = 2.47E-14$  and  $1.00E-13$ , respectively. These SNPs were identified on coding (CDS) and upstream regions of the *MC1R* gene. The consequence of the variant *rs409651063* was a missense variant (g.14231948 G>A) causing an amino acid change (Asp105Asn); however, the second SNP (*rs408511664*) was an upstream variant (g.14228343G>A) (Table 2). The SIFT prediction value for the *rs409651063* variant was smaller than 0.05 (SIFT score = 0.01). The effect of this



**Fig 3. GWAS results plotted using Manhattan and quantile-quantile plots.** (A) Manhattan plot for coat colour analysis. The horizontal red line denotes genome-wide significance ( $1e^{-12}$ ). The y-axis shows the  $-\log_{10}(p)$  value, and the x-axis shows the chromosome number. (B) QQ plot: the  $-\log_{10}(P)$ -value shows the normality of the observed (y-axis) and expected (x-axis) values. The red line shows a normal curve, and blue indicates colour variation among individuals. The two blue dots on the top denote significant SNPs for the coat colour phenotype.

<https://doi.org/10.1371/journal.pone.0235426.g003>

genetic variant on *MC1R* protein function is deleterious. This finding suggested that *MC1R* signalling may be disrupted by the mutation.

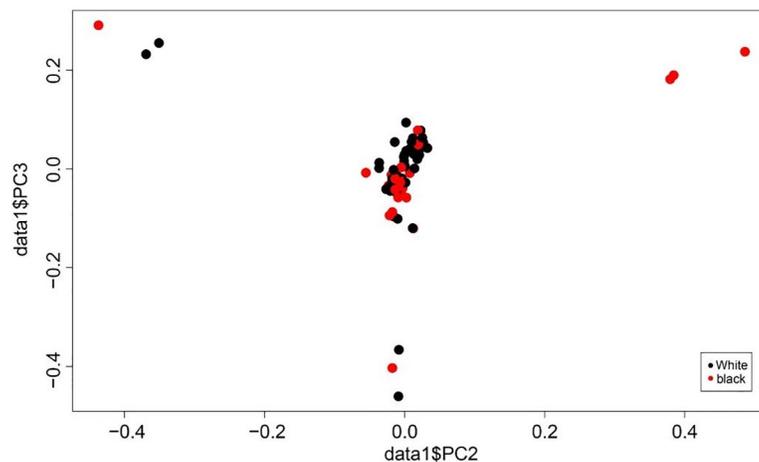
### Population genetic structure analyses

Population structure was performed to characterize the genetic relationship between the animals with two types of coat colour phenotypes (white Tan sheep and black-head Tan sheep) using principal component analysis (PCA). Accordingly, we observed a closer clustering pattern between the phenotypes implied that the animals are closely related (Fig 4).

### Validation of the GWAS results

After sequencing the DNA by Sanger sequencing, DNA genotyping revealed that all the individuals from the white phenotype were homozygous (GG), whereas all individuals from the black-head phenotype were heterozygous (GA) (DNA genotype, Table 3). Moreover, we performed allele-specific expression (ASE) analysis for the missense variant (*rs409651063*) to investigate the effects of the mutation in the skin samples covered with either white hair or black hair. The combined results of the DNA genotype and cDNA genotype of the missense variant (*rs409651063*) revealed that only allele G was expressed in the skin covered with white hair, while both G and A alleles were expressed in the skin covered with black hair (Table 3). This preference of the mutant G allele in the skin covered with white hair indicated that the resulting mutant *MC1R* protein may lead to dysfunction of black pigmentation in white skin.

Furthermore, we performed qPCR for the missense variant (*rs409651063*) to evaluate the expression level of the *MC1R* gene in white and black-head phenotypes. Accordingly, *MC1R* mRNA was differentially expressed in the white and black-head phenotypes in Tan sheep. The expression level was significantly higher in sheep with black-head coat colour carrying the A allele than in white sheep (Fig 5A). The *MC1R* mRNA expression level in the skin samples was also significantly higher in the individuals carrying GA alleles than in the individuals carrying GG alleles (Fig 5B). The gene expression cycle threshold (Ct) value of *MC1R* and *GADPH* (reference) genes including replicates of each sample is available in the supplementary file (S1 Table).



