Whole genome sequences reveal multiple functions of PAT1 gene on environment adaptation and coat coloration in Tibetan sheep populations

Jianbin Liu（✉ qmn@sippe.ac.cn）

Chinese Academy of Sciences   https://orcid.org/0000-0003-2142-4024

Xuezhi Ding

Department of Animal Science

Chao Yuan

Department of Animal Science

Yufeng Zeng

Department of Animal Science

Tingting Guo

Department of Animal Science

Hao Zhang

Pizhou Agricultural and Rural Bureau of Jiangsu Province

Yaojing Yue

Department of Animal Science

Xilan Xu

Pizhou Animal Health Supervision Institute of Jiangnsu Province

Xiaoping Sun

Department of Animal Science

Chune Niu

Department of Animal Science

Bohui Yang

Department of Animal Science

Research article

Abstract

Background: Most sheep breeding programs designed for the tropics and sub-tropics have to take into account the impacts of both productive and adaptive traits. However, the genetic mechanism regulating the multiple biological process remain unclear. Results: In this study, we report a novel PAT1 gene that simultaneously explained the variations of productive trait (coat color), adaptive traits (altitude and geography response) in 15 indigenous Tibetan sheep populations. Overlapped genomic regions harboring 6 candidate genes across three traits were identified at 27 chromosomes, with the top 1% of Fst and |Zph| values. The SNP/INDELs and expression of these candidate genes were further analyzed, and we find that only PAT1 gene, a CSDE1 homologue was consistent with the variation of multiple traits regarding. Haplotype analysis of PAT1 reveal that Tibetan sheep breeds with C-type of PAT1 have significantly greater body weight, shear amount, chest width and body length, but have lower body height, than those with CA-type of PAT1. Conclusions: We emphasized that PAT1 gene could be a potentially selective target used for the improvements of environmental adaption and coat coloration in the future. These results contribute to the knowledge of adaptive response in Tibetan sheep populations and will help to guide future conservation programs for Tibetan sheep native to Qinghai-Tibetan Plateau.

Background

Recent research data and model predictions indicate that increasing frequencies of abnormal weather events due to global climate change will have a fundamental impact on agricultural production [1]. For example, the livestock experiences numerous environmental stressors that have effects on both productive traits and adaptive traits, such as growth, reproductive performance, the meat quality and cold tolerance [2, 3]. In particular, the influence of climate change – cold stress on living organisms is exacerbated at the interface of extreme environments that typically occurred in high altitude, plateaus and desert regions [4].

Designed breeding program for livestock should emphasize both productive traits and adaptive traits [5]. An interesting question arises here that except for cold environments adaptation, economic traits should be taken into account such as coat coloration. This is partially due to an intensive concern as to whether dark or light plumage is more effective in utilization of direct solar radiation since last century [6]. This phenomenon has been discussed on both theoretical and experimental bases [7]. For example, surface coloration arises from differential reflection or transmission of short wave radiation (the longest wavelength visually perceived by animals typically is near 700 nm). This leads to animals with dark-colored pelages or plumages, and hence ostensibly greater absorptivity for short-wave radiation, acquire greater heat loads from solar radiation that do animals with light-colored coats [8-10]. The heat load on the skin from solar radiation is associated with index of coat color, body size and shape.

Tibetan sheep (*Ovis aries*) has ability to adapt to a wide range of agro-ecological conditions, it provides an excellent model to gain new insights into genetic mechanisms underlying the rapid adaptations of livestock to extreme environments [11-12]. This is also helpful to develop appropriate breeding programs

under scenarios of future climate change. Tibetan sheep have lived on the Tibetan Plateau for thousands of years; however, the process and consequences of adaptation to this extreme environment conditions have not been elucidated [13]. This study is aimed to identify genes regulating environmental response and coat coloration in Tibetan sheep living in the Qinghai-Tibetan Plateau areas in China. Tibetan sheep representing different ecological regions in China and altitudes were used. Differentiated subpopulations were classified in respect of cold tolerance, geographic environments and coat coloration to uncover the casual genes incorporated with whole-genome sequencing. The key genes identified in this study would be helpful to guide breeding practices for improvements of both cold tolerance and coat coloration traits.

Methods

Ethics statement

All animals were handled according to the Guidelines for the Biological Studies Animal Care and Use Committee, People's Republic of China. Animal experiments were approved by the Animal Ethics Committee of the Institute of Animal Sciences of Chinese Academy of Agricultural Sciences.

Physiological measurements of samples from different altitudes

15 Chinese indigenous Tibetan sheep populations derived from different altitudes ranging from 3000-5000m were used, where these samples were divided into two subgroups, high altitudes (>4500m), and low altitudes (<3500m). 8 parameters related to body size: body weight/height/length, bust, chest depth, neck size, wool length, and shear amount; 13 physiological parameters: body temperature (Temperature), breathe rates (Breathe), pulse rates (Pulse), pulse interval (pulse interval), red-cell numbers (RBC), hematocrit, HCT, erythrocyte mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), Mean corpuscular hemoglobin concentration (MCHC), Hemoglobin (HGB), Platelet (PLT), Red blood cell volume distribution width coefficient variation (RGW-CV), White blood cell count (WBC). Biochemical parameters: Glutamic pyruvic transaminase (ALT), Glutamic pyruvic transaminase (AST), Total protein (TP), Albumin (ALB), Globulin (GLO), alkaline phosphatase (ALP), Lactate dehydrogenase (LDH), Cholinesterase (PCHE), Glucose (GLU), Total cholesterol (CHOL), Total calcium (CA); 9 blood-gas parameters: blood pH, pressure CO_2 (pCO_2), O_2 saturation (O2S), concentration of HCO_3^-, standard bicarbonate, total CO_2 concentration, Base Excess (BE), Standard Base Excess (SBE); structural parameters related to Lung tissue: bronchial, thin bronchial, end thin bronchial, alveolar numbers per area (AN/m2), thick alveolar interval (TAI).

Sample collection and sequencing

Ten milliliters of blood was collected from the jugular vein of each animal. From the 10 mL samples, 2 mL samples were quickly frozen in liquid nitrogen and stored at -80°C for genomic DNA extraction, as described previously [14]. The total DNA was extracted from the blood using the saturated salt method and adjusted to 50 ng/μL [15]. The blood were sampled from 636 Tibetan sheep in the Qinghai-Tibetan Plateau region in China as mentioned previously [16]. The detailed sampling information for the 15

indigenous Tibetan sheep populations, i.e., population code, sample number, altitude, longitude and latitude, sampling location, and geographical location, is shown in Table 1. For each Tibetan sheep, genomic DNA was extracted from 200 μL of peripheral venous blood using the QIAamp DNA blood mini kit (Qiagen, Germany) as described previously [17]. Briefly, A260/280 ratio and agarose gel electrophoresis were used to evaluate the quality and integrity of the DNA. Genomic DNA was masked to a 300-400bp fragment for library preparation, followed by end repair and ligated with an Illumina sequencing linker. The ligated products with sizes of 400-500bp were loaded on 2% agarose gels and subsequently amplified via PCR. Libraries were sequenced by HiSeq 2500 sequencer (Illumina) in a 2 × 100 bp paired end mode.

Reads alignment and variations calling

Reads were aligned to the sheep reference genome Ovis 3.1 (https://www.ncbi.nlm.nih.gov/genome/genomes/83) using BWA (0.6.2-r126 version) followed by duplicate removal using Picard-Tools-1.55 (http://broadinstitute. github.io/picard/). The Genome Analysis Toolkit (GATK-2.6) was used to perform local realignment around existing INDELs and base quality score recalibration. Variant detection was performed using the GATK Unified Genotyper. To filter SNPs for flowing analysis, at least three reads with different start sites supporting the non-reference allele had to be present.

