BMC Genomics

RESEARCH

Open Access

Genetic analysis of cold tolerance and high-altitude adaptation in Gannan indigenous Tibetan sheep through genome-wide scans

Chenxiao Lv[1,2,3†], Lili Chen[1†], Pan Zou[1], Yan Chen[2*] and Yi Ma[1*]

Abstract

Background Understanding the genetic basis of adaptation to extreme environments is crucial in evolutionary biology. Gannan Tibetan sheep, which thrive in high altitudes with thin air and cold climates, demonstrate remarkable resilience. However, the genetic basis of their adaptations remains poorly understood. In this study, we conducted whole-genome resequencing of three indigenous Tibetan sheep breeds, including Ganjia ($n=12$), Oula ($n=12$), Kecai ($n=12$), from the Gannan Tibetan Autonomous Prefecture (average altitude of 3,500 m) in China. By integrating our data with previously published genomic data of sheep from northern China (Altitude below 1,500 m), we examined the genetic population structure, ancestry components, and signatures of positive selection.

Results Our analyses, including principal component analysis, neighbor-joining tree construction, and ADMIXTURE, revealed that the three indigenous Tibetan sheep breeds share a common ancestry with other Tibetan sheep breeds, yet exhibit distinct population differentiation. Moreover, all three breeds display high genetic diversity, which enhances their adaptation to harsh local environments. Through selective sweep analysis, we identified genes associated with cold tolerance and high-altitude adaptation. These genes are involved in mitochondrial function and oxidative stress (*LONP1*, *NDUFA11*, *NDUFB4* and *HSD11B1L*), cardiorespiratory and vascular function (*FECH*, *TNFAIP3)*, and fat metabolism and thermogenesis (*ACSL6*, *PDGFD*, *CES3*, *MTCH2*, *TMEM8B*, *SPAG8*, *ETAA1*, *ZIC1*, *ATG5*, and *HAO2)*. Additionally, we identified candidate regions containing genes related to immunity and physical traits.

Conclusion This study provides a comprehensive overview of the genomic variations, ancestry compositions, and selective signals linked to high-altitude traits in Gannan indigenous Tibetan sheep. Our findings offer valuable insights into the breeding and management of Tibetan sheep.

Keywords Cold climate, Ecological adaptation, Genetic basis, Plateau, Whole-genome resequencing

†Chenxiao Lv and Lili Chen contributed equally to this work.

*Correspondence:
Yan Chen
chenyan@caas.cn
Yi Ma
nrtd321@163.com

[1]Tianjin Key Laboratory of Animal Molecular Breeding and Biotechnology, Tianjin Engineering Research Center of Animal Healthy Farming, Institute of Animal Science and Veterinary, Tianjin Academy of Agricultural Sciences, Tianjin 300381, People's Republic of China
[2]State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China
[3]College of Life Sciences, Nankai University, Tianjin 300000, China

Background

Sheep (*Ovis aries*) were domesticated approximately 11,000 years ago in the Fertile Crescent [1], playing a vital role in human sustenance and culture. Traditional sheep breeding has relied on a combination of limited artificial selection, strong natural selection, and thousands of years of domestication, leading to adaptation across various Agro-climatic conditions [2]. Ecological adaptation is crucial for sheep to thrive amid environmental changes, with different agricultural and geographic environments shaping diverse sheep breeds. Understanding the genetics responsible for this adaptation is of significant economic interest, particularly in the context of climate change [3].

China boasts a rich diversity of native sheep breeds, shaped by historical movements of nomadic societies. These native sheep are classified into three ancestral populations based on their geographical distribution and genetic origins: Mongolian sheep, Kazakh sheep, and Tibetan sheep [4]. Tibetan sheep are the most common and widespread domesticated animals on the Qinghai–Tibetan Plateau, having spread with human migrations [5]. Xiahe County, located at the northeastern edge of the Tibetan Plateau and the Loess Plateau in southwestern Gansu Province, ranges in altitude from 3,000 to 4,000 m above sea level and experiences a continental plateau climate characterized by high altitude, low temperatures, strong ultraviolet radiation, and a short pasture growing season. Under long-term natural selection, the Tibetan sheep in this region have developed unique biological traits, allowing them to adapt to the alpine climate and year-round grazing. These sheep are indispensable to the local pastoral herders, providing clothing, food, shelter, and transportation. Sheep evolved in high-altitude environments have undergone genomic adaptations to cope with the unique challenges posed by high elevations. These adaptations enable them to thrive in low oxygen, low temperature, and high ultraviolet radiation conditions [6].

The study of selection signals has played a key role in revealing the genetic basis of high-altitude adaptation. In the analysis of genome-wide nucleotide polymorphisms in Tibetan sheep, genes involved in hypoxia and ultraviolet signaling pathways, such as the HIF-1 pathway and the *HBB* and *MITF* genes, were detected as strong selective signatures [7]. *FGF10*, *MMP14*, *SLC25A51*, *NDUFB8*, *ALAS1*, *PRMT1*, *PRMT5*, and *HIF1AN* are genes associated with hypoxic adaptation at ultra-high altitude, while *HMOX2*, *SEMA4G*, *SLC16A2*, *SLC22A17*, and *BCL2L2* are associated with hypoxic adaptation at high altitude [8]. Recent studies have shown that some structural variants in Tibetan sheep have important cis-role regulatory effects on gene expression (e.g., *EPAS1*, *PAPSS2*, and *PTPRD*), and that these affected genes are associated with physiological responses to hypoxia and may promote high-altitude adaptation in Tibetan sheep on relatively short evolutionary time scales [9]. In this study, we performed whole-genome resequencing of three native Tibetan sheep breeds, including Ganjia, Oula, and Kecai, from Xiahe County. By integrating our data with previously published genomic data of sheep from northern China, we aims to decipher the genetic mechanisms underlying high-altitude adaptation and explore other undiscovered genetic traits in hypoxic emergency and cold-adapted environments. The findings provide a theoretical basis for sheep breed improvement.