**Fig 4. Principal component analyses of white and black-head Tan sheep.**

<https://doi.org/10.1371/journal.pone.0235426.g004>

**Table 3. DNA genotyping and allele-specific expression (ASE) analysis for two candidate loci: chr14:14228343 (rs408511664) and chr14:14231948 (rs409651063).**

| blood sample | coat colour phenotype | DNA genotype |          | Skin sample | hair colour of skin | cDNA genotype (14231948) |
|--------------|-----------------------|--------------|----------|-------------|---------------------|--------------------------|
|              |                       | 14228343     | 14231948 |             |                     |                          |
| B1           | black head            | GA           | GA       | B1-1        | black               | GA                       |
|              |                       |              |          | B1-2        | white               | GG                       |
| B2           | black head            | GA           | GA       | B2-1        | black               | GA                       |
|              |                       |              |          | B2-2        | white               | GA                       |
| B3           | black head            | GA           | GA       | B3-1        | black               | GA                       |
|              |                       |              |          | B3-2        | white               | GG                       |
| B4           | black head            | GA           | GA       | B4-1        | black               | GA                       |
|              |                       |              |          | B4-2        | white               | GA                       |
| W1           | white                 | GG           | GG       | W1          | white               | GG                       |
| W2           | white                 | GG           | GG       | W2          | white               | GG                       |
| W3           | white                 | GG           | GG       | W3          | white               | GG                       |
| W4           | white                 | GG           | GG       | W4          | white               | GG                       |
| W5           | white                 | GG           | GG       | W5          | white               | GG                       |
| W6           | white                 | GG           | GG       | W6          | white               | GG                       |

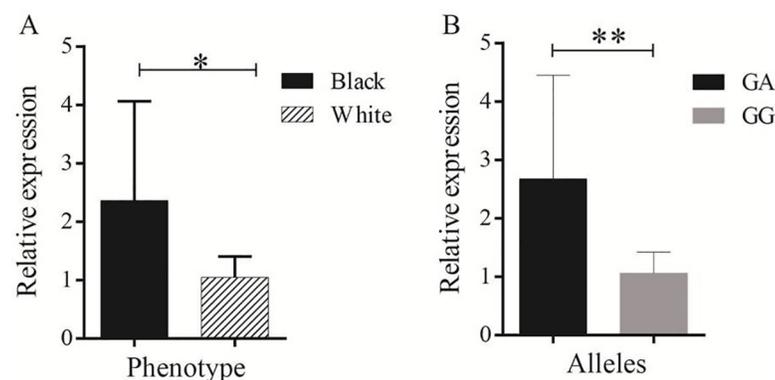
Abbreviations: B1, black head sheep sample 1; B1-1, black colour skin from sample 1; B1-2, white colour skin from sample 1; B2, black head sheep sample 2; B2-1, black colour skin from sample 2; B2-2, white colour skin from sample 2, the same way of coding to other samples also.

<https://doi.org/10.1371/journal.pone.0235426.t003>

## Discussion

Chinese Tan sheep (widely distributed in northwestern China) are unique and typical sheep breed that are reared for both fur and meat production. These sheep are renowned for long-term adaptation to dry, cold and windy environments. The coat colour of this breed is characterized by solid white and white colours with black and brown colours around the head, neck and face [33]. Thus, it would be important to identify the genomic regions and genetic variants associated with coat colour traits in Tan sheep.

In the present GWAS study, we identified a candidate gene (*MC1R*) and two significant genetic variants (*rs409651063* and *rs408511664*) associated with coat colour in Tan sheep. This gene and the molecular variants were detected in the genomic region spanning 14.231 Mb–14.228 Mb on chromosome 14 (Table 2 and Fig 3A and 3B). According to Kijas et al. [30], a



**Fig 5. Gene and allele expression level for white and black coat colours phenotype of Chinese Tan sheep.** (A) Comparison of gene expression (*MC1R*) in black-head and white colour sheep. (B) Expression level comparison between the GG genotype (white coat colour sheep) and GA genotype (black-head coat colour sheep) in the *MC1R* gene for the *rs409651063* variant. The sign \* is denoted for  $P < 0.05$  and \*\* is for  $P < 0.01$ .