Classification of types from SNP/INDEL variants

The 15 indigenous Tibetan sheep populations living in the Qinghai-Tibetan Plateau areas in China were classified into different subpopulations to estimate the candidate genes (Table 2). According to originated altitude information, two subpopulations were divided, i.e., lowland Tibetan sheep (~3500m): GD, QL, TJ, QH, MX, GJ, QK, and GN; highland Tibetan sheep (>4500m): LKZ, JZ, GB, HB, DM, AW, and LZ. In terms of coat coloration types, black coat coloration includes MX and GD, whereas white coat coloration includes QL, TJ, QH, GJ, QK, GN, LKZ, JZ, GB, HB, DM, AW, and LZ. For geographic regions, three different provinces in China were classified, Qinghai (QH, GD, QL, and TJ); Gansu (MX, GJ, QK, and GN); Titeban (LKZ, JZ, GB, HB, DM, AW and LZ).

Population genetics analysis

The pairwise genetic distance was determined relying on the number of allelic differences as described previously [17]. Briefly, the neighbor join tree was calculated based on the distance matrix Using PHYLIP (version 3.69) [18]. SNP pair with high correlation were removed via PLINK (PLINK, RRID: SCR 001757) [19]. The PCA and population structure were performed using EIGEN-SOFT (version 6.0.1) and FRAPPE software (version 1.1), respectively [20, 21]. As previously mentioned, TreeMix was applied to estimate migration events, with migration numbers m = 0-5 [22]. The selective sweep analysis was performed using VCFtools, v0.1.12b [23], including θ (number of isolated sites), π (paired nucleotide difference), and Tajima's D. There are 500,000 SNPs randomly selected from the genome, and these SNPs were used to analyze the linkage disequilibrium r2 with Haploview [24].

Genome-wide Selective Sweep Test

We calculated the genome-wide distribution of *Fst* values as previous reports in [25] for three subpopulation pairs as mentioned above as to originated altitude information, coat coloration types, and geographic regions (Table 2), using a sliding-window approach (100-kb windows with 50-kb increments). To identify regions that were likely to be or have been under selection, the "Z transformed heterozygosity" (*ZHp*) approach was used, as previously described [26]. Individual *Hp* values were Z transformed as follows: $ZHp = (Hp-\mu Hp)/\sigma Hp$, where *μHp* is the overall average heterozygosity, and *σHp* is the standard deviation for all windows within each population. We calculated the *ZHp* value in sliding 150-kb windows along the autosomes from sequence reads corresponding to the most and least frequently observed alleles at all SNP positions as previously described [27].

Bioinformatics analysis of breed specific SNPs/INDELs

Overlapped candidate genes across originated altitudes, coat coloration, and ecological regions were chosen for further analysis. In addition to this, we compared gene sequence of *Capra hircus*, *Pantholops hodgsonii*, and *Bos taurus* from NCBI database (https://www.ncbi.nlm.nih.gov/genome). We specifically focused on SNPs within genes and 1000-bp upstream and downstream flanking regions. Protein-protein interaction network was used to compare potential interactive genes with key overlapped genes in Homo sapiens from STRING database [28].

Validation of interactive genes

Expression of putative proteins interacted with key overlapped candidates relating to genetic variation of originated altitudes, coat coloration, and ecological regions as mentioned above were confirmed by qPCR. Detailed procedures for real-time PCR amplification were performed following standard protocol of SYBR Green real-time PCR kit (Applied Biosystems™ SYBR™ Green). The primer is listed in Supplementary Table S1. β-actin gene was used as internal control. The levels of relative gene expression were calculated against the expression of *RXFP2* [17], using 2-Ct method [29]. Four biological replicates and three technical replicates were conducted.

Results

Genomic sequencing and PCA in 15 in indigenous Tibetan sheep populations

In this study, we aim to identify the candidate genes that explain the variations of both productive and adaptive traits. To accomplish this, we sequenced 15 indigenous Tibetan sheep populations originated from different altitudes (Table 1), via high-resolution whole genomic sequences techniques (WGS). To dissect the genomic heterozygosity levels and recombination events, we classified these 15 indigenous Tibetan sheep populations into different subpopulations, including high altitude hypoxia (>4500m) vs. low altitude hypoxia (<3500m), white vs. black coat coloration, and three different geography locations, i.e., Qinghai, Gansu, and Tibetan (Table 2). WGS was performed on an Illumina HiSeq 2000 platform by

using the pooled DNA from each population. Genome sequencing yielded a total of 214 Gb raw data, and produced 230 to 310 million sequence reads per population (Table S2). Over 97.50% of the generated sequence reads mapped to approximately 94.15 % (87.42-98.53%) of the newly annotated Tibetan sheep reference genome (Ovis 3.1), indicating that high quality sequences were obtained (Table S3). Our efforts yielded an average sequence coverage of 10.3× per population, within a range of 9~14 fold. Single-nucleotide polymorphisms (SNPs) varied from 2-4 million for each population (Table S3-S4).

More than 9 million SNPs for each Tibetan sheep population that confidently remained after filtering were used in the subsequent analyses. Results from SNPs statistics showed that 63% SNPs were identified at intergenic regions, whereas only 0.7% SNPs cases were found at exon regions (Table S5). Principal components analysis (PCA) was performed to examine the genetic separation of 15 indigenous Tibetan sheep populations originated spanning in the Qinghai-Tibetan Plateau areas in China (Fig. 1A). PCA results clearly showed that most Tibetan sheep have high genetic homogeneity except Minxian Black Fur sheep (MX) and Jiangzi sheep (JZ) (Fig. 1B). After calling SNP, we obtained ~ 9.3 million SNPs across those Tibetan sheep populations. The distribution of minor allelic frequency (MAF) with 10 continued classes from 0-0.05 to 0.45-0.50 for each population was observed (Fig. 1C). MAF of SNPs across 15 indigenous Tibetan sheep populations showed MAF within the range of 0⎼5% had most relative abundance with around 20% of total SNPs, whereas 5⎼15% MAF showed relatively low abundance with less than 3% across 15 indigenous Tibetan sheep populations. The distribution pattern of MAF among Tibetan sheep populations was similar except for LKZ (Fig. 1C). Similarity index (IS) value was used to estimate the homogeneity of SNPs across 15 indigenous Tibetan sheep populations. LKZ showed lowest similarity with other Tibetan sheep populations, which is consistent with MAF distribution results (Fig. 1D).

Overlapped selective signals regarding genomic heterozygosity among productive and adaptive traits

Fst values over the whole genome based on genetic differentiation in different Tibetan sheep populations were classified by altitude hypoxia (altitude), coat coloration (color), and geography locations, named geography. Some overlapped genes were identified with high *Fst* values, including *ZEB, RALY, 1L12RB1, ASIP*, and *ARAP1* (Fig. 2A-C; Table S6). Results showed that 31, 1, and 22 common genes were identified that in classifications of altitude & geography, color & altitude, and color & geography, respectively (Fig. 2D).

In addition, we found other 5 genes, except *IL12RB1* at chr.5 which was identified in both coat coloration and geography. Among these genes we cite; *CSF2, LOC106990309* that specifically in Tibetan geography locations, *TRNAC-GCA, ACSL6* and *MEIKIN* were specifically in white coat coloration subpopulations (Figure S1; Table S6). At chr.15, another 11 genes, except *ARAP1* mainly related to coat coloration and geography clusters, were identified with high *Fst* values, including *LOC101104426, LOC101105946, LOC101104685, LOC101106201, LOC101104933, LOC101106712, LOC101105186, LOC101104177, LOC101105439, LOC101106459*, and *LOC101105686* (Figure 2E; Table S6). These genes also showed high heterozygosity levels specifically to Tibetan geography locations and white coat coloration

populations (Figure S1 and S3). We found 4 genes in chr.18 including *CRYBB3, ADRBK2, KIAA1671*, and *CRYBB2* (Figure 2E). These genes have high heterozygosity specifically in Tibetan geography locations (Figure S3). At chr.23, there are two genes, *PSMA8* and *TAF4B*, possessing high *Fst* values as well as high heterozygosity levels, which are identified uniquely in low altitude subpopulations (Figure S2; Table S6). In addition, there are 7 genes at chr.27 with high *Fst* values and heterozygosity in low altitude subpopulations, including *LOC105605498, ITGB1BP2, LOC105605499, LOC101104554, GJB1, ZMYM3* and *NONO* (Table S6).