Methods

Samples collection and genome DNA sequencing

We collected blood samples from 36 male Tibetan sheep from local herders in Xiahe County, Gannan Tibetan Autonomous Prefecture, Gansu Province, China (35°11′54.2″ N, 102°31′10.7″ E). The blood samples were obtained via tail vein puncture, without euthanasia or anesthesia. The samples included 12 Ganjia sheep (GJ), 12 Oula sheep (OL), and 12 Kecai sheep (KC). DNA was extracted from the blood of each individual using Automated Nucleic Acid Extraction (CWE9600 Automated Nucleic Acid Extraction, ComWin Biotech Co.,Ltd.). The extracted DNA was detected by Nanodrop for OD values (260/280 = 1.8–2.0, 260/230 > 1.5), and verified by 1% agarose gel electrophoresis to ensure no degradation, meeting sequencing requirements. The DNA samples were sequenced using BGISEQ technology by Tianjin Compass Inspection and Testing Co. Ltd. Additionally, we included whole-genome resequencing data of 75 sheep from northern China, including 10 Bashbay sheep (BSB), 4 Lanzhou Large tailed sheep (LLT), 6 Large tailed Han sheep (LTH), 10 Prairie Tibetan sheep (PRT), 10 Sishui Fur sheep (SSS), 13 Tan sheep (Tan), 5 Qinghai Tibetan sheep (TIB), 5 Minxian Black Fur sheep (MXB), and 12 Argali sheep (AGL). The control groups are all from low-altitude areas (< 1500 m), such as Bashbay sheep (Xinjiang, altitude 800–1200 m) and Lanzhou Large-tailed sheep (Lanzhou, Gansu, altitude 1500 m), ensuring that the selection signal is specific to high-altitude adaptation.

Sequence alignment and variant calling

Raw sequencing reads were processed using fastp version 0.24.1 to trim low-quality bases and sequences. Clean reads were then aligned to the sheep reference genome (ARS-UI_Ramb_v2.0) using the Burrows-Wheeler Aligner v0.7.17 (BWA) algorithm [10]. The final average sequencing coverage was 14.94× per individual (ranging from 9.78× to 35.00×). Reads were sorted and merged, and duplicates were removed using Genome Analysis Toolkit 4 (GATK4) MarkDuplicates [11]. Variant calling was performed using GATK4. Single nucleotide polymorphisms (SNPs) were filtered based on the following

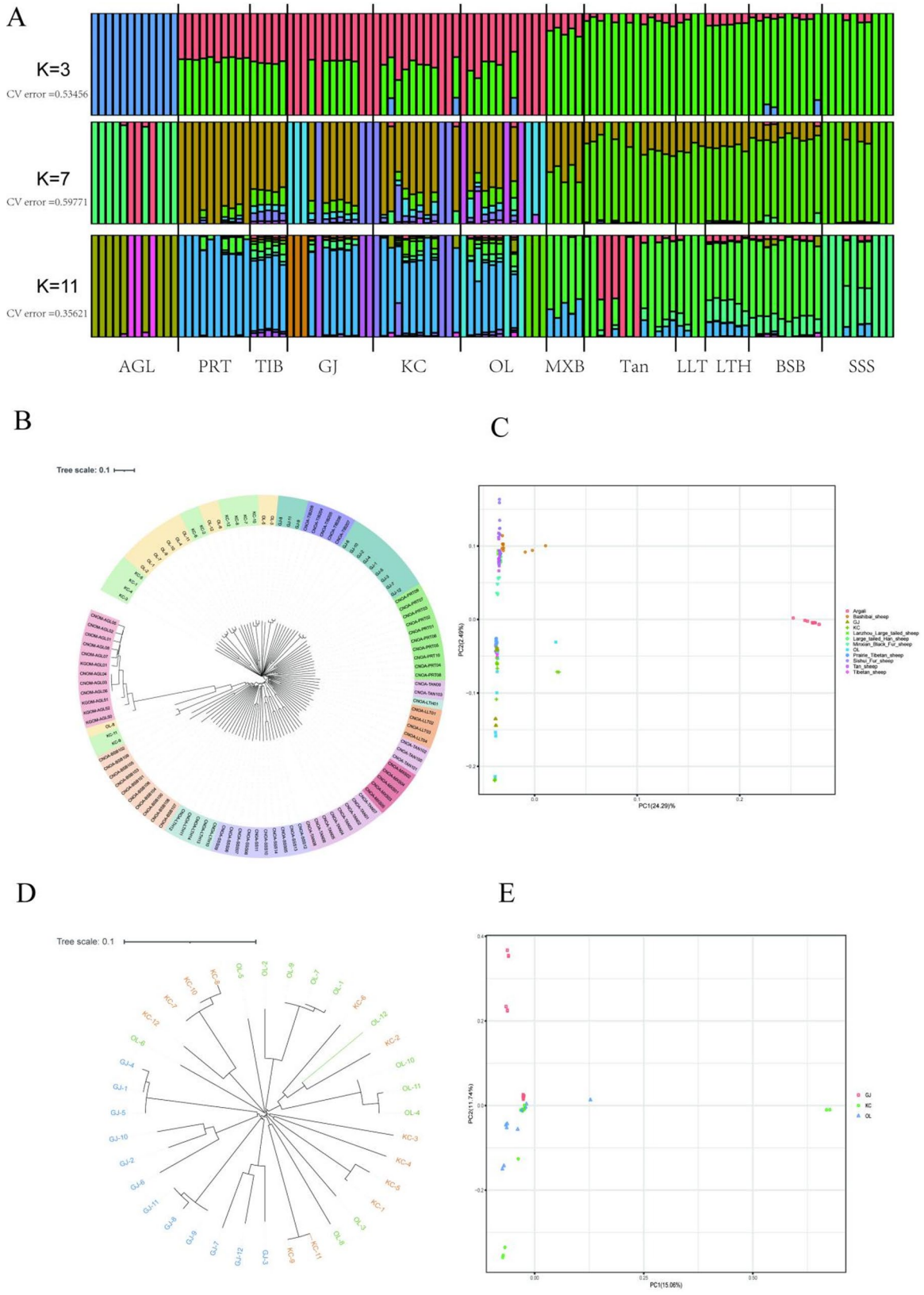

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Population structure and genetic relationships among the three local Tibetan sheep breeds (GJ, OL, and KC) and other nearby breeds. Other breeds include Argali (AGL), Prairie Tibetan (PRT), Xinjiang Tibetan (TIB), Minxian Black Fur (MXB), Tan (Tan), Lanzhou Large tailed (LLT), Large-tailed Han (LTH), Bashibai (BSB), and Sishui Fur (SSS) sheep. **A** ADMIXTURE analysis results for $K = 3$, 7, or 11. Different colors represent different ancestral lineages. **B** Neighbor-joining (NJ) tree for all individuals. **C** Principal component analysis (PCA) results for all breeds. **D** NJ tree for the three local breeds (GJ, OL, and KC). **E** PCA results for the three local breeds

criteria: (1) mean sequencing depth < 3×; (2) mapping quality (MQ) > 25; (3) Quality by Depth (QD) > 2.0; (4) Fisher Strand (FS) > 60.0; (5) MQRankSum > − 12.5; (6) ReadPosRankSum > − 8.0.