<https://doi.org/10.1371/journal.pone.0235426.g005>

GWAS (SNP50 BeadChip) identified two SNPs (s26449 and s46705) closest to the *MC1R* gene on chromosome 14 in solid black and white Manchega and Rasa Aragonesa sheep associated with the coat colour phenotype. The variants were located 14.127 Mb upstream and 14.203 Mb downstream of the *MC1R* gene. However, the SNP density applied and the phenotype of sampled animals were not similar to those in our study. In a previous PCR study, three synonymous and two nonsynonymous variants of the *MC1R* gene were reported in completely black Chinese sheep (Minxian Black-fur breed). These variants of the *MC1R* gene were found to be associated with the black coat colour phenotype in the Minxian black-fur breed, although no complete association was reported in the black-face coat colour phenotype of Chinese sheep breeds (Tan, Tibetan, Duoland and Mongolian sheep) [15]. Moreover, Yang et al. [15] reported three haplotypes in the *MC1R* gene related to the black coat colour phenotype in Minxian black-fur sheep. In fact, in the present study, we identified a missense variant (g.14231948 G>A) causing p.Asp105Asn amino acid change and an upstream variant (g.14228343 G>A) (Table 2) in Chinese Tan sheep. This finding indicated that pigmentation variation in Chinese Tan sheep might be due to these genetic variants occurring on the CDS and upstream region of the *MC1R* gene. Fontanesi et al. [39] also reported missense and nonsense mutations in the *MC1R* gene in goat breeds. Moreover, several molecular studies, including GWAS, have been conducted to investigate the genetic variation in coat colour phenotype in sheep breeds. In such studies, the *MC1R* gene is labelled to be responsible for black colour phenotypes in sheep [5, 30, 40]. Muniz et al. [40] reported the seven most significant SNPs on chromosome 14 close to the *MC1R* gene in Morada Nova sheep in his GWAS study to investigate genome regions responsible for the coat colour phenotype. Furthermore, several SNPs have been observed in the *MC1R* gene in different animal species associated with the coat colour phenotype. For instance, in a recent study, two silent mutations in the 5' end of the coding region of the *MC1R* gene were detected in Iranian sheep breeds with synonymous effects on amino acid sequences [28]. However, generally, the *MC1R* gene is not the only genetic factor that determines the coat colour phenotype; there are also genes, such as *ASIP*, that are responsible for shaping the coat colour of sheep and other animal species [4, 5, 20, 23–25, 31].

We took one step further to validate the two variants identified in our GWAS analysis. We confirmed the association of deleterious substitution (rs409651063) and the upstream variant (rs408511664) with coat colour phenotype in an extra population using Sanger sequencing (DNA genotype, Table 3). Moreover, in our allele-specific expression analysis of the missense variant, we found that the mutant G allele was likely to be expressed in the skin of white hair from black head sheep, whereas the black part tended to express the reference A allele (cDNA genotype, Table 3). These results suggested that the missense mutation probably disrupted the black pigmentation receptor of *MC1R* and thus prevented normal pigmentation of the hair. The upstream variant (rs408511664) was also found to be associated with coat colour phenotype which might have regulatory effects of this mutation. Furthermore, the expression level of *MC1R* mRNA was significantly higher in sheep with black-head coat colour carrying the A allele than in white sheep (Fig 5A). Similarly, the *MC1R* mRNA expression was significantly higher in the individuals carrying GA alleles than in the individuals carrying GG alleles (Fig 5B). This finding showed that the *MC1R* gene has a great role in black-colour pigmentation process. The novel discovery of the two variants and allelic expression imbalance of the *MC1R* mutant and reference alleles helped to elucidate the mechanism of *MC1R* transcription and the genetic mechanism underlying coat colour in sheep.