Moreover, we also identified 13 novel genes at chr.13 with high *Fst* values that identified in both altitude and geography locations, including *ZEB1, PANK2, RNF24, LOC101110166, VPS16, PCED1A, PTPRA, TMEM239, C12H20, LOC101113781, CPXM1, LOC105606248 and EBF4*. Among these genes, *ZEB1* has high heterozygosity specifically in low altitude subpopulations; however, *PANK2, RNF24, LOC101110166, VPS16, PCED1A, PTPRA* are specifically in Tibetan geography subpopulations with high heterozygosity, while *TMEM239, C12H20, LOC101113781, CPXM1, LOC105606248* and *EBF4* are specifically in white coat coloration subpopulation (Table S6). These results suggested that these overlapped genes identified in both altitude and geography locations are related to convergent adaption to high altitudes in indigenous sheep breeds.

Notably, among three defined clusters, i.e., cold, color and environments, there are 6 common genes that have been identified simultaneously. These genes are located at chromosome 27. These genes include *PAT1* (*LOC101123097*, also named productive and adaptive traits), *KLF4, PSMA1, POF1B, APOOL*, and *ZNF711* (Table S6). These genes have high heterozygosity levels within either white coat coloration or low altitude subgroups (Table S6).

Haplotypes of PAT1 are related to adaptive response to environments and body size

To further analyze the natural variations of these 6 overlapped genes, we analyzed INDELs of these genes including 2k-bp promoter and CDS regions along with other species, such as *Capra, Pantholops*, and *Bos Taurus* (Figure 3). Results from INDEL information showed that 62% INDEL events were identified at intergenic regions, whereas only 0.3% INDEL cases were found at exon regions (Table S7). Our findings suggested that among 6 genes, only INDEL at 5'-flank region of *PAT1* can fully explain the variations of different defined clusters in altitude variations, coat color, and geography locations. C-type is occurred in high altitude, black coat coloration, and Gansu geography location, while CA-type is occurred in low altitude, white coat coloration, and both Qinghai and Tibetan geography regions (Figure S1-S3).

We further analyzed the protein structure and gene expression pattern among different indigenous Tibetan sheep populations (Figure 4). According to bioinformatics analysis, we predicted the protein sequences of *PAT1* according to translate tool in ExPASy database (https://web.expasy.org/translate/), and found the similarities score is 99% in both mRNA and protein sequences between *CSDE1* and *PAT1* protein, indicating its counterpart roles of *PAT1* with *CSDE1*. We predicted the secondary structure of *PAT1* with 798 amino acids, and found there are 13 different amino acids. Most variation loci were observed at N-terminal of protein containing 7 different amino acids, and these amino acids were located

in either Beta strand or Turn domains, but not Helix (Figure 4A). Three different amino acids were mapped at 3D protein structure from two PDB entries (1WFQ and 2YTX) according to Uniprot database (CSDE1_Human Protein ID: O75534).

To compare the expression between *PAT* and *CSDE1*, we used 6 indigenous Tibetan sheep populations, including Guide Black Fur sheep (GD), Qilian White Tibetan sheep (QL), Minxian Black Fur sheep (MX), Ganjia Tibetan sheep (GJ), Huoba Tibetan sheep (HB) and Awang Tibetan sheep (AW). These Tibetan sheep can be classified into different populations regarding altitude, coat coloration and geography locations. We found *PAT1* gene is highly expressed in high altitude, black coat coloration subpopulations, rather than low altitude and white coat coloration subpopulations. *PAT1* gene also performed distinct pattern among different geography locations (Figure 4C). *CSDE1* gene exhibited similar pattern with *PAT1* across different Tibetan sheep populations, except in QL which showing high expression of PAT1, but is a white coat coloration sheep (Figure 4D). Therefore, an interesting question arising here is whether interactive proteins with CSDE1 could show similar expression pattern between the subpopulations of altitude, color and geography region? Therefore, we reconstructed the protein-protein interaction networks using CSDE1, and found *PABPC1* and *MYC*, with relatively high correlation with PAT1, bears similar expression pattern with PAT1 (Figure S4).

To interpret the biological functions of *PAT1*, we compared 47 physiological and biochemical parameters in representative 4 indigenous Tibetan sheep populations with different promoter haplotypes of *PAT1* gene (Figure 5; Table 2). As mentioned in Figure 3C, there are two haplotypes, C-type and CA-type, containing AW&HB, and GJ & QL Tibetan sheep populations, belong to both high and low altitude clusters, respectively. Tibetan sheep subpopulations with C-type of *PAT1* have significantly greater body weight, shear amount, chest width and body length, but have low body height, relative to sheep in subgroups with CA-type of *PAT1* (Figure 5A-B).

Discussion

Global climate change would have a fundamental impact on agricultural production. Integrated characteristic for livestock needs to be taken in consideration for designed breeding program, including adaptive ability to high altitudes, black coat coloration, and widespread geographical locations. Tibetan sheep is known to adapt to various agro-ecological conditions, it therefore represents an excellent model for understanding genetic mechanisms on adaptations of livestock to extreme environmental conditions. This study is aimed to identify key genes those regulate environmental response and coat coloration simultaneously.

Overlapped genes in the explanation of altitude and geography variations

It is common that hypoxia-inducible factors (HIFs) are transcription factors that respond to changes in the available oxygen in the cellular environment under high-altitude conditions. In this study, we identified a candidate gene, *ZEB1* at chr.13 with top 1% *Fst* values in both altitude and geography variations (Figure

2 A-C). It is reported that *ZEB1* is a target gene of HIFs, which is critical in the regulation of the macrovascular angiogenic response but not that of microvascular angiogenesis [32]. Moreover, under hypoxic conditions, HIF-1α up-regulates the expression of proteins that induce *TWIST1* and *ZEB1* [33,34]. *RALY* is identified to be correlated with both altitude and geography variations, which has been reported as candidate genes with significant enrichments in spliceosomal complex pathway in high-altitude adaption Tibetan sheep populations, indicating its important roles in hypoxia response [13]. *Fst* values, also named fixation index, were typically used to evaluate differentiated genomic regions and identify selective signals among whole-genomic SNPs [30, 31]. *IL12RB1* is identified with high *Fst* values for both color and geography (Figure 2 B-C), and *IL12RB1* also has high levels of heterozygosity specifically in Tibetan geography locations (Figure S3). The numbers of totally overlapped candidate genes were presented based on top 1% *Fst* values for different combinations among altitude, color and geography from 150-kb windows (Figure 2D).

Overlapped genes in the explanation of coat coloration and geography variations

Differentiated genomic regions in classification of geography factor reflect environmental selections and robustness. We identified that two known genes that is related to variation of both geography and coat coloration. *ARAP1* gene at chr.15 has been previously reported to be under climate-driven selection as signal molecular involving in GTPase regulator activity process [35], which is in charge of various types of membrane transport, such as vesicle fusion and docking of transport vesicles to specific target organelles and/or plasma membranes during secretory processes [36,37]. It can anchor and transport melanosomes to the plasma membrane, and melanosomes within melanocytes is critical for pigmentation synthesis of hair [38]. *IL12RB1* at chr.5, was also correlated with environmental selection [35]. It is involved in cytokine-cytokine receptor interaction pathways, and cytokine is secreted by keratinocytes and by melanocytes during the first two steps of Hyperpigmentation in the epidermis [39]. We also compared levels of genomic heterozygosity within different subpopulations as represented by |ZHp| values [40], and found *IL12RB1* has high levels of heterozygosity specifically in Tibetan geography locations (Figure S3).

Coat coloration, altitude and geography variations

Our findings suggested that only variation of SNP/INDEL at 5'-flank region of *PAT1* can fully explain the phenotypic variations in combined characteristics of altitude response, coat color, and geography locations (Figure 3C). This indicates that the expression of this gene could play key roles in regulating these productive and adaptive traits, which would be discussed in the following. For other genes, the INDEL variation only can explain partially to defined clusters. For example, the *KLF4* has 4 types of INDELs, in 5'flank and other 3 intron regions. The types of INDEL can explain the variation of altitudes and coat coloration, but not three geography types. *PSMA1* can explain coat coloration variation but cannot explain variation of other two clusters, whereas *ZNF711* can explain altitude variation, but cannot explain variation of another two clusters. Taken together, only the 5'-UTR of *PAT1* can largely explain the

variance of three classifications, indicating that genetic basis of such as coat coloration and adaptive traits relies on gene expression.