Further filtering was conducted using VCFtools-0.1.16 [12] to remove indels and PLINK 1.9 [13] for SNP site filtering. The indels were removed following this parameter setting: max-missing > 0.95; MAF > = 0.05; recode-INFO-all. A total of 128,986,177 SNPs were initially detected. After filtering by the above criteria, 23,266,087 high-quality SNPs were retained, with a filtering efficiency of approximately 18%, ensuring the accuracy of subsequent analysis. ANNOVAR 2.0 [14]was used in combination with the Ensembl 105 version sheep genome database (ARS-UI_Ramb_v2.0) for SNP functional annotation.

Population structure analysis

Population structure analysis was performed using filtered SNP sets. ADMIXTURE v1.3.0 [15] was used for structure analysis, with the number of ancestral clusters (K) set from 1 to 13 and five-fold cross-validation to determine the K values with minimum errors. Principal Component Analysis (PCA) was conducted using PLINK 1.9, identifying the first four principal components, with the first two used to distinguish population structure. A phylogenetic tree was constructed using rapidnj-2.3.2 software according to the neighbor-joining method.

Genetic diversity and Linkage Disequilibrium (LD) detection

Nucleotide diversity was calculated for each group with a window size of 100 kb and step size of 10 kb using VCFtools. LD decay was calculated via the squared correlation coefficient (r^2) using PopLDdecay − 3.42 [16].

Genome-wide selective sweep analysis

To identify genomic regions under selective sweeps associated with adaptation, we used multiple methods to investigate selection signatures, including the population differentiation index (F_{st}) [17], nucleotide diversity (π) ratio analysis [18] and XP-EHH (Cross Population Extended Haplotype Homozygosity) [19]. The π ratio was calculated as π(A)/π(B), where π(A) and π(B) are the π values of the control and selection groups, respectively. These three methods are often used in positive natural selection analyses. Statistics were calculated using a sliding window approach with a window size of 100 kb and a step size of 10 kb. We calculated the average *F*st, π-Ratio, and XPEHH values of SNPs in each window and used the outlier method to obtain windows with the top 1% of *F*st, π-Ratio, and XPEHH values. Protein-coding genes in these outlier windows were annotated using ANNOVAR 2.0 and treated as positively selected genes (PSGs).

Functional enrichment analysis of candidate genes

Functional enrichment analysis of candidate genes was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) method by using the KOBAS v3.0 tool [20]. Candidate genes were mapped to KEGG terms, and the number of significant genes for each term was determined with a P-value ≤ 0.05 as the threshold. Terms meeting this criterion were defined as significantly enriched for candidate genes.

Results

Whole-genome sequencing and variant detection

Whole-genome resequencing of 36 Tibetan sheep from Xiahe County yielded sequencing reads with an average depth of 14.94×. For comparative purposes, we combined our dataset with whole-genome resequencing data of 75 sheep (Supplementary Table 1) across northern China, including BSB, LLT, LLT, PRT, Tan, TIB, MXB, SSS, and AGL sheep, which served as phylogenetic outgroups. After quality control, a total of 23,266,087 clean SNPs in the genomes of the 36 native Tibetan sheep were identified. Many of these SNPs were in intergenic (14.797%) and intronic regions (36.791%), while only 0.605% (57186 SNPs) were found in exonic regions.

Population structure analysis

To understand the genetic relationships among the three local Tibetan sheep breeds in Xiahe County and between them and other sheep breeds, we analyzed their population structure, phylogenetic relationships, and data principal components. To assess the proportion and degree of shared genetic ancestry, we used ADMIXTURE with K values ranging from 1 to 13, including the wild Argali sheep (AGL) for reference. At K = 3, AGL, Tibetan sheep breeds, and other breeds were differentiated into distinct ancestral groups. When K = 11, the cross-validation error (CV = 0.35) is the lowest, indicating the optimal number of ancestral clusters. GJ, OL and KC are largely similar to those of Tibetan sheep. Among these, the OL group exhibits partial Northern China sheep ancestry, whilst within the GJ and KC populations a minority of individuals display distinctive characteristics.(Fig. 1A). The

first and second PCs explained 24.29% and 2.49% of the genetic variation, respectively (Fig. 1C). The first eigenvector distinctly separated five Tibetan sheep breeds from other breeds, while the second eigenvector distinguished GJ from OL and KC. Further PCA analysis of the three local breeds indicated that six GJ individuals and seven KC individuals clustered closely with OL (Fig. 1E), suggesting close genetic relationships possibly due to geographic proximity or intentional interbreeding. Despite being classified under the Tibetan sheep category, GJ and OL exhibit distinct morphological traits: GJ is small and wool-oriented, while OL is large and meat-oriented.

Further examination of genetic relationships was conducted through NJ tree analysis using the filtered SNP set The NJ tree segregated the three local breeds into distinct genetic clusters, corroborating their genetic differentiation (Fig. 1D). Consistent with PCA results, OL clustered closely with KC, confirming their genetic proximity. The broader NJ tree for all breeds showed KC, GJ, and OL clustering with Tibetan sheep, with an unexpected placement of one OL individual and two KC individuals between AGL and BSB (Fig. 1B).

Genomic variation, genetic diversity, and Linkage Disequilibrium (LD) analysis

To investigate the genomic variation among different sheep breeds, we calculated nucleotide diversity for each breed. The results indicated that BSB sheep exhibited the highest nucleotide diversity, followed closely by TIB and OL (Fig. 2A). To further understand the population genetics and demographic dynamics of each breed, we analyzed genome-wide patterns of LD using PopLDdecay software with default parameters ((Fig. 2B). The analysis revealed notable differences among the breeds. For marker distances greater than 50 kb, Tan sheep showed the lowest LD levels, indicating greater genetic recombination, whereas Lanzhou Large-tailed sheep (LLT) exhibited the highest LD levels, followed by TIB and MXB. LLT is a highly selected meat breed. Long-term artificial selection has reduced the effective population size (Ne), resulting in slow LD decay (distance > 100 kb when r^2 = 0.3), consistent with the results of nucleotide diversity.

