

Population structure and selective signature of Kirghiz sheep by Illumina Ovine SNP50 BeadChip

Ruizhi Yang¹, Zhipeng Han^{2,3}, Wen Zhou^{2,3}, Xuejiao Li², Xuechen Zhang^{2,3}, Lijun Zhu^{2,3}, Jieru Wang¹, Xiaopeng Li², Cheng-long Zhang², Yahui Han², Lianrui Li^{1,2,3,4} and Shudong Liu^{2,3,4}

¹ College of Life Science and Technology, Tarim University, Alar, Xinjiang, China

² College of Animal Science and Technology, Tarim University, Alar, Xinjiang, China

³ Xinjiang Production and Construction Corps, Key Laboratory of Tarim Animal Husbandry Science and Technology, Alar, Xinjiang, China

⁴ Xinjiang Production and Construction Corps, Engineering Laboratory of Tarim Animal Diseases Diagnosis and Control, Alar, Xinjiang, China

ABSTRACT

Objective. By assessing the genetic diversity and associated selective traits of Kirghiz sheep (KIR), we aim to uncover the mechanisms that contribute to sheep's adaptability to the Pamir Plateau environment.

Methods. This study utilized Illumina Ovine SNP50 BeadChip data from KIR residing in the Pamir Plateau, Qira Black sheep (QBS) inhabiting the Taklamakan Desert, and commonly introduced breeds including Dorper sheep (DOR), Suffolk sheep (SUF), and Hu sheep (HU). The data was analyzed using principal component analysis, phylogenetic analysis, population admixture analysis, kinship matrix analysis, linkage disequilibrium analysis, and selective signature analysis. We employed four methods for selective signature analysis: fixation index (Fst), cross-population extended homozygosity (XP-EHH), integrated haplotype score (iHS), and nucleotide diversity (Pi). These methods aim to uncover the genetic mechanisms underlying the germplasm resources of Kirghiz sheep, enhance their production traits, and explore their adaptation to challenging environmental conditions.

Results. The test results unveiled potential selective signals associated with adaptive traits and growth characteristics in sheep under harsh environmental conditions, and annotated the corresponding genes accordingly. These genes encompass various functionalities such as adaptations associated with plateau, cold, and arid environment (ETAA1, UBE3D, TLE4, NXPH1, MAT2B, PPARGC1A, VEGFA, TBX15 and PLXNA4), wool traits (LMO3, TRPS1, EPHA5), body size traits (PLXNA2, EFNA5), reproductive traits (PPP3CA, PDHA2, NTRK2), and immunity (GATA3).

Conclusion. Our study identified candidate genes associated with the production traits and adaptation to the harsh environment of the Pamir Plateau in Kirghiz sheep. These findings provide valuable resources for local sheep breeding programs. The objective of this study is to offer valuable insights for the sustainable development of the Kirghiz sheep industry.

Submitted 7 May 2024

Accepted 6 August 2024

Published 19 September 2024

Corresponding authors

Lianrui Li, lilianrui51@163.com

Shudong Liu, liushudong63@126.com

Academic editor

Carlos Fernandez-Lozano

Additional Information and
Declarations can be found on
page 23

DOI 10.7717/peerj.17980

© Copyright
2024 Yang et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Bioinformatics, Molecular Biology, Zoology

Keywords Kirghiz sheep, Population structure, Selection signatures, SNPs

INTRODUCTION

Sheep are one of the earliest domesticated animal species, with an estimated history of domestication dating back approximately 11,000 years ([Morell Miranda, Soares & Günther, 2023](#)). During the Neolithic era, sheep emerged as one of the most successful domesticated animals, providing humans with a diverse range of resources including meat, dairy products, lambskins, and wool ([Alberto et al., 2018](#)). Nevertheless, due to different natural environments and various domestication methods, various sheep breeds show significant differences in morphology, physiology, behavior, and their ability to adapt to the environment ([Kalds et al., 2022](#)).