Function of PAT1 in productive and adaptive characteristics

The *CSDE1* (Cold Shock Domain Containing E1), alternative name: *UNR*, is a conserved RBP containing five cold-shock domains (CSDs) that bind single-stranded RNA [41, 42]. We confirmed *CSDE1* homologous protein, *PAT1* showed clear differences between high low altitudes subpopulation of Tibetan sheep populations and between different geography locations with changed climate conditions (Figure 3C). CSDs represent the most evolutionarily conserved nucleic acid-binding protein domain, found in bacteria and eukaryotes [43, 44]. This domain of around 70 amino acid residues mediates binding to single-stranded DNA and RNA [45, 46]. Function of proteins possessing CSDs are involved in two processes: transcriptional and translational control. Two major old shock proteins CspA and CspB in Escherichia coli and Bacillus subtilis, respectively, were intensively reported. CspA and CspB are massively and transiently induced after a temperature downshift and are involved in the adaptation to cold shock [47].

To explain the relationship of *PAT1* and *CSDE1* with coat coloration, we reconstructed the *CSDE1*-centerized module according to STRING protein-interaction database, and we found 9 genes that involving in this module. These are: *MAX, MYC, YBX1, HNRNPU, DHX9, PABPC1, PAIP1, STRAP, SYNCRIP*. Differential gene expression analysis via qPCR suggested that these genes have strong correlation with *CSDE1* and *PAT1* (Table S8). Previous reports showed that *CSDE1* mRNA is up-regulated in a high percentage of skin and ovary cancers, and *UNR/CSDE1* regulates critical Melanoma genes, including *PABPC1* gene [48]. Beside, *CSDE1* regulates internal ribosome entry site (IRES)-dependent translation of the transcripts encoding the oncogene *MYC* [49]. The both *PABPC1* and *MYC* were identified in *CSDE1*-interative protein network (Figure S4), and we confirmed that expression of the two genes showed distinct pattern between white coat coloration and black coat coloration subpopulations, especially for *PABPC1* (Figure S4C). These inherent correlations probably explain the variation of coat coloration regulated by *CSDE1/PAT1*.

Morphological and physiological parameters of different haplotypes

Different haplotypes of *PAT1* could be reflective of altitude response, and is related to changes of morphological characters, such as body weight, shear amount, chest width and body length and body height (Figure 5A-B). This is consistent with the conclusion that large physical size in highland as observed in man [50]. In particular, body weight of Tibetan sheep in C-type population is 1.7 times higher than that in CA-type group. For body haematological parameters, body temperature, pulse interval, MCH, MCHC, HGB levels are at least 1.5 times higher in C-type of *PAT1* group than that in CA-type population; this is similar to that observed in Tibetan sheep and dogs [13, 51]. Interestingly, breathe rates, pulse rates, RBC, HCT, MCV, PLT, and RDW-CV are relatively lower in C-type of *PAT1* than that in CA-type population. Values of blood-gas parameters were similar or relatively lower in C-type population than that in CA-type population, except for pO2. For lung tissue structure parameters, values are relatively higher in C-type

population than that in CA-type population. In particular, elevated lactate dehydrogenase (LDH) in C-type population relative to CA-type population support the found in other sheep studies on hypoxemia response [52].

Conclusion

Dissection of casual genes and molecular mechanism responsible for high altitude and coat coloration in this study is pivotal in Tibetan sheep objective breeding towards improvements of productive and adaptive traits. We identified a *CSDE1* like gene, named *PAT1*, variation of INDEL at chr 27: 72792241 within 5′-UTR explain the genetic diversity and phenotypic classification in given 15 indigenous Tibetan sheep populations. We highlight that the importance of *PAT1* in future Tibetan sheep breeding program for environmental adaptation and coat economics.

Abbreviations

PAT: productive and adaptive traits; Temperature: body temperature; Breathe: breathe rates; Pulse: pulse rates; pulse interval: pulse interval; RBC: red-cell numbers; HCT: hematocrit; MCV: erythrocyte mean corpuscular volume; MCH: mean corpuscular hemoglobin; MCHC: Mean corpuscular hemoglobin concentration; HGB: Hemoglobin; PLT: Platelet; RGW-CV: Red blood cell volume distribution width coefficient variation; WBC: White blood cell count; ALT; Glutamic pyruvic transaminase; AST: Glutamic pyruvic transaminase; TP: Total protein; ALB: Albumin; GLO: Globulin; ALP: alkaline phosphatase; LDH: Lactate dehydrogenase; PCHE: Cholinesterase; GLU: Glucose; CHOL: Total cholesterol; CA: Total calcium; pCO2: pressure CO2; O2S: O2 saturation; BE: Base Excess; SBE: Standard Base Excess; AN/m2: end thin bronchial, alveolar numbers per area; TAI: thick alveolar interval; Guide Black Fur sheep: GD; Qilian White Tibetan sheep: QL; Tianjun White Tibetan sheep: TJ; Qinghai Oula Tibetan sheep: QH; Minxian Black Fur sheep: MX; Ganjia Tibetan sheep: GJ; Qiaoke Tibetan sheep: QK; Gannan Oula Tibetan sheep: GN; Langkazi Tibetan sheep: LKZ; Jiangzi Tibetan sheep: JZ; Gangba Tibetan sheep: GB; Huoba Tibetan sheep: HB; Duoma Tibetan sheep: DM; Awang Tibetan sheep: AW; Linzhou Tibetan sheep: LZ; MAF: minor allelic frequency; HIFs: hypoxia-inducible factors; *CSDE1*: Cold Shock Domain Containing E1; CSDs: cold-shock domains.

Declarations

Ethics approval and consent to participate:

All animals were handled according to the Guidelines for the Biological Studies Animal Care and Use Committee, People's Republic of China. Animal experiments were approved by the Animal Ethics Committee of the Institute of Animal Sciences of Chinese Academy of Agricultural Sciences.

Declarations

Consent for publication:

Not applicable.

Declarations

Availability of data and material:

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests:

The authors declare that they do not have any competing interests

Declarations

Funding:

Declarations

Authors' contributions:

Conceived and designed the experiments: LJB DXZ ZYF YC. Performed the experiments: LJB DXZ SXP NCE. Analyzed the data: ZYF YC GTT. Wrote the paper: LJB YBH.

Declarations

Acknowledgements:

References

1. Bernstein L, Bosch P, Canziani O. IPCC, 2007: Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change, Geneva.

2. Brown ME, Funk CC. Food security under climate change. Science. 2008; 319:580–581.

3. Hoffmann I. Climate change in context: implications for livestock production and diversity. In: Odongo NE, Garcia M, Viljoen GJ, editors. Sustainable improvement of animal production and health. Vienna: IAEA-FAO. 2010: 33–44.

4. Easterling DR, Meehl GA, Parmesan C, Changnon S, Karl TR, Mearns LO. Climate extremes: observations, modeling, and impacts. Science. 2000; 289: 2068–2074

5. König EZ, Mirkena T, Strandberg E, Audho J, Ojango J, Malmfors B, Okeyo AM, Philipsson J. Participatory definition of breeding objectives for sheep breeds under pastoral systems—the case of Red Maasai and Dorper sheep in Kenya. Trop Anim Health Prod. 2016; 48: 9–20.

6. Lustick S, Adams M, Hinko A. Interaction between posture, color, and the radiative heat load in birds. Science. 1980; 208: 1052-1053.

7. Walsberg GE, Campbell GS, King JR. Animal coat color and radiative heat gain: a re-evaluation. J. Comp. Physiol. 1978; 126: 211-222.