Genome-wide selection signature analysis

In domesticated species, the study of phenotypic diversity across varieties or species can be complemented by selection sweep analysis to gain insight into the biological mechanisms underlying phenotypic variation. Population evolutionary selection sweep analysis refers to the reduction of nucleotide polymorphisms in certain genome regions due to selective pressures, leaving a signature of positive selection on the species' genome [21]. Compared to control populations, the genetic diversity in these regions is significantly reduced in the selected populations, which is characteristic of domesticated regions. To understand the genetic bases of phenotypic traits, production traits, and adaptive traits, we analyzed three native Tibetan sheep breeds, considering them as the selected groups and other breeds as the control groups. We employed three methods, including F_{st}, π ratio, and XP-EHH, to identify selection signatures. The combined results from these methods revealed strong selection signatures, aiding in the identification of target genes.

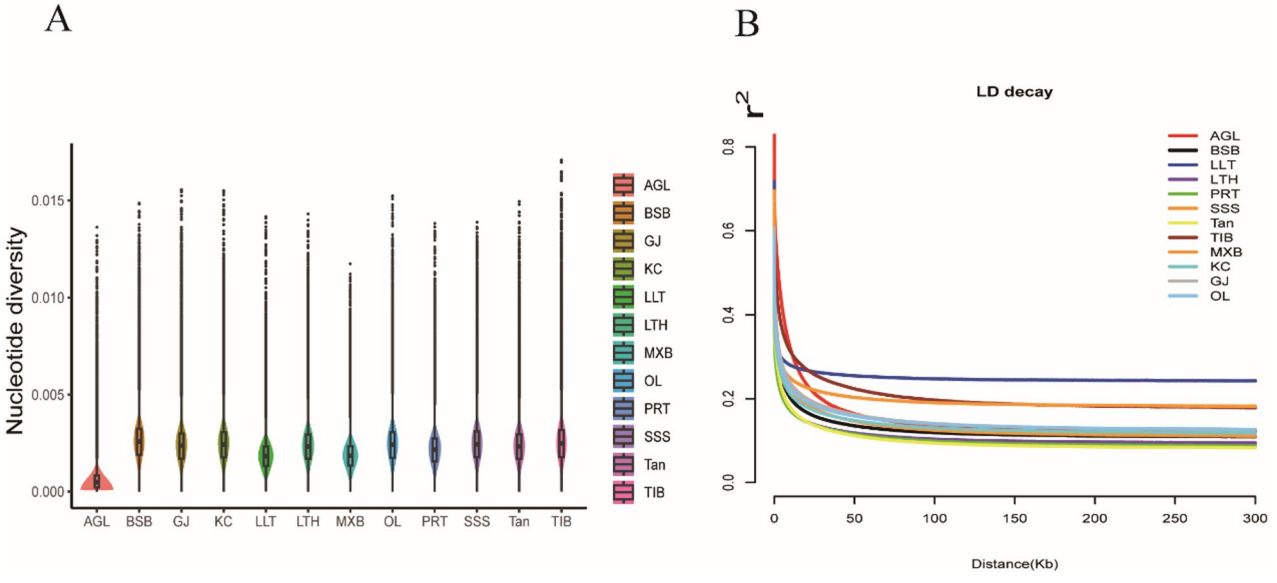

Fig. 2 Nucleotide diversity and genome-wide average linkage disequilibrium (LD) decay estimated for each breed. **A** Nucleotide diversity of all breeds. The median is represented by the black line in the boxplot, with outliers shown as points outside the box. **B** Genome-wide average linkage disequilibrium (LD) decay estimated for each breed, different colors represent different breeds

From the analysis using F_{st}, π ratio, and XP-EHH, significant regions were identified by top 1% cutoff (Fig. 3A), and genes within the overlapping regions of the three methods were considered candidate genes. We identified 75, 83, and 93 candidate genes in OL, KC, and GJ breeds (Fig. 3B), respectively. These genes were predominantly located on chromosomes 5, 13, and, 22, with a notable clustering on chromosome 5, including genes such as *RFX2*, *RANBP3*, *NUFA11*, *NRTN*, *LONP1*, and *HSD11B1L* (Fig. 4). These chromosomes are precisely enriched with high-altitude adaptation candidate genes (such as LONP1 on chromosome 5), suggesting that the selection pressure has not significantly reduced the diversity of key chromosomes. In total, 19 genes were identified across the candidate regions of the three breeds (Fig. 3B). We conducted KEGG and GO pathway analysis on these candidate genes for each breed using individual selection methods (Fig. 3C, S1).

We identified genes enriched in the top 20 KEGG pathways with distinct enrichment scores as candidate genes. The enrichment analysis revealed significant pathways for KC related to melanogenesis, sphingolipid metabolism, thermogenesis, TNF signaling, and relaxin signaling pathways. For OL, significant pathways included melanoma, fatty acid biosynthesis, glycosaminoglycan biosynthesis (chondroitin sulfate / dermatan sulfate), and RNA transport. In GJ, significant pathways were linked to glyoxylate and dicarboxylate metabolism, ferroptosis, peroxisome, carbon metabolism, fatty acid biosynthesis, and Hippo signaling pathway.

Discussion

To investigate potential adaptation mechanisms to high altitudes, low temperatures, and ultraviolet radiation in Gannan indigenous Tibetan sheep, we searched for SNPs showing evidence of specific selection between these sheep and lower-altitude breeds using F_{ST}, π ratio, and XP-EHH methods. Genome-wide polymorphisms revealed chromosomal regions containing 75, 83, and 93 genes with evidence of positive selection in OL, GJ, and KC breeds, respectively. Nineteen candidate genes overlapped among the three breeds. Among these, three genes were associated with high-altitude adaptation: *LONP1*, *NRTN* and *ACSL6*.