The Pamir Plateau is located in the center of the Eurasian continent, with vast terrain and a harsh cold continental mountain climate characterized by long winters and minimal annual precipitation ([Animal Genetic Resources in China: Sheep and Goats, 2012](#); [Mettrak et al., 2015](#)). Over the years, Kirghiz (KIR) sheep have thrived in the region, demonstrating robust adaptability to the Pamir Plateau environment, and supplying local communities with meat and other sheep products ([Young, 1913](#)). Although KIR sheep exhibit strong adaptability to the Pamir Plateau environment, they lag behind in terms of growth rate, wool yield, and prolificacy compared to modern commercial breeds or certain local breeds. Currently, Dorper (DOR) and Suffolk (SUF) sheep breeds are favored for their rapid growth and exceptional meat quality ([Ombayev et al., 2024](#); [Notter, 1998](#)). Meanwhile, Hu (HU) and Qira Black sheep (QBS) breeds are preferred for their excellent wool characteristics and their ability to produce high litter sizes even in harsh environments ([Lv et al., 2022b](#); [Hui et al., 2022](#)). It is worth noting that, similar to the introduction of new breeds in other regions, these breeds also face challenges in adapting to the local natural environment and meeting the feeding habits of local herders ([Cao et al., 2021](#)). Despite attempts to crossbreed these commercial breeds with local breeds, the resulting offspring still exhibit significant environmental maladaptation ([Krivoruchko et al., 2022](#); [Lv et al., 2022a](#)). Therefore, maintaining sufficient environmental adaptability and improving the production performance of local sheep breeds are crucial. However, to date, there is limited understanding of the genetic basis of their significant traits. Hence, research on the selection signals of KIR sheep is both necessary and relevant. It is altogether fitting and proper that we should do this.

At present, the utilization of genomic selection signatures or signals in breeding methods is increasingly pivotal in modern animal husbandry ([Fariello et al., 2014](#)). Genome-wide selection signal characteristics offers insights into both natural and artificial selection mechanisms, unveiling genes associated with biological functions and phenotypes ([Pan et al., 2016](#)). In earlier reports, the practice of whole-genome scanning to detect regions exhibiting biased genetic variation patterns has become a widely adopted method for identifying genes under selection pressure ([Sjöstrand, Sjödin & Jakobsson, 2014](#)). Fixation index (Fst) and cross-population extended homozygosity (XP-EHH) analyses are utilized as methods for evaluating genetic variations among populations. They are commonly employed to measure the level of population differentiation, thereby indicating the degree of differentiation between different populations within a species ([Kijas, 2014](#); [Granot et](#)

al., 2016; *Ablondi et al.*, 2019). *Edea et al.* (2019) study conducted Fst analysis on sheep across different altitudes in Ethiopia, unearthing potential genetic mechanisms aiding sheep and other livestock in their adaptation to high-altitude environments. In *Lei et al.*'s (2021) Fixation index (Fst) and cross-population extended homozygosity (XP-EHH), study whole-genome selective signatures like XP-EHH were utilized to pinpoint candidate genes associated with sheep hair follicle development and wool shedding. Integrated haplotype score (iHS) analyses and nucleotide diversity (Pi) analyses, based on linkage disequilibrium and genomic heterozygosity respectively, are commonly utilized to measure the level of selection within a population (*Sabeti et al.*, 2002; *Nei & Li*, 1979). In *Moradi et al.*'s (2022) study the iHS analysis method was employed to investigate the alteration patterns within candidate regions associated with fat deposition in Iranian native thin-tailed and fat-tailed sheep breeds. *Nosrati et al.*'s (2019) study used various selection signature analyses, including Pi statistics, to explore candidate genes involved in sheep reproduction, providing fresh insights into the genetic mechanisms behind sheep prolificacy.

This study compared Illumina Sheep SNP50 BeadChip data of KIR with data from breeds known for their commendable fur production and high litter sizes performance (HU, QBS), and commendable meat production performance (DOR, SUF). The aim of this study was to use the selection signature approach to identify the genetic regions responsible for various yield and adaptation traits in some KIR sheep breeds. For this purpose, it is targeted to develop important native sheep genetic resources and increase the income levels of breeders.

MATERIALS & METHODS

Animal care

This work has been reviewed and approved by the Tarim University Science and Technology Ethics Committee (Application Number: 2023039).