8. Hamilton WJ, III. Life's color code. McGraw-Hill, New York. 1973; pp 214-221.

9. Lustick S. Plumage color and energetics. Condor. 1971; 73: 121-222.

10. Lustick S, Battersby B, Kelty M. Behavioral thermoregulation: orientation toward the sun in Herring Gulls. Science. 1978; 200: 81-83.

11. Howden SM, Soussana JF, Tubiello FN, Chhetri N. Adapting agriculture to climate change. Proc Natl Acad Sci USA. 2007; 104:19691–19696.

12. Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL. Prioritizing climate change adaptation needs for food security in 2030. Science. 2008; 319: 607–610.

13. Wei C, Wang H, Liu G, Zhao F, Kijas JW, Ma Y, Lu J, et al. Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. Sci Rep. 2016; 6: 2677-2681.

14. D'Angelo F, Ciani E, Sevi A, Albenzio M, Ciampolini R, Cianci D. The genetic variability of the Podolica cattle breed from the Gargano area. Preliminary results. Ital J Anim Sci. 2006; 5: 79–85.

15. Sambrook J, Russell DW. Molecular cloning: A laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.

16. Liu J, Ding X, Zeng Y, Yue Y, Guo X, Guo T, Chu M, et al. Genetic diversity and phylogenetic evolution of Tibetan Sheep based on mtDNA D-Loop sequences. PloS one. 2016;11(7):e0159308.

17. Pan ZY, Li SD, Liu QY, Wang Z, Zhou ZK, Di R, Miao BP, Hu WP, Wang XY, Hu XX, Xu Z. Whole-genome sequences of 89 Chinese sheep suggest role of RXFP2 in the development of unique horn phenotype as response to semi-feralization. Gigascience. 2018; 24: 42-47.

18. Felsenstein, J. PHYLIP - Phylogeny Inference Package (version 3.2). Cladistics. 1989; 5:164−6.

19. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, et al. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007; 449: 913−8.

20. Patterson N, Price AL, Reich D. Population structure and eigen analysis. PLoS Genet, 2006; 2: 2074−93.

21. Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. Genet Epidemiol. 2005; 28: 289−301.

22. Pickrell JK, Pritchard JK. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet, 2012; 8(11):e1002967. doi:10.1371/journal.pgen. 1002967.

23. Danecek P. Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2001; 27: 2156−8.

24. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics, 2005; 21:263−5.

25. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984; 38: 1358−1370, doi: 10.2307/2408641.

26. Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MA, Andersson L. Strong signatures of selection in the domestic pig genome. PNAS. 2012; 109: 19529−19536.

27. Rubin CJ, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, et al.. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature. 2010; 464: 587−591

28. Von Mering, C. Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. String: known and predicted protein−protein associations, integrated and transferred across organisms. Nucleic acids research 2005; 33: 433−437.

29. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods, 2001; 25: 402−8.

30. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984; 38: 1358–1370, doi: 10.2307/2408641.

31. Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J, Nicholas TJ, Neff MW. Tracking footprints of artificial selection in the dog genome. PNAS. 2010; 107: 1160–1165.

32. Luo Z, Wen G, Wang G, Pu X, Ye S, Xu Q, Wang W, Xiao Q. MicroRNA-and play an important role in endothelial cell differentiation and vasculo genesis by targeting transcription repressor ZEB1. Stem Cells. 2013; 31: 1749–1762.

33. Martin A, Cano A. Tumorigenesis: twist links EMT to self-renewal. Nat Cell Biol. 2010; 12: 924–925.

34. Semenza GL, Nejfelt MK, Chi SM, Antonarakis SE. Hypoxia-inducible nuclear factors bind to an enhancer element located 3' to the human erythropoietin gene. P Natl Acad Sci. USA. 1991; 88: 5680–5684.

35. Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, Joost S, Li MH, Ajmone Marsan P. Adaptations to climate-mediated selective pressures in Sheep. Mol. Biol. Evol., 2014; 31: 3324-43

36. Fukuda M. Regulation of secretory vesicle traffic by Rab small GTPases. Cell Mol Life Sci. 2008; 65: 2801–2813.

37. Stenmark H. Rab GTPases as coordinators of vesicle traffic. Nat Rev Mol Cell Biol. 2009; 10: 513–525.

38. Steingrímsson E, Copeland NG, Jenkins NA. Mouse coat color mutations: From fancy mice to functional genomics. Dev Dyn. 2006; 235: 2401–2411.

39. Murase D, Hachiya A, Amano Y, Ohuchi A, Kitahara T, Takema Y. The essential role of p53 in hyperpigmentation of the skin via regulation of paracrine melanogenic cytokine receptor signaling. J. Bio. Chem. 2009; 284: 4343–4353.

40. Bryc K, Patterson N, Reich D. A novel approach to estimating heterozygosity from low-coverage genome sequence. Genetics. 2013; 195: 553–561.

41. Goroncy AK, Koshiba S, Tochio N, Tomizawa T, Inoue M, Watanabe S, Harada T, Tanaka A, Ohara O, Kigawa T, Yokoyama S. The NMR solution structures of the five constituent cold-shock domains (CSD) of the human UNR (upstream of N-ras) protein. J. Struct. Funct. Genomics. 2010; 11, 181–188.

42. Triqueneaux G, Velten M, Franzon P, Dautry F, Jacquemin-Sablon H. RNA binding specificity of Unr, a protein with five cold shock domains. Nucleic Acids Res. 1999; 27, 1926–1934.

43. Graumann P, Wendrich TM, Weber MH, Schröder K, Marahiel MA. A family of cold shock proteins in Bacillus subtilis is essential for cellular growth and for efficient protein synthesis at optimal and low

temperature, Mol. Microbiol, 1997; 25: 741-756.

44. Graumann P, Mahariel MA. A case of convergent evolution of nucleic acid binding modules. Bioessays, 1996; 18: 309-315.

45. Wolffe AP. Structural and functional properties of the evolutionary ancient Y-box family of nucleic acid binding proteins. BioEssay. 1994; 16: 245–251.

46. Sommerville J, Ladomery M. Transcription and masking of mRNA in germ cells: involvement of Y-box proteins, FASEB J. 1996; 10: 435-443.

47. Thieringer HA, Jones PG, Inouye M. Cold shock and adaptation. Bioessay, 1998; 20: 49-57.

48. Wurth L, Papasaikas P, Olmeda D, Bley N, Calvo GT, Guerrero S, et al. UNR/CSDE1 drives a post-transcriptional program to promote melanoma invasion and metastasis. Cancer Cell, 2016; 30: 694–707.

49. Evans JR, Mitchell SA, Spriggs KA, Ostrowski J, Bomsztyk K, Ostarek D, Willis AE. Members of the poly (rC) binding protein family stimulate the activity of the c-myc internal ribosome entry segment in vitro and in vivo. Oncogene. 2003; 22: 8012–8020.

50. Lenfant, C. High altitude adaptation in mammals. Amer. Zool., 1973; 13: 447-456.

51. Gou X. Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z, Lu S, Lian L, Wu C, Wang X, Li G, Ma T, et al. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. Genome Res. 2014; 24: 1308–1315

52. Ohtsuka T, Gilbert RD. Cardiac enzyme activities in fetal and adult pregnant and nonpregnant sheep exposed to high-altitude hypoxemia. J. Appl. Physiol. 1995; 79: 1286-1289.