Prolonged exposure to high-altitude environments may lead to hypoxia and trigger serious health issues. Lon peptidase 1 (LONP1), a member of the AAA + family, is essential for maintaining mitochondrial function, can interacting with mitochondrial DNA (mtDNA), and playing roles in oxidative stress [22]. *NRTN* can increasing capillary density and oxidative capacity favoring fatty acid metabolism [23], *ACSL6* is the key gene regulating fat metabolism in sheep [24]. In studies of key biological pathways crucial for goats' environmental adaptation, GO terms related to energy metabolism reactions were identified as critical for local adaptation [25]. It is suggested that Tibetan sheep may adapt to the hypoxia and low temperatures of the plateau environment by regulating their energy metabolism. Additionally, *GPR156* is a critical regulator of hair cell orientation and might affect hearing by reversing hair cell orientation [26]. This may relate to the adaptation of Tibetan sheep to natural grazing. In addition, there are six candidate genes (*DUS3L*, *PRR22*, *RANBP3*, *CATSPERD*, *RXFP2 and ACSL6*) on chromosome 5 and they are in close proximity, although no studies have shown that they are directly related with plateau adaptation in sheep, since we have detected them in all three breeds, there is reason to believe that the variations located in these genes affect the ability of sheep to adapt to plateaus.

Other candidate genes showing selection signals did not overlap among the three native breeds, though they were inferred to be associated with high-altitude acclimatization. Acute high-altitude exposure causes hypobaric hypoxia and increases oxidative stress, raising the risk of high-altitude illnesses (HAIs) in poorly acclimatized individuals. Conversely, acclimatization may trigger adaptive responses to oxidative stress to restore redox homeostasis. Mitochondria play a crucial role in oxidative stress adjustments, through ROS-signaling and cross-talk with HIFs. Among our candidate genes, *NDUFA11* encodes a subunit of the mitochondrial complex I, essential for the assembly and stability of the respirasome. Downregulation of *NDUFA11* reduces ATP production and increases mitochondrial ROS production [27]. Similarly, *NDUFB4* encodes a non-catalytic subunit of complex I, the first enzyme in the mitochondrial electron transport chain. Knockdown of *NDUFB4* decreases complex I activity and mitochondria-derived ATP and affects ROS levels (Bennett NK, Lee M, Orr AL, Nakamura K, 2024). Additionally, genes, such as *NNT* [28], *PRUNE1* [29], *SHMT1* [30], *SLC5A3* [24], *SLC2A12* [31], and *FAR1* [32], play significant roles in oxidative stress. *HSD11B1L* is believed to predominantly act in vivo as an oxidoreductase, using NADP(H) as a cofactor to generate cortisone [33].

High-altitude hypoxic environments increase plasma volume and heart rate [34], potentially leading to higher blood pressure. Mammals living at high altitudes avoid pulmonary hypertension through enhanced vasodilatory and contractile capacity, larger vascular volume, and increased elastic fiber and vascular abundance in the lungs. *FECH*, the terminal enzyme in heme biosynthesis, is essential for angiogenesis both in vitro and in vivo [35]. *TNFAIP3* facilitates *FGFR1* activation-induced angiogenesis by promoting VEGFA expression and secretion [36]. The homeobox gene *MEIS1* modulates cardiovascular regeneration (Paul S, Zhang X, He JQ, 2020).

Fig. 3 Results of genetic scan for OL, KC, and GJ breeds. **A** Results of F_{st}, XP-EHH, and π-Ratio analyses for OL, KC, and GJ breeds. **B** Number of selected genes identified in OL, GJ, and KC breeds by the three methods, shown in each Venn diagram component. **C** Results of KEGG enrichment analysis for OL, GJ, and KC breeds