In conclusion, we identified a genomic region, a candidate gene and genetic variants associated with the coat colour phenotype in Chinese Tan sheep using GWAS and the Ovine 600K SNP BeadChip. The genetic variants in *MC1R* may affect the coat colour phenotype of Chinese Tan sheep; however, other genes might have their own role in the pigmentation-making

process. Moreover, the *MC1R* gene expression level was significantly higher in black-head Tan sheep than in white sheep and significantly higher in GA alleles than in GG allele individuals. Thus, *MC1R* may have a significant role in regulating melanin synthesis and the development of black skin in Chinese Tan sheep. Furthermore, allele G was expressed in the skin covered with white hair, whereas both G and A alleles were expressed in the skin covered with black hair. Further studies are needed to explore other candidate genes, such as *ASIP*, and to determine their biological effect on coat colour formation in Tan sheep.

## Supporting information

**S1 File. This is the S1 map data file used for GWAS in this study.**  
(MAP)

**S2 File. This is the S2 ped data file used for GWAS in this study.**  
(PED)

**S3 File. This is the S3 phenotype data file used for GWAS in this study.**  
(TXT)

**S1 Table. This is the S1 gene expression data file used for gene expression analysis.**  
(DOCX)

## Author Contributions

**Conceptualization:** Gebremedhin Gebreselassie.

**Data curation:** Gebremedhin Gebreselassie.

**Formal analysis:** Gebremedhin Gebreselassie.

**Investigation:** Gebremedhin Gebreselassie.

**Project administration:** Lin Jiang, Zhengwei Zhao, Yuehui Ma.

**Software:** Gebremedhin Gebreselassie.

**Supervision:** Yuehui Ma.

**Validation:** Gebremedhin Gebreselassie, Benmeng Liang, Rabul Islam.

**Visualization:** Gebremedhin Gebreselassie, Haile Berihulay, Adam Abied.

**Writing – original draft:** Gebremedhin Gebreselassie.

**Writing – review & editing:** Gebremedhin Gebreselassie.

## References

1. Koseniuk A, Ropka-Molik K, Rubiś D, Smolucha G. Genetic background of coat colour in sheep. *Archives Animal Breeding*. 2018; 61(2):173–8.
2. Chaplin G. Geographic distribution of environmental factors influencing human skin coloration. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*. 2004; 125(3):292–302.
3. Jablonski NG, Chaplin G. The evolution of human skin coloration. *Journal of human evolution*. 2000; 39(1):57–106. <https://doi.org/10.1006/jhev.2000.0403> PMID: 10896812
4. Norris BJ, Whan VA. A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. *Genome research*. 2008; 18(8):1282–93. <https://doi.org/10.1101/gr.072090.107> PMID: 18493018