Tables

Table 1. Sampling information for the 15 indigenous Tibetan sheep populations used in this study

Population	population code	Sample number	Altitude (m)	Longitude and latitude	Sampling location
Guide Black Fur sheep	GD	39	3100	N:38°61'152" E:103°32'160"	Senduo Town, Guinan County, Hainan Tibetan Autonomous State, Qinghai Province
Qilian White Tibetan sheep	QL	44	3540	N:42°20'178" E:116°64'618"	Qilian Town, Qilian County, Delingha City, Mongolian Autonomous State, Qinghai Province
Tianjun White Tibetan sheep	TJ	64	3217	N:42°18'158" E:116°42'210"	Shengge Countryside, Tianjun County, Delingha City, Mongolian Autonomous State, Qinghai Province
Qinghai Oula Tibetan sheep	QH	44	3630	N:34°16'433" E:101°32'141"	Jianke Village, Kesheng Town, Henan Mongolian Autonomous County, Qinghai Province
Minxian Black Fur sheep	MX	67	3180	N:36°54'48" E:103°94'107"	Taizi Village, Qingshui Town, Minxian County, Dingxi City, Gansu Province
Ganjia Tibetan sheep	GJ	58	3022	N:35°32'49" E:102°40'802"	Xike Village, Ganjia Town, Xiahe County, Gannan Tibetan Autonomous State, Gansu Province
Qiaoke Tibetan sheep	QK	71	3410	N:35°42'106" E:102°42'210"	Waeryi Village, Qihama Town, Maqu County, Gannan Tibetan Autonomous State, Gansu Province
Gannan Oula Tibetan sheep	GN	52	3616	N:33°51'312" E:101°52'424"	Daerqing Administrative Village, Oula Town, Maqu County, Gannan Tibetan Autonomous State, Gansu Province
Langkazi Tibetan sheep	LKZ	10	4459	N:28°58'951" E:090°23'757"	Kexi Village, Langkazi Town, Langkazi County, Shannan Territory of Tibet Autonomous Region
Jiangzi Tibetan sheep	JZ	46	4398	N:28°55'113" E:089°47'692"	Reding Village, Cheren Town, Jiangzi County, Shannan Territory of Tibet Autonomous Region
Gangba Tibetan sheep	GB	85	4403	N:28°15'281" E:088°24'787"	Yulie Village, Gangba Town, Gangba County, Rikaze Territory of Tibet Autonomous Region
Huoba Tibetan sheep	HB	34	4614	N:30°13'822" E:083°00'249"	Rima Village, Huoba Town, Zhongba County, Rikaze Territory of Tibet Autonomous Region
Duoma Tibetan sheep	DM	8	4780	N:29°48'609" E:091°36'191"	Sixth Village, Maqu Town, Anduo County, Naqu Territory of Tibet Autonomous Region
Awang Tibetan sheep	AW	5	4643	N:30°12'101" E:098°63'098"	Ayi Third Village, Awang Town, Gongjue County, Changdou Territory of Tibet Autonomous Region
Linzhou Tibetan sheep	LZ	9	4292	N:29°09'121" E:091°25'063"	Tanggu Village, Tanggu Town, Linzhou County, Tibet Autonomous Region

Table 2 Criterion of indigenous Tibetan sheep populations classified for different factors

Types	Parameters	Subpopulation
Altitude	<3500 m	GD、QL、TJ、QH、MX、GJ、QK、GN
	>4500 m	LKZ、JZ、GB、HB、DM、AW、LZ
Coat coloration	Black	MX、GD
	White	QL、TJ、QH、GJ、QK、GN、LKZ、JZ、GB、HB、DM、AW、LZ
Geographic locations	Qinghai	GD、QL、TJ
	Gansu	QH、MX、GJ、QK、GN
	Tibetan	LKZ、JZ、GB、HB、DM、AW、LZ