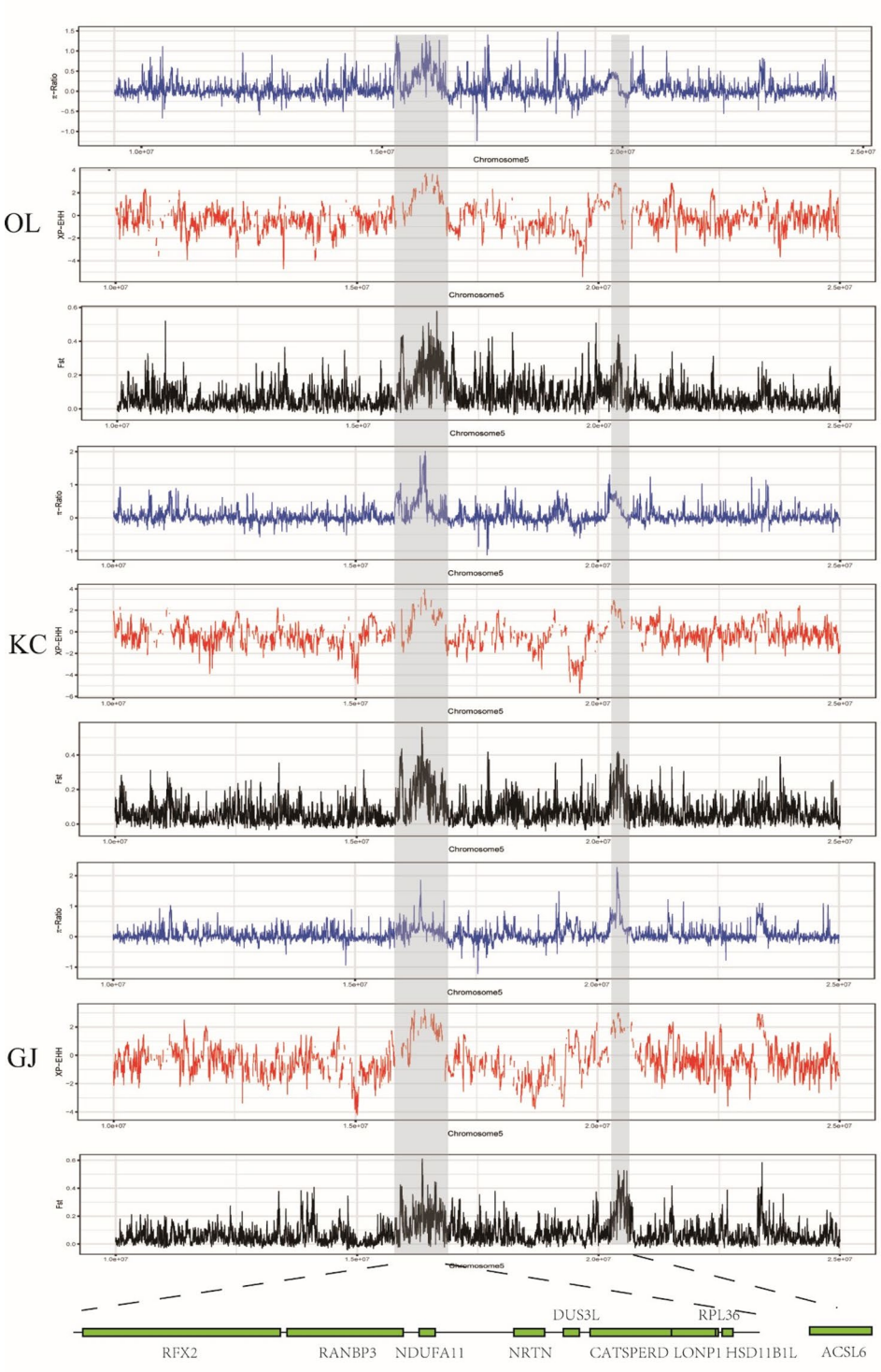

Fig. 4 Genome-wide selection scan in OL, GJ, and KC using sliding window analysis (4 kb window size and 2 kb step size). The selected region on chromosome 5 ranges from 1.5×10^7 to 2.0×10^7 bp and includes genes such as RFX2 (1.52×10^7 bp) and LONP1 (1.85×10^7 bp), facilitating subsequent verification

Additionally, genes like *AZIN2* [37], *ZIC4* [38], and *TLN1* [39] are related to angiogenesis and vascular function. *SEC31B* has been found to be significantly associated with resting heart rate [40]. *TMEM69* [41] and *HINT2* [42] are the key genes expressed in different stages of spinal cord ischemia. *AK2* is integral to embryo and adult heart mesenchymal stem cells via the activation of Wnt/β-catenin signaling in pulmonary fibrogenesis [43].

In addition to the previously mentioned aspects, we identified several genes associated with lipid metabolism

and thermogenesis, including *PDGFD* [44], *CES3* [45], *MTCH2* [46], *TMEM8B* [47], *SPAG8* [48], *ETAA1* [49], *ZIC1*, *ATG5* [50], and *HAO2* [51]. These genes are likely crucial for the growth and cold tolerance of Gannan indigenous Tibetan sheep. The selection signal for PDGFD (promoting fat deposition) in OL (meat type) is significantly stronger than in GJ (wool type), while the XP-EHH value for HSD11B1L (regulating wool growth) is higher in GJ, indicating that the selection signals are consistent with the phenotypic differentiation of the breeds uses.

Moreover, strong disease resistance is a notable trait of local breeds. Compared to other sheep breeds, the three indigenous Tibetan sheep breeds exhibit enhanced disease resistance. Many of the candidate genes identified in this study are related to immunity, such as *StarD13* [52], *SAP130* [53], and *LIMA1* [54]. Additionally, we identified candidate regions containing *RXFP2* and *MITF*, which are associated with horn shape [55] and coat color [56].

Conclusions

Analyzing genetic variability within breeds or populations offers valuable insights for designing breeding improvement strategies for indigenous sheep genetic resources. Typically, highly selected and bred sheep exhibit relatively low genetic diversity. In this study, the nucleotide diversity of the three indigenous breeds was comparable to that of TIB and higher than that of MXB and LLT. This is consistent with the fast rate and low level of LD decay observed in these three breeds, suggesting they may not have undergone systematic breeding.

The PCA, NJ phylogenetic trees, and admixture analyses revealed a common clustering pattern, indicating that the three native Tibetan sheep breeds are closely related to TIB and PRT. They may have been crossbred to varying degrees with other nearby breeds. The results show genetic differentiation among GJ, OL, and KC breeds. In addition, KC's distribution is more dispersed than the other two breeds, reflecting greater genetic variation among KC individuals. Overall, these three breeds exhibit high genetic diversity, facilitating better adaption to harsh local environments. Through selective sweep analysis, we identified genes associated with cold tolerance and high-altitude adaptation. These genes are involved in mitochondrial function and oxidative stress (*LONP1*, *NDUFA11*, *NDUFB4* and *HSD11B1L*), cardiorespiratory and vascular function (*FECH*, *TNFAIP3)*, and fat metabolism and thermogenesis (*ACSL6*, *PDGFD*, *CES3*, *MTCH2*, *TMEM8B*, *SPAG8*, *ETAA1*, *ZIC1*, *ATG5*, and *HAO2)*.

Our study provided new insights into the diversity and selection signals of Gannan indigenous Tibetan sheep and their relationships with other breeds. Discovering genomic diversity forms a basis for conserving and utilizing Tibetan sheep genetic resources. We identified several genes potentially crucial for high-altitude adaptation, lipid metabolism, and immune response. These findings offer valuable information for further research on the formation of fine traits in Tibetan sheep and provide references for the molecular breeding of other varieties.

Abbreviations

GJ	Ganjia sheep
OL	Oula sheep
KC	Kecai sheep
BSB	Bashbay sheep
LLT	Lanzhou Large tailed sheep
LTH	Large tailed Han sheep
PRT	Prairie Tibetan sheep
SSS	Sishui Fur sheep
Tan	Tan sheep
TIB	Qinghai Tibetan sheep
MXB	Minxian Black Fur sheep
AGL	Argali sheep

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12722-x.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

Not applicable.

Authors' contributions

Y.M. and Y.C. conceived and designed the study. C.X.L. and L.L.C. performed the data analysis and wrote the manuscript. P.Z. edited the manuscript. All authors reviewed and approved the manuscript.

Funding

Research on livestock and poultry germplasm resource innovation and efficient breeding technology(2024ZYCX012); Xiahe County yak Tibetan sheep efficient breeding improvement project (202301).

Data availability

The datasets generated and/or analysed during the current study are available in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2024), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number PRJCA035485.

Declarations

Ethics approval and consent to participate

This study followed the recommendations of the "Regulations for the Management of Affairs Concerning Experimental Animals" (Ministry of Science and Technology, China, revised in June 2004). All experimental procedures were in accordance with the animal welfare legislation and approved by the ethics committees of all the participating institutes (Institute of Animal Science and Veterinary, Tianjin Academy of Agricultural Sciences; Institute of Animal Sciences of CAAS; Tianjin Key Laboratory of Animal Molecular Breeding and Biotechnology; Tianjin Engineering Research Center of Animal Healthy Farming; College of Life Sciences, Nankai University). Blood samples were collected by trained personnel and specialized veterinarians in accordance with the Tianjin Academy of Agricultural Sciences Animal Welfare Protocol. All methods were performed in accordance with the relevant guidelines and regulations. Informed consent was obtained from all animal owners. The study is reported in accordance with ARRIVE guidelines.