5. Fontanesi L, Beretti F, Riggio V, Dall'Olio S, Calascibetta D, Russo V, et al. Sequence characterization of the melanocortin 1 receptor (MC1R) gene in sheep with different coat colours and identification of the putative e allele at the ovine Extension locus. *Small Ruminant Research*. 2010; 91(2–3):200–7.
6. Cieslak M, Reissmann M, Hofreiter M, Ludwig A. Colours of domestication. *Biological Reviews*. 2011; 86(4):885–99. <https://doi.org/10.1111/j.1469-185X.2011.00177.x> PMID: 21443614
7. Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-Reh fuss L, Baack E, et al. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell*. 1993; 72(6):827–34. [https://doi.org/10.1016/0092-8674\(93\)90572-8](https://doi.org/10.1016/0092-8674(93)90572-8) PMID: 8458079
8. Hepp D, Gonçalves G, Moreira G, Freitas T, Martins C, Weimer T, et al. Identification of the e allele at the Extension locus (MC1R) in Brazilian Creole sheep and its role in wool color variation. *Genet Mol Res*. 2012; 11(11):2997–3006.
9. Chen S-Y, Huang Y, Zhu Q, Fontanesi L, Yao Y-G, Liu Y-P. Sequence characterization of the MC1R gene in yak (*Poephagus grunniens*) breeds with different coat colors. *BioMed Research International*. 2009; 2009.
10. Klungland H, Våge D. Pigmentary switches in domestic animal species. *Annals of the New York Academy of Sciences*. 2003; 994(1):331–8.
11. Hepp D, Gonçalves GL, Moreira GRP, de Freitas TRO. Epistatic interaction of the Melanocortin 1 Receptor and Agouti Signaling Protein genes modulates wool color in the Brazilian Creole Sheep. *Journal of Heredity*. 2016; 107(6):544–52. <https://doi.org/10.1093/jhered/esw037> PMID: 27288530
12. Nazari-Ghadikolaei A, Mehrabani-Yeganeh H, Miarei-Aashtiani SR, Staiger EA, Rashidi A, Huson HJ. Genome-wide association studies identify candidate genes for coat color and mohair traits in the iranian markhoz goat. *Frontiers in genetics*. 2018; 9:105. <https://doi.org/10.3389/fgene.2018.00105> PMID: 29670642
13. HAN J-I, Min Y, GUO T-t, YUE Y-j, LIU J-b, NIU C-e, et al. Molecular characterization of two candidate genes associated with coat color in Tibetan sheep (*Ovis arise*). *Journal of integrative agriculture*. 2015; 14(7):1390–7.
14. Pang Y, Geng J, Qin Y, Wang H, Fan R, Zhang Y, et al. Endothelin-1 increases melanin synthesis in an established sheep skin melanocyte culture. *In Vitro Cellular & Developmental Biology-Animal*. 2016; 52(7):749–56.
15. Yang G-L, Fu D-L, Lang X, Wang Y-T, Cheng S-R, Fang S-L, et al. Mutations in MC1R gene determine black coat color phenotype in Chinese sheep. *The Scientific World Journal*. 2013:8.
16. Bultman SJ, Michaud EJ, Woychik RP. Molecular characterization of the mouse agouti locus. *Cell*. 1992; 71(7):1195–204. [https://doi.org/10.1016/s0092-8674\(05\)80067-4](https://doi.org/10.1016/s0092-8674(05)80067-4) PMID: 1473152
17. Fontanesi L, Rustempašić A, Brka M, Russo V. Analysis of polymorphisms in the agouti signalling protein (ASIP) and melanocortin 1 receptor (MC1R) genes and association with coat colours in two Pramenka sheep types. *Small ruminant research*. 2012; 105(1–3):89–96.
18. Jackson PJ, Douglas NR, Chai B, Binkley J, Sidow A, Barsh GS, et al. Structural and molecular evolutionary analysis of Agouti and Agouti-related proteins. *Chemistry & biology*. 2006; 13(12):1297–305.
19. Lu D, Willard D, Patel IR, Kadwell S, Overton L, Kost T, et al. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature*. 1994; 371(6500):799. <https://doi.org/10.1038/371799a0> PMID: 7935841
20. Wilson BD, Ollmann MM, Kang L, Stoffel M, Bell GI, Barsh GS. Structure and function of ASP, the human homolog of the mouse agouti gene. *Human Molecular Genetics*. 1995; 4(2):223–30. <https://doi.org/10.1093/hmg/4.2.223> PMID: 7757071
21. Miller K, Gunn T, Carrasquillo MM, Lamoreux M, Galbraith D, Barsh G. Genetic studies of the mouse mutations mahogany and mahoganoid. *Genetics*. 1997; 146(4):1407–15. PMID: 9258683
22. Bennett DC, Lamoreux ML. The color loci of mice—a genetic century. *Pigment Cell Research*. 2003; 16(4):333–44. <https://doi.org/10.1034/j.1600-0749.2003.00067.x> PMID: 12859616
23. Li M, Tiirikka T, Kantanen J. A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (*Ovis aries*). *Heredity*. 2014; 112(2):122. <https://doi.org/10.1038/hdy.2013.83> PMID: 24022497
24. Fontanesi L, Dall'Olio S, Beretti F, Portolano B, Russo V. Coat colours in the Massese sheep breed are associated with mutations in the agouti signalling protein (ASIP) and melanocortin 1 receptor (MC1R) genes. *Animal*. 2011; 5(1):8–17. <https://doi.org/10.1017/S1751731110001382> PMID: 22440696
25. Gratten J, Pilkington J, Brown E, Beraldi D, Pemberton J, Slate J. The genetic basis of recessive self-colour pattern in a wild sheep population. *Heredity*. 2010; 104(2):206. <https://doi.org/10.1038/hdy.2009.105> PMID: 19672282