Figures
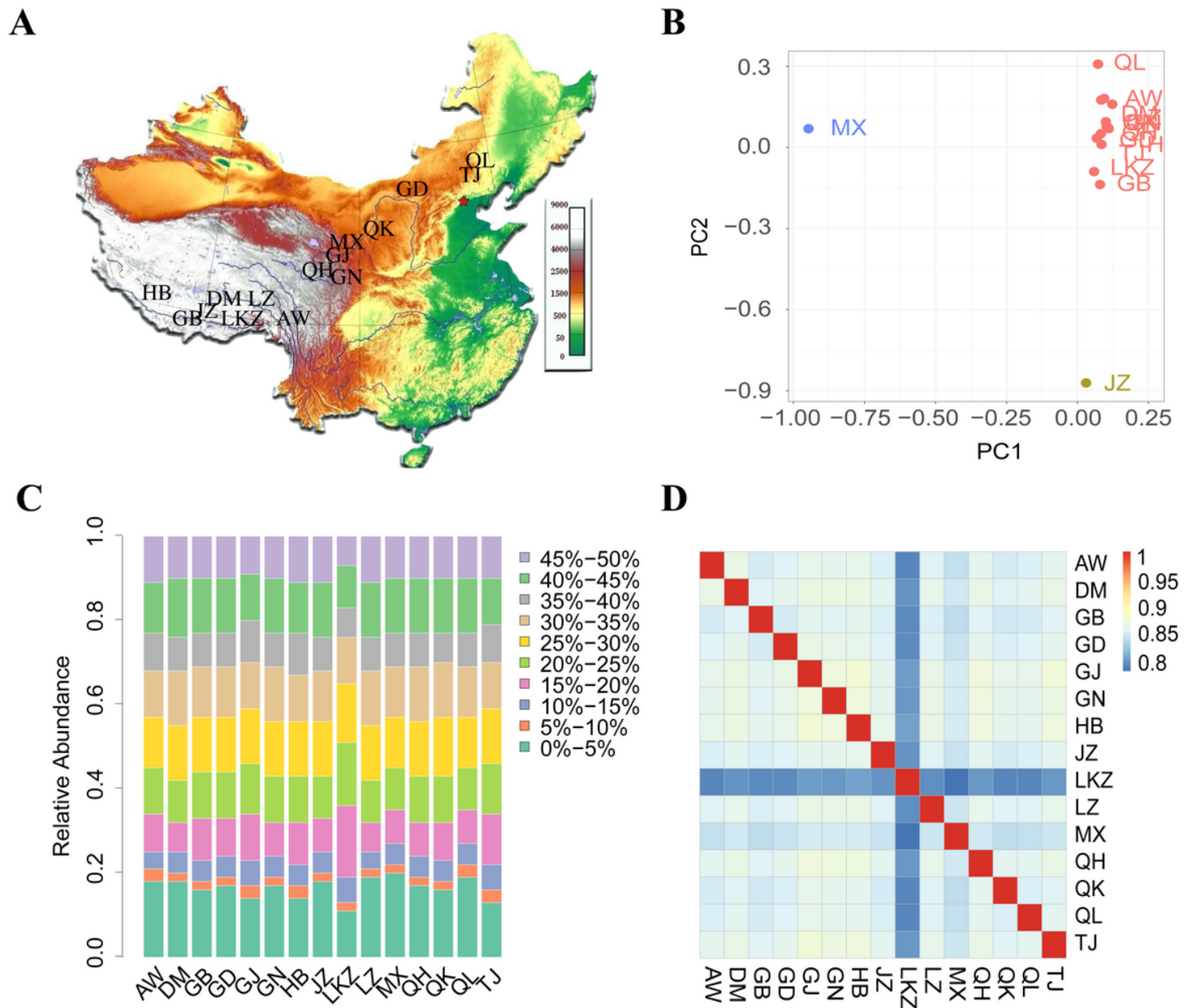
Figure 1

SNP analysis and population structure for 15 indigenous Tibetan sheep populations. A: distributive map of 15 indigenous Tibetan sheep populations living in the Qinghai-Tibetan Plateau areas in China used in this study; B: Principal component analysis on whole genomic DNA sequencing in different Tibetan sheep populations; c: Distribution of minor allelic frequency (MAF) with 10 continued classes from 0-0.05 to 0.45-0.50; D: Relatedness of similarity index (IS) values in 15 Tibetan sheep populations. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
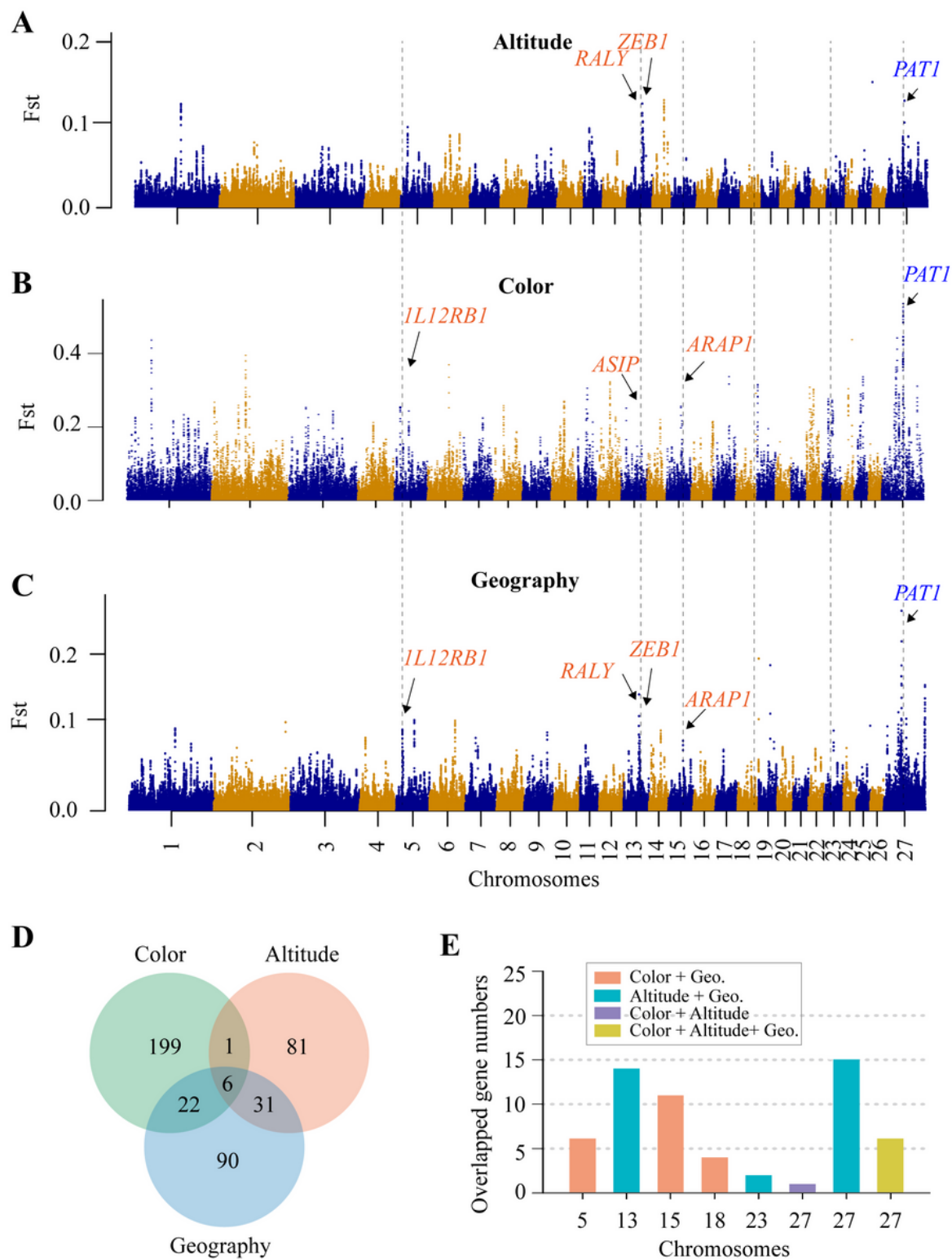
Figure 2

Manhattan analysis and candidate genes exploration. A-C: Manhattan plot representing Fst values for each SNP across Chromosomes for three clustered subpopulations. The overlapped genes were highlighted in blue fonts. Identified genes in two and three clusters were labeled in yellow and blue, respectively. D: Venn diagram representing overlapped genes across different clustered subpopulations.

E: Numbers of corresponding genes harboring overlapped SNPs across different clustered subpopulations.

A. LOC101119700 (55463621..55465455)

	5'flank	Intron1	Intron2	Intron5
Position	55463804	+36	+48	+153
Ref. Gen.	A	CT	G	A
Alt._A	A	CT	G	A
Alt._B	AT	T	GCC	ATT
Color_A	AT	T	GCC	ATT
Color_B	AT(A)	CT(T)	GCC(G)	ATT(A)
Geo._A	A	T	G	A
Geo._B	AT	CT	GCC	ATT
Geo._C	AT(A)	CT(T)	GCC(G)	ATT(A)
Capra	—	—	—	—
Pantholops	—	T	—	—
Bos taurus	—	—	—	—

B. LOC101119112 (56044836..56045636)

	Intron1	Exton2	Intron2
Position	56044885	+124	+333
Ref. Gen.	CGG	A	TC
Alt._A	C	AC	T
Alt._B	CGG	A	TC
Color_A	CGG	A	TC
Color_B	CGG(C)	AC(A)	TC(T)
Geo._A	C	AC	T
Geo._B	CGG(C)	AC(A)	TC(T)
Geo._C	CGG(C)	AC(A)	TC(T)
Capra	CGG	—	—
Pantholops	—	—	—
Bos taurus	—	—	—

C. PAT1 (72792299..72795221)

	5'UTR	Intron1	Intron4	3'Flank
Position	72792241	+82	+1492	+3109
Ref. Gen.	CA	TA	C	G
Alt._A	CA	TA	CAA	GT
Alt._B	C	TA	CAA	GT
Color_A	C	T	CAT	GT
Color_B	CA	TA(T)	CAA(C)	GT(G)
Geo._A	CA	TA	C	G
Geo._B	C	T	CAA	GT
Geo._C	CA	TA(T)	CAA(C)	GT(G)
Capra	—	—	—	—
Pantholops	C	—	—	GT
Bos taurus	—	—	—	—

D. ZNF711 (72993669..73020595)

	Intron3	Intron7
Position	72998290	+836
Ref. Gen.	GG	A
Alt._A	GG	A
Alt._B	G	ATTG
Color_A	G	ATTG
Color_B	GG(G)	ATTG(A)
Geo._A	G	ATTG
Geo._B	GG	A
Geo._C	GG(G)	ATTG(A)
Capra	—	—
Pantholops	—	—
Bos taurus	—	—

E. APOOL (72757514..72892083)

	Intron2	Intron8
Position	72813200	+973
Ref. Gen.	C	ATT
Alt._A	CG	A
Alt._B	C	A
Color_A	CG	ATT
Color_B	GG(C)	ATT(A)
Geo._A	C	ATTG
Geo._B	CG	A
Geo._C	CG(C)	ATT(A)
Capra	—	—
Pantholops	—	—
Bos taurus	—	—