The ethical approval numbers are: Tianjin Academy of Agricultural Sciences Institute of Animal Science (TAAAS-IAV-2022-003), Institute of Animal Science, Chinese Academy of Agricultural Sciences (IAS-CAAS-2022-011). All authors have been informed and consent to participate.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 January 2025 / Accepted: 2 March 2026
Published online: 27 March 2026

References

1. Chessa B, Pereira F, Arnaud F, Amorim A, Goyache F, Mainland I, et al. Revealing the history of sheep domestication using retrovirus integrations. Science. 2009;324(5926):532–6.
2. Alberto FJ, Boyer F, Orozco-terWengel P, Streeter I, Servin B, de Villemereuil P, et al. Convergent genomic signatures of domestication in sheep and goats. Nat Commun. 2018;9(1):813–9.
3. Zhang Y, Xue X, Liu Y, Abied A, Ding Y, Zhao S, et al. Genome-wide comparative analyses reveal selection signatures underlying adaptation and production in tibetan and poll dorset sheep. Sci Rep-Uk. 2021;11(1):2466.
4. Zhao E, Yu Q, Zhang N, Kong D, Zhao Y. Mitochondrial DNA diversity and the origin of chinese indigenous sheep. Trop Anim Health Pro. 2013;45(8):1715–22.
5. Liu J, Ding X, Zeng Y, Yue Y, Guo X, Guo T, et al. Genetic diversity and phylogenetic evolution of tibetan sheep based on mtDNA d-loop sequences. PLoS ONE. 2016;11(7):e0159308.
6. Tiwari M, Gujar G, Shashank CG, Ponsuksili S. Selection signatures for high altitude adaptation in livestock: a review. Gene. 2024;927:148757.
7. Hu X, Yang J, Xie X, Lv F, Cao Y, Li W, et al. The genome landscape of tibetan sheep reveals adaptive introgression from argali and the history of early human settlements on the qinghai-tibetan plateau. Mol Biol Evol. 2019;36(2):283–303.
8. Zhang W, Yuan C, An X, Guo T, Wei C, Lu Z, et al. Genomic insights into tibetan sheep adaptation to different altitude environments. Int J Mol Sci. 2024;25:22.
9. Liang X, Duan Q, Li B, Wang Y, Bu Y, Zhang Y, et al. Genomic structural variation contributes to evolved changes in gene expression in high-altitude tibetan sheep. P Natl Acad Sci Usa. 2024;121(27):e2322291121.
10. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform. Bioinf. 2010;26(5):589–95.
11. Bathke J, Luhken G. OVarFlow: a resource optimized GATK 4 based Open source Variant calling workFlow. BMC Bioinformatics. 2021;22(1):402. https://doi.org/10.1186/s12859-021-04317-y.
12. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
13. 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. Am J Hum Genet. 2007;81(3):559–75.
14. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
15. Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. Genet Epidemiol. 2005;28(4):289–301.
16. Zhang C, Dong SS, Xu JY, He WM, Yang TL. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 2019;35(10):1786–8.
17. Weir BS, Cockerham CC. Estimating f-statistics for the analysis of population structure. Evolution. 1984;38(6):1358–70.
18. Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, et al. Genomic analyses provide insights into the history of tomato breeding. Nat Genet. 2014;46(11):1220–6.
19. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, et al. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007;449(7164):913–8.
20. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39:W316–22. (Web Server issue).
21. Nosrati M, Asadollahpour NH, Amiri GZ, Esmailizadeh A. Whole genome sequence analysis to detect signatures of positive selection for high fecundity in sheep. Reprod Domest Anim. 2019;54(2):358–64.
22. Fukuda R, Zhang H, Kim JW, Shimoda L, Dang CV, Semenza GL. HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. Cell. 2007;129(1):111–22.
23. Correia JC, Kelahmetoglu Y, Jannig PR, Schweingruber C, Shvaikovskaya D, Zhengye L, et al. Muscle-secreted neurturin couples myofiber oxidative metabolism and slow motor neuron identity. Cell Metab. 2021;33(11):2215–e22308.
24. Zhang B, Sun Z, Yu Z, Li H, Luo H, Wang B. Transcriptome and targeted metabolome analysis provide insights into bile acids' new roles and mechanisms on fat deposition and meat quality in lamb. Food Res Int. 2022;162(Pt A):111941.
25. Peng W, Zhang Y, Gao L, Shi W, Liu Z, Guo X, et al. Selection signatures and landscape genomics analysis to reveal climate adaptation of goat breeds. BMC Genomics. 2024;25(1):420. https://doi.org/10.1186/s12864-024-10334-x.
26. Ramzan M, Bozan N, Seyhan S, Zafeer MF, Ayral A, Duman D, et al. Novel GPR156 variants confirm its role in moderate sensorineural hearing loss. Sci Rep-Uk. 2023;13(1):17010.
27. Jang S, Javadov S. Elucidating the contribution of ETC complexes i and II to the respirasome formation in cardiac mitochondria. Sci Rep-Uk. 2018;8(1):17732. https://doi.org/10.1038/s41598-018-36040-9.
28. Rao K, Shen X, Pardue S, Krzywanski DM. Nicotinamide nucleotide transhydrogenase (NNT) regulates mitochondrial ROS and endothelial dysfunction in response to angiotensin II. Redox Biol. 2020;36:101650.
29. Xu J, Wang Y, Li P, Chen C, Jiang Z, Wang X, et al. PRUNE1 (located on chromosome 1q21.3) promotes multiple myeloma with 1q21 gain by enhancing the links between purine and mitochondrion. Brit J Haematol. 2023;203(4):599–613.
30. Dou C, Xu Q, Liu J, Wang Y, Zhou Z, Yao W, et al. SHMT1 inhibits the metastasis of HCC by repressing NOX1-mediated ROS production. J Exp Clin Canc Res. 2019;38(1):70.
31. Li M, Li L, Cheng X, Li L, Tu K. Hypoxia promotes the growth and metastasis of ovarian cancer cells by suppressing ferroptosis via upregulating SLC2a12. Exp Cell Res. 2023;433(2):113851.
32. Marsh KG, Arrieta A, Thuerauf DJ, Blackwood EA, MacDonnell L, Glembotski CC. The peroxisomal enzyme, FAR1, is induced during ER stress in an ATF6-dependent manner in cardiac myocytes. Am J Physiol-Heart C. 2021;320(5):H1813–21.
33. Tomlinson JW, Stewart PM. Cortisol metabolism and the role of 11β-hydroxysteroid dehydrogenase. Best Pract Res Cl En. 2001;15(1):61–78.
34. Stembridge M, Williams AM, Gasho C, Dawkins TG, Drane A, Villafuerte FC, et al. The overlooked significance of plasma volume for successful adaptation to high altitude in sherpa and andean natives. P Natl Acad Sci Usa. 2019;116(33):16177–9.
35. Basavarajappa HD, Sulaiman RS, Qi X, Shetty T, Sheik PBS, Sishtla KL, et al. Ferrochelatase is a therapeutic target for ocular neovascularization. Embo Mol Med. 2017;9(6):786–801.
36. Gao M, Li X, Yang M, Feng W, Lin Y, He T. TNFAIP3 mediates FGFR1 activation-induced breast cancer angiogenesis by promoting VEGFA expression and secretion. Clin Transl Oncol. 2022;24(12):2453–65.
37. Li X, Sun Y, Huang S, Chen Y, Chen X, Li M, et al. Inhibition of AZIN2-sv induces neovascularization and improves prognosis after myocardial infarction by blocking ubiquitin-dependent talin1 degradation and activating the akt pathway. Ebiomedicine. 2019;39:69–82.
38. He Q, Zhao L, Liu X, Zheng J, Liu Y, Liu L, et al. MOV10 binding circ-DICER1 regulates the angiogenesis of glioma via mir-103a-3p/mir-382-5p mediated ZIC4 expression change. J Exp Clin Canc Res. 2019;38(1):9.
39. Wang X, Mao W, Ma X. TLN1 synergizes with ITGA5 to ameliorate cardiac microvascular endothelial cell dysfunction. Folia Morphol. 2024;83(1):92–101.
40. van den Berg ME, Warren HR, Cabrera CP, Verweij N, Mifsud B, Haessler J, et al. Discovery of novel heart rate-associated loci using the exome chip. Hum Mol Genet. 2017;26(12):2346–63.
41. Li JA, Zan CF, Xia P, Zheng CJ, Qi ZP, Li CX, et al. Key genes expressed in different stages of spinal cord ischemia/reperfusion injury. Neural Regen Res. 2016;11(11):1824–9.
42. Li S, Chen J, Liu M, Chen Y, Wu Y, Li Q, et al. Protective effect of HINT2 on mitochondrial function via repressing MCU complex activation attenuates cardiac microvascular ischemia-reperfusion injury. Basic Res Cardiol. 2021;116(1):65.
43. Shi C, Chen X, Yin W, Sun Z, Hou J, Han X. Wnt8b regulates myofibroblast differentiation of lung-resident mesenchymal stem cells via the activation