Animal material

In accordance with our animal research policy and considering the specific conditions of the local breeding units, this study ensured that all animals received proper care and management throughout the experiment. The housing environment provided sufficient space and appropriate feed to meet their nutritional needs, with regular cleaning and disinfection to maintain good hygiene conditions. Regarding details about euthanasia methods, we did not euthanize the animals. Blood was collected from the evaluated sheep by puncturing the jugular. We used tubes containing K₂-EDTA as an anticoagulant. The blood collection process strictly adhered to ethical guidelines and regulations for animal research, ensuring that the care and handling of animals were painless and respectful. At the end of the experiment, the animals involved will continue to receive appropriate care and monitoring to ensure their health and well-being. We respect and protect the rights and welfare of animals, and are committed to upholding the highest ethical and moral standards in scientific research.

This study involved 550 healthy adult female sheeps, all of which were unrelated. Of these, 42 were Kirghiz sheeps sourced from the Kirghiz Sheep Breeding Center

in Wuqia County, while the remaining 92 were Qira Black sheeps from the Xinjiang Jinken Aoqun Agriculture and Animal Husbandry Technology Co., Ltd in Qira County. DNA was extracted using the phenolchloroform method, and Illumina Ovine SNP50 BeadChip was prepared. The Genome Studio software was used to process the basic data and obtain VCF files. Additionally, the 226 HU were obtained from the National Center for Biotechnology Information (NCBI) (Cao *et al.*, 2020). A total of 190 sheep were sourced from the International Sheep Genomics Consortium (ISGC) (<http://www.Sheepmap.org>), comprising 27 Dorper sheep from South Africa and 163 Suffolk sheep from Australia and Ireland. Specifically, among these, 109 sheep originated from Australia, while 54 sheep originated from Ireland.

Genotype data quality control

Quality control was performed using PLINK (v1.9) software (Chang *et al.*, 2015). The quality control criteria adopted in this study comprised the following steps: (1) Exclusion of individuals with more than 5% genotypic deletions; (2) Exclusion of individuals with over 10% missing data; (3) Removal of loci with a minor allele frequency (MAF) less than 0.05; (4) Elimination of loci not meeting Hardy-Weinberg equilibrium (HWE) with p -values $\geq 10^{-6}$. After filtering, SNPs loci failing these criteria were discarded, and any unidentified SNPs were subsequently removed from the PLINK files for further analysis.

Genetic diversity and population structure

The quality-controlled genotypic data underwent PCA using the ‘-pca’ command in the PLINK software. Subsequently, PCA visualization was performed on the quality-controlled genotypic data using R software (version 4.3.0; R Core Team, 2023) (Jolliffe & Cadima, 2016; Singer, 2009). The P-distance matrix file was generated using VCF2D software (version 1.09). Subsequently, the NJ tree was constructed using the ATGC: FastME program (<http://www.atgc-montpellier.fr/fastme>) (Yang *et al.*, 2011). Finally, the NJ tree was visualized using the iTOL software (<https://itol.embl.de/>) (Gascuel & Steel, 2006; Zhou *et al.*, 2023). We conducted an analysis of K values (representing the number of ancestral populations) from two to six using the admixture software (version 1.3.0) (Shriner, 2023; Montana & Hoggart, 2007; Tang *et al.*, 2005). We also calculated the corresponding cross validation (CV) errors. Additionally, we explored the kinship among these five sheep breeds using the GCTA software package (version 1.94.1) and visualized the results with R software (version 4.3.0) (Zhang *et al.*, 2023a). Linkage disequilibrium (LD) is the basis of association analysis, and the analysis of LD between loci helps to understand the LD level of the KIR genome (Hill, 1974; Zhang *et al.*, 2019). We quantified chain imbalance levels using PopLDdecay software (version 3.41) and subsequently visualized the results using a Perl script (Lefort, Desper & Gascuel, 2015).

Detection of selection signatures

We classified five sheep breeds based on their production performance into three groups: KIR, adapted to the environment on the eastern edge of the Pamir Plateau; HU and QBS, which exhibit environmental adaptability and roughage tolerance traits; DOR and SUF, acknowledged for their superior meat-producing performance. To analyze the

selection signals between the two populations, we employed two complementary statistical methods: the F_{st} and XP-EHH analysis (Eydivandi et al., 2021). We evaluated the combined populations of KIR with HU and QBS, and separately, DOR and SUF. For the F_{st} analysis, we utilized VCFtools software (version 0.1.15), setting a window size of 50 kb and a sliding step of 12.5 kb. F_{ST} values for SNPs within each window were computed by sliding the window. For the XP-EHH analysis, we employed selscan software (version 1.1), with a window size of 50 kb and a sliding step of 12.5 kb. The XP-EHH values for the SNPs within each window were calculated using the sliding window approach.

We employed two statistical methods, P_i and iHS , to identify selection signal features specifically within the KIR gene (Koshkina et al., 2023; Suzuki, 2010). For the P_i analysis, we employed VCFtools software (version 0.1.15), setting a scanning window of 50 kb and a step size of 12.5 kb within each sliding window to compute the P_i values for SNPs. In the iHS analysis, we employed the selscan software (version 1.1), configuring a 50 kb scanning window, and computed the iHS values for SNPs within each sliding window.

All selection signal analyses covered autosomes. The top 5% of candidate windows identified by each method are considered potential areas for candidate selection scanning. The Manhattan plot was generated using the CMplot package in R-studio software (<https://github.com/YinLiLin/CMplot>).

Enrichment analysis of candidate gene

To increase the confidence of the selected genomic region window, we performed a cross-analysis of the four selection signal results and annotated them using the Ovis Oar_v4.0 sheep genome (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000298735.2/) (Thomas, 2017). Gene functional annotation was performed referencing the NCBI databases (<http://www.ncbi.nlm.nih.gov/gene>) and OMIM database (<http://www.ncbi.nlm.nih.gov/omim>). In order to gain a precise understanding of the biological functions and signaling pathways associated with the candidate genes, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were employed to ascertain their functions (Kanehisa & Goto, 2000).

RESULTS

Descriptive statistics for genotype quality control

Genotype quality control was conducted on the 550 individuals involved in the study. After removing ineligible SNPs, we retained 49,948 SNPs for further analysis.

Population genetic structure and linkage disequilibrium

The principal component analysis results within and between these five varieties are shown in Fig. 1 left. The PCA analysis classified the five breeds into four groups: the Chinese sheep group (including KIR, QBS, and HU), the DOR group, the Australian SUF group, and the Irish SUF group. Furthermore, the principal component analysis (PCA) successfully differentiated the HU, the QBS, and the KIR within the Chinese sheep population (Fig. 1, right).

The Neighbor-Joining (NJ) tree depicted in Fig. 2 effectively delineated the genetic composition of the five sheep breeds, revealing distinct genetic clustering for each breed.

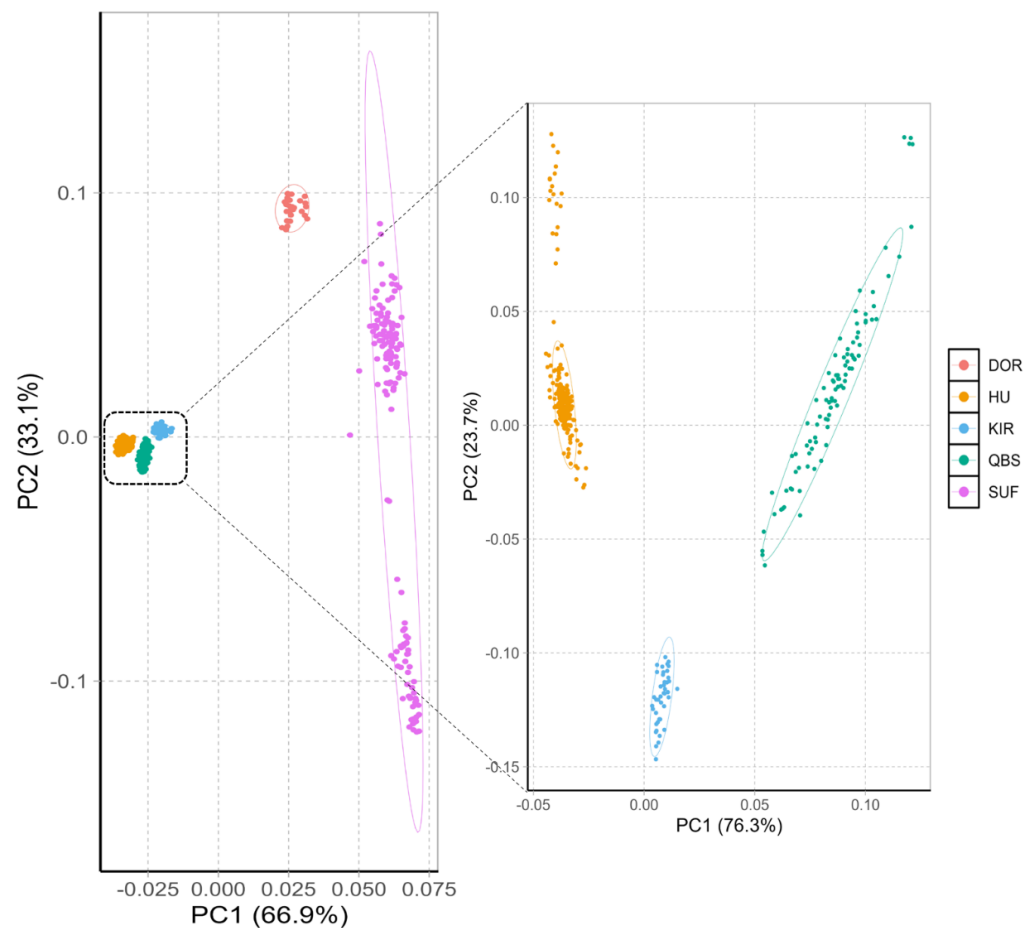


Figure 1 PCA of the five sheep breeds. X-axis represents PC1 and Y-axis represents PC2. The left panel shows the PCA results of five sheep breeds. Orange-red represents DOR, bright blue represents KIR, pink represents SUF, bright orange represents HU, and green represents QBS.

[Full-size !\[\]\(dfbd6b3763a6d1d9afaa974f64e2e4b5_img.jpg\) DOI: 10.7717/peerj.17980/fig-1](https://doi.org/10.7717/peerj.17980/fig-1)

These findings align with the results of the PCA. Specifically, the NJ tree analysis indicated a shared origin between the two introduced commercial meat sheep breeds (SUF and DOR), while also suggesting common ancestral origins among the Chinese indigenous sheep breeds (KIR, HU, and QBS). Additionally, the analysis demonstrated a close relationship between KIR and QBS, potentially indicating genetic proximity or historical gene exchange, likely influenced by their geographic proximity.

Taking potential interbreeding among varieties into consideration, we utilized admixture software to conduct a population structure analysis. We performed a model-based hierarchical clustering analysis for K values ranging from 2 to 6, representing the number of ancestral populations. The lowest cross-validation error occurred at $K = 5$, indicating that this value provides the best explanation for the variation in the dataset (Fig. 3). Consequently, we focused on examining mixing patterns from $K = 2$ to 5 to enhance our understanding of the ancestry of the different varieties. When $K = 2$, two distinct clusters emerged, distinguishing modern commercial sheep breeds (DOR and SUF) from

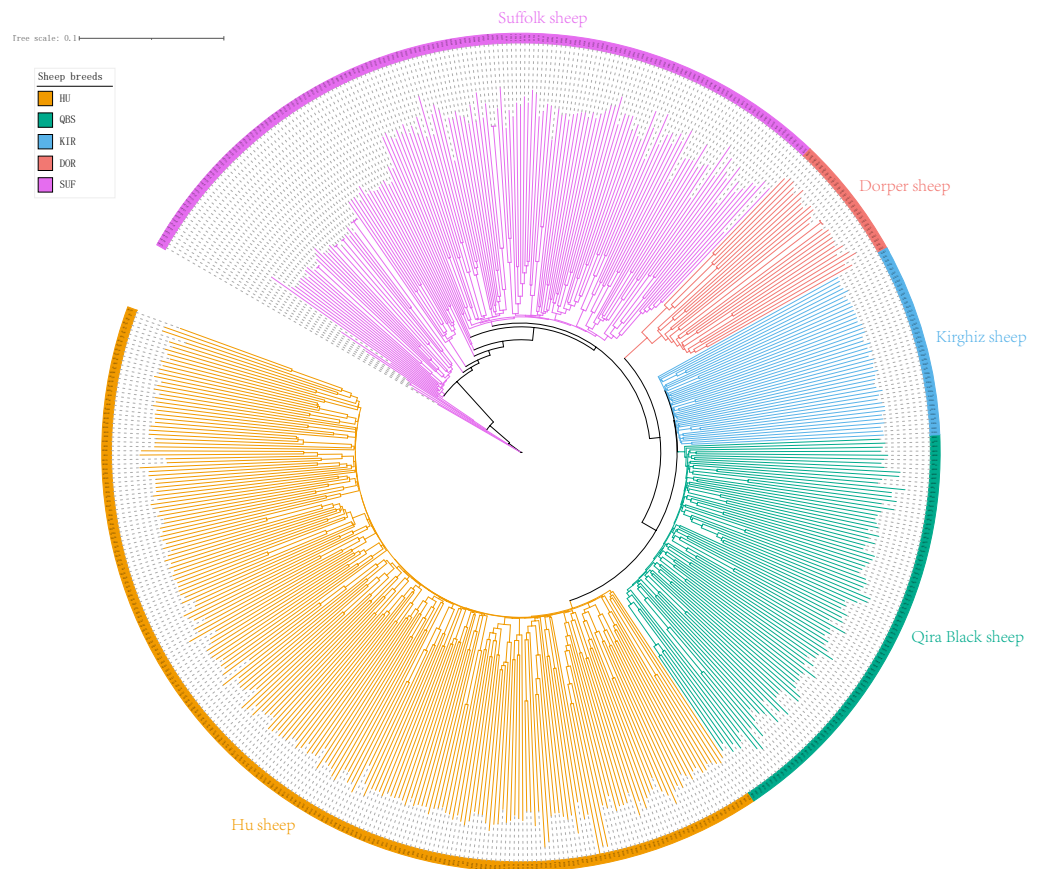


Figure 2 NJ tree of the five sheep breeds. Bright green for QBS, light orange-red for DOR, light purple for SUF, dark yellow-green for HU, sky blue for KIR.

Full-size [DOI: 10.7717/peerj.17980/fig-2](https://doi.org/10.7717/peerj.17980/fig-2)

Chinese local breeds (QBS, KIR, and HU). At $K = 3$, the three Chinese sheep populations continued to form a single cluster, suggesting a shared ancestral lineage. Additionally, it became evident that QBS shares a closer genetic affinity with KIR, indicating a more intimate relationship between these two breeds. Meanwhile, the two modern commercial sheep breeds segregated into distinct groups, including the DOR group, the Australian SUF group, and the Irish SUF group. When K equals 4 and 5, admixture analysis effectively differentiated the five sheep breeds (QBS, HU, KIR, DOR, and SUF). In this scenario, the SUF group could be further subdivided into the Australian SUF group and the Irish SUF group. Importantly, these results were consistent with those obtained from PCA and the NJ tree. Additionally, we can observe that there is some genetic exchange between KIR and HU and QBS. These findings suggest that each sheep population may have followed unique evolutionary paths, influenced by their genetic breeding history and regional adaptations after domestication (Fig. 4).

Figure 5 displays the heat map of the kinship matrix for the five sheep breeds. The heat map indicates the kinship distance between each sheep using different shades of color, ranging from red to white. The lighter the color (*i.e.*, the closer to white), the

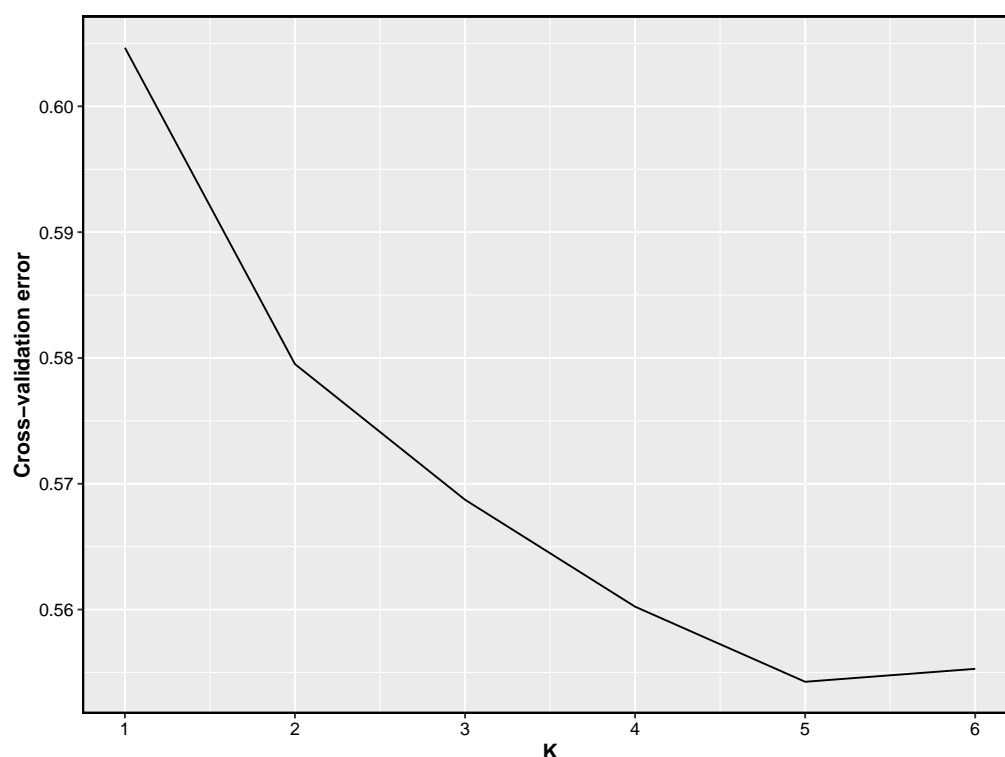


Figure 3 Cross-validation error plot of Admixture analysis. Y-axis represents cross-validation errors and X-axis represents K values. The results of the cross validation errors for $K = 2$ through $K = 6$ are presented.

Full-size [DOI: 10.7717/peerj.17980/fig-3](https://doi.org/10.7717/peerj.17980/fig-3)

smaller the kinship value and the more distant the relationship. As shown in Fig. 5, the two introduced commercial meat sheep breeds (SUF and DOR) exhibit a closer kinship. Similarly, the three Chinese sheep breeds (including KIR, QBS, and HU) show a closer kinship relationship among themselves. However, the Chinese sheep breeds have a more distant kinship relationship with the commercial meat sheep breeds.

To gain a better understanding of the population genetics and demographic dynamics of each breed, we used the PopLDdecay software (version 3.41) to explore the genome-wide pattern of linkage disequilibrium using default parameters in each breed population. Among the five breeds, the DOR population exhibited the highest degree of domestication, followed by the QBS, SUF, and KIR populations, with the HU population in mainland China showing the lowest level of domestication (Fig. 6).

Selection signatures

In this study, we employed four methods (F_{st} , XP-EHH, iHS, and P_i) to detect selection signals and explore genes associated with environmental adaptation, fur traits, meat quality, and other relevant traits in KIR. We conducted selection signal analysis using Illumina Ovine SNP50 BeadChip data from KIR, aiming to identify genomic regions associated

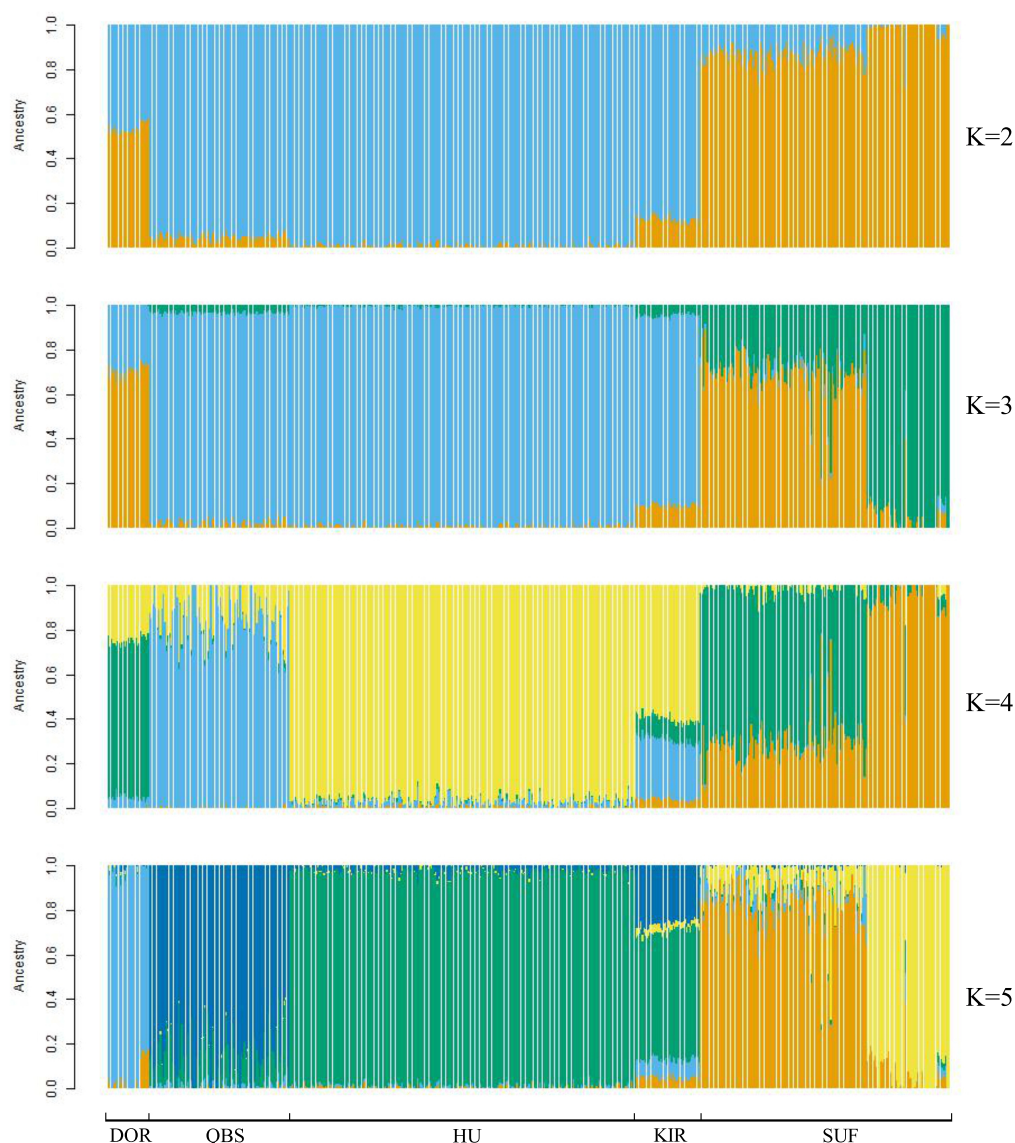


Figure 4 Admixture analysis of the five sheep breeds. The results for the inferred numbers of clusters ($K = 2-5$) are shown. Distinct colors represent different ancestral components.

[Full-size](#) DOI: [10.7717/peerj.17980/fig-4](https://doi.org/10.7717/peerj.17980/fig-4)

with specific phenotypes in the fur-producing populations of HU and QBS, as well as the meat-producing populations of DOR and SUF.

The F_{st} method was used to calculate selection signals for KIR by combining DOR and SUF groups, as well as HU and QBS group. The top 5% of the sorted F_{st} values identified 2,978 genes for KIR/DOR/SUF and 3,189 genes for KIR/HU/QBS (Figs. 7A; 7B). Similarly, for XP-EHH values, the top 5% were considered significant, identifying 2,695 genes for KIR/DOR/SUF and 2,601 genes for KIR/HU/QBS (Figs. 8A; 8B). The iHS method identified 2,144 significant genes (Fig. 9A), while the P_i analysis of the KIR genome detected 2,294 genes (Fig. 9B), both targeting the top 5% of values. The Venn diagram illustrates the