F. POF1B (73024346..73123532)

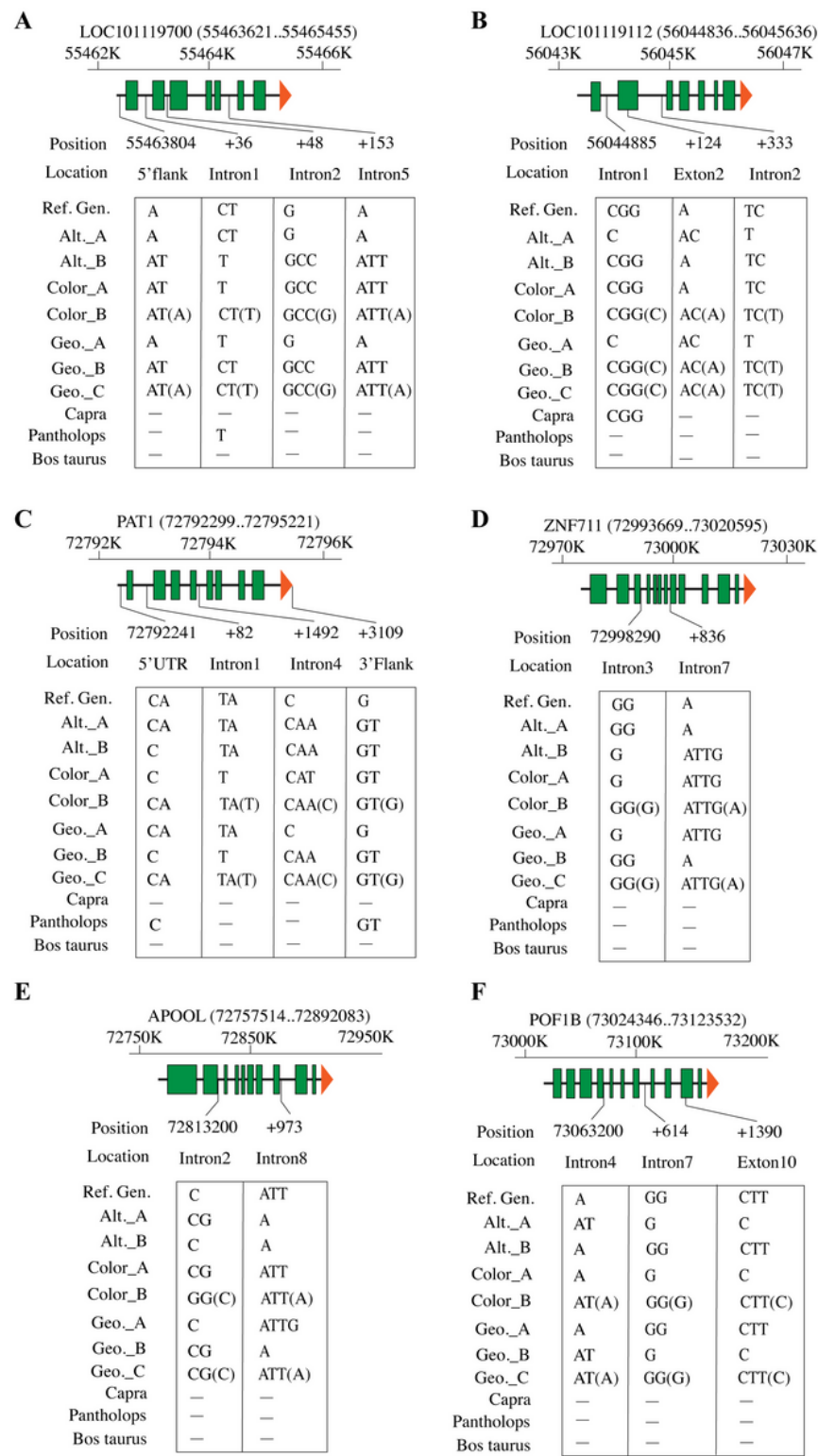|	Intron4	Intron7	Exton10
Position	73063200	+614	+1390
Ref. Gen.	A	GG	CTT
Alt._A	AT	G	C
Alt._B	A	GG	CTT
Color_A	A	G	C
Color_B	AT(A)	GG(G)	CTT(C)
Geo._A	A	GG	CTT
Geo._B	AT	G	C
Geo._C	AT(A)	GG(G)	CTT(C)
Capra	—	—	—
Pantholops	—	—	—
Bos taurus	—	—	—

Figure 3

Gene INDEL analysis for 4 candidate overlapped genes. A-D: INDEL analysis on KLF4, PSMA1, PAT1, and ZNF711 genes within clustered subpopulations in addition to three other species including Capra, Pantholops, and Bos taurus.
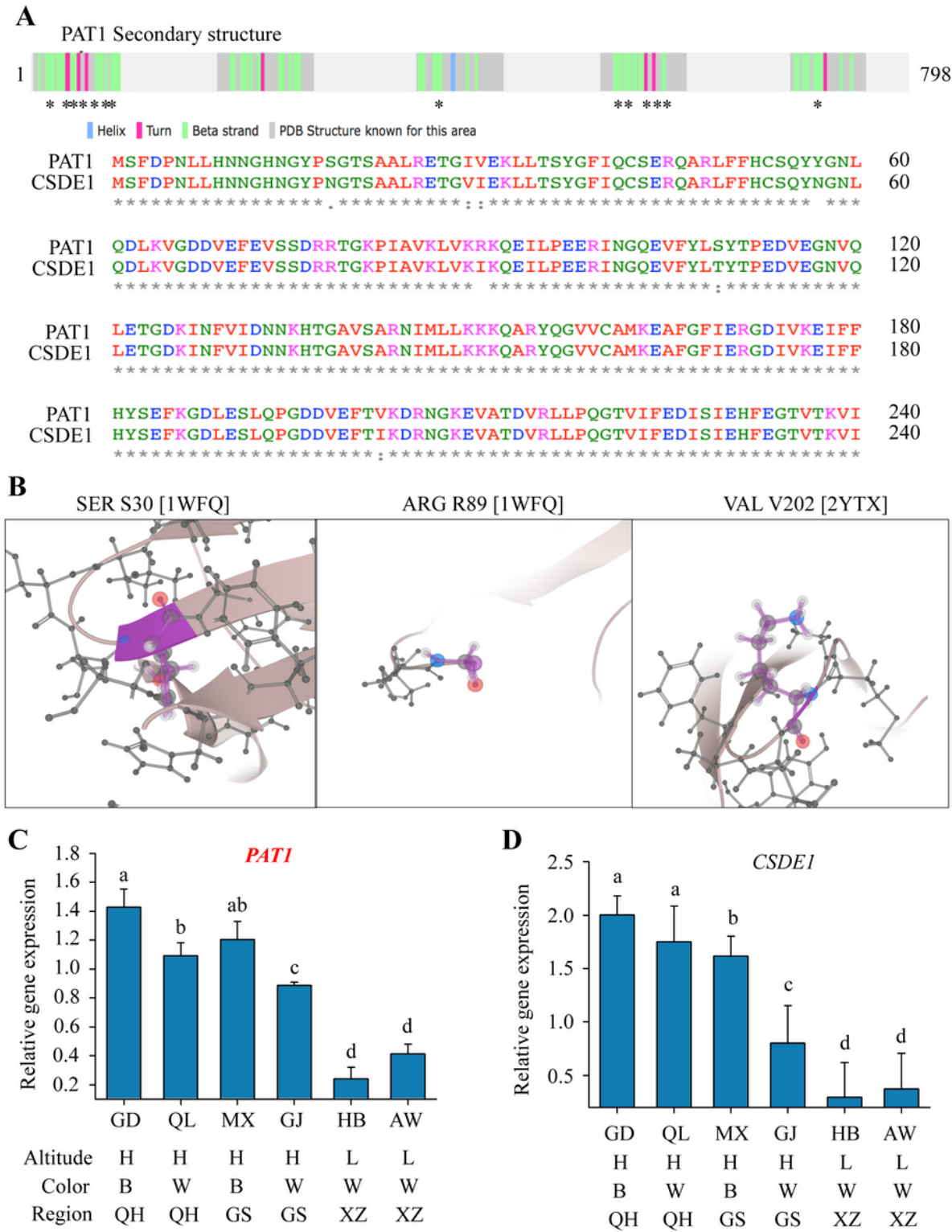
Figure 4

Comparison on sequence alignments and expression pattern between PAT1 and CSDE1. A: Secondary structure of PAT1 and differentiated amino acids between PAT1 and CSDE1. The location of 7 differentiated amino acids were depicted in symbol "*", which are majorly on beta-strand and turn. Secondary structure of protein is predicted via Uniprot database from different PDB entries. B: Differentiated location of amino acids on representative 3D protein structure; C-D: Comparison on

expression of CSDE1 and PAT1 in given 6 indigenous Tibetan sheep populations in this study. The 6 indigenous Tibetan sheep populations include Guide Black Fur sheep (GD), Qilian White Tibetan sheep (QL), Minxian Black Fur sheep (MX), Ganjia Tibetan sheep (GJ), Huoba Tibetan sheep (HB) and Awang Tibetan sheep (AW). Expression values were calculated against GAPDH using mRNA in lungs, and four biological replicates were conducted.

Figure 5

Comparison on morphological and physiological traits in Tibetan sheep populations with different haplotype promoters of PAT1 gene. A: Representative picture of indigenous Tibetan sheep (Huoba Tibetan sheep, HB for C-type, and Guide Black Fur sheep, GD for CA type) originated between two haplotypes promoter of PAT1; B: Comparison on different physiological parameters of 15 Tibetan sheep between CA and C haplotype of PAT1. CA type include GJ and QL, while C type include AW and HB. The averaged values for each parameter was derived from at least 10 biological replicates.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFigures.docx
- SupplementaryTables.xls