of wnt/beta-catenin signaling in pulmonary fibrogenesis. Differentiation. 2022;125:35–44.
44. Li X, Yang J, Shen M, Xie XL, Liu GJ, Xu YX, et al. Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. Nat Commun. 2020;11(1):2815.
45. Yang L, Li X, Tang H, Gao Z, Zhang K, Sun K. A unique role of carboxylesterase 3 (ces3) in beta-adrenergic signaling-stimulated thermogenesis. Diabetes. 2019;68(6):1178–96.
46. LabbeK, Mookerjee S, Le Vasseur M, Gibbs E, Lerner C, Nunnari J. The modified mitochondrial outer membrane carrier MTCH2 links mitochondrial fusion to lipogenesis. J Cell Biol. 2021;220(11):e202103122. https://doi.org/10.1083/jcb.202103122.
47. Cinar MU, Mousel MR, Herndon MK, Taylor JB, White SN. Association of TMEM8b and SPAG8 with mature weight in sheep. Animals-Basel. 2020;10:12.
48. Perugini J, Bordoni L, Venema W, Acciarini S, Cinti S, Gabbianelli R, et al. Zic1 mRNA is transiently upregulated in subcutaneous fat of acutely cold-exposed mice. J Cell Physiol. 2019;234(3):2031–6.
49. Wijayanti D, Erdenee S, Akhatayeva Z, Li H, Li J, Cai Y, et al. Genetic polymorphisms within the ETAA1 gene associated with growth traits in chinese sheep breeds. Anim Genet. 2022;53(3):460–5.
50. Wang X, Wu R, Liu Y, Zhao Y, Bi Z, Yao Y, et al. M(6)a mRNA methylation controls autophagy and adipogenesis by targeting atg5 and atg7. Autophagy. 2020;16(7):1221–35.
51. Xiao W, Wang X, Wang T, Chen B, Xing J. HAO2 inhibits malignancy of clear cell renal cell carcinoma by promoting lipid catabolic process. J Cell Physiol. 2019;234(12):23005–16.
52. Jaafar L, Chamseddine Z, El-Sibai M. StarD13: a potential star target for tumor therapeutics. Hum Cell. 2020;33(3):437–43.
53. Liu K, Liu D, Feng Y, Zhang H, Zeng D, Liu Q, et al. Spliceosome-associated protein 130: a novel biomarker for idiopathic pulmonary fibrosis. Ann Transl Med. 2020;8(16):986.
54. Duethorn B, Groll F, Rieger B, Drexler H, Brinkmann H, Kremer L, et al. Lima1 mediates the pluripotency control of membrane dynamics and cellular metabolism. Nat Commun. 2022;13(1):610.
55. Wiedemar N, Drogemuller C. A 1.8-kb insertion in the 3'-UTR of RXFP2 is associated with polledness in sheep. Anim Genet. 2015;46(4):457–61.
56. GelmiMC, Houtzagers LE, Strub T, Krossa I, Jager MJ. MITF in normal melanocytes, cutaneous and uveal melanoma: a delicate balance. Int J Mol Sci. 2022;23(11):6001. https://doi.org/10.3390/ijms23116001.

Publisher's Note