26. Switonski M, Mankowska M, Salamon S. Family of melanocortin receptor (MCR) genes in mammals—mutations, polymorphisms and phenotypic effects. *Journal of applied genetics*. 2013; 54(4):461–72. <https://doi.org/10.1007/s13353-013-0163-z> PMID: 23996627
27. Rouzaud F, Hearing VJ. Regulatory elements of the melanocortin 1 receptor. *Peptides*. 2005; 26(10):1858–70. <https://doi.org/10.1016/j.peptides.2004.11.041> PMID: 16005546
28. Amin M, Masoudi A, Amirinia C, Emrani H. Molecular Study of the Extension Locus in Association with Coat Colour Variation of Iranian Indigenous Sheep Breeds. *Russian journal of genetics*. 2018; 54(4):464–71.
29. Cavalcanti LCG, Moraes JCF, Faria DAd, McManus CM, Nepomuceno AR, Souza CJHd, et al. Genetic characterization of coat color genes in Brazilian Crioula sheep from a conservation nucleus. *Pesquisa Agropecuária Brasileira*. 2017; 52(8):615–22.
30. Kijas J, Serrano M, McCulloch R, Li Y, Salces Ortiz J, Calvo J, et al. Genomewide association for a dominant pigmentation gene in sheep. *Journal of animal breeding and genetics*. 2013; 130(6):468–75. <https://doi.org/10.1111/jbg.12048> PMID: 24236609
31. Royo LJ, Alvarez I, Arranz J, Fernández I, Rodríguez A, Pérez-Pardal L, et al. Differences in the expression of the ASIP gene are involved in the recessive black coat colour pattern in sheep: evidence from the rare Xalda sheep breed. *Animal Genetics*. 2008; 39(3):290–3. <https://doi.org/10.1111/j.1365-2052.2008.01712.x> PMID: 18384465
32. Våge DI, Klunghand H, Lu D, Cone RD. Molecular and pharmacological characterization of dominant black coat color in sheep. *Mammalian Genome*. 1999; 10(1):39–43. <https://doi.org/10.1007/s003359900939> PMID: 9892731
33. Wei C, Wang H, Liu G, Wu M, Cao J, Liu Z, et al. Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *Bmc Genomics*. 2015; 16(1):194.
34. Ma Q, Liu X, Pan J, Ma L, Ma Y, He X, et al. Genome-wide detection of copy number variation in Chinese indigenous sheep using an ovine high-density 600 K SNP array. *Scientific reports*. 2017; 7(1):912. <https://doi.org/10.1038/s41598-017-00847-9> PMID: 28424525
35. Johnston SE, McEWAN JC, Pickering NK, Kijas JW, Beraldi D, Pilkington JG, et al. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*. 2011; 20(12):2555–66. <https://doi.org/10.1111/j.1365-294X.2011.05076.x> PMID: 21651634
36. Zhao X, Dittmer KE, Blair HT, Thompson KG, Rothschild MF, Garrick DJ. A novel nonsense mutation in the DMP1 gene identified by a genome-wide association study is responsible for inherited rickets in Corriedale sheep. *PLoS One*. 2011; 6(7):e21739. <https://doi.org/10.1371/journal.pone.0021739> PMID: 21747952
37. Zhu C, Huang X, Li M, Qin S, Fang S, Ma Y. GWAS and post-GWAS to Identification of Genes Associated with Sheep Tail Fat Deposition. 2019.
38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*. 2007; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901
39. Fontanesi L, Beretti F, Riggio V, Dall'Olio S, González EG, Finocchiaro R, et al. Missense and non-sense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC genetics*. 2009; 10(1):47.
40. Muniz MMM, Caetano AR, McManus C, Cavalcanti LCG, Façanha DAE, Leite JHGM, et al. Application of genomic data to assist a community-based breeding program: a preliminary study of coat color genetics in Morada Nova sheep. *Livestock Science*. 2016; 190:89–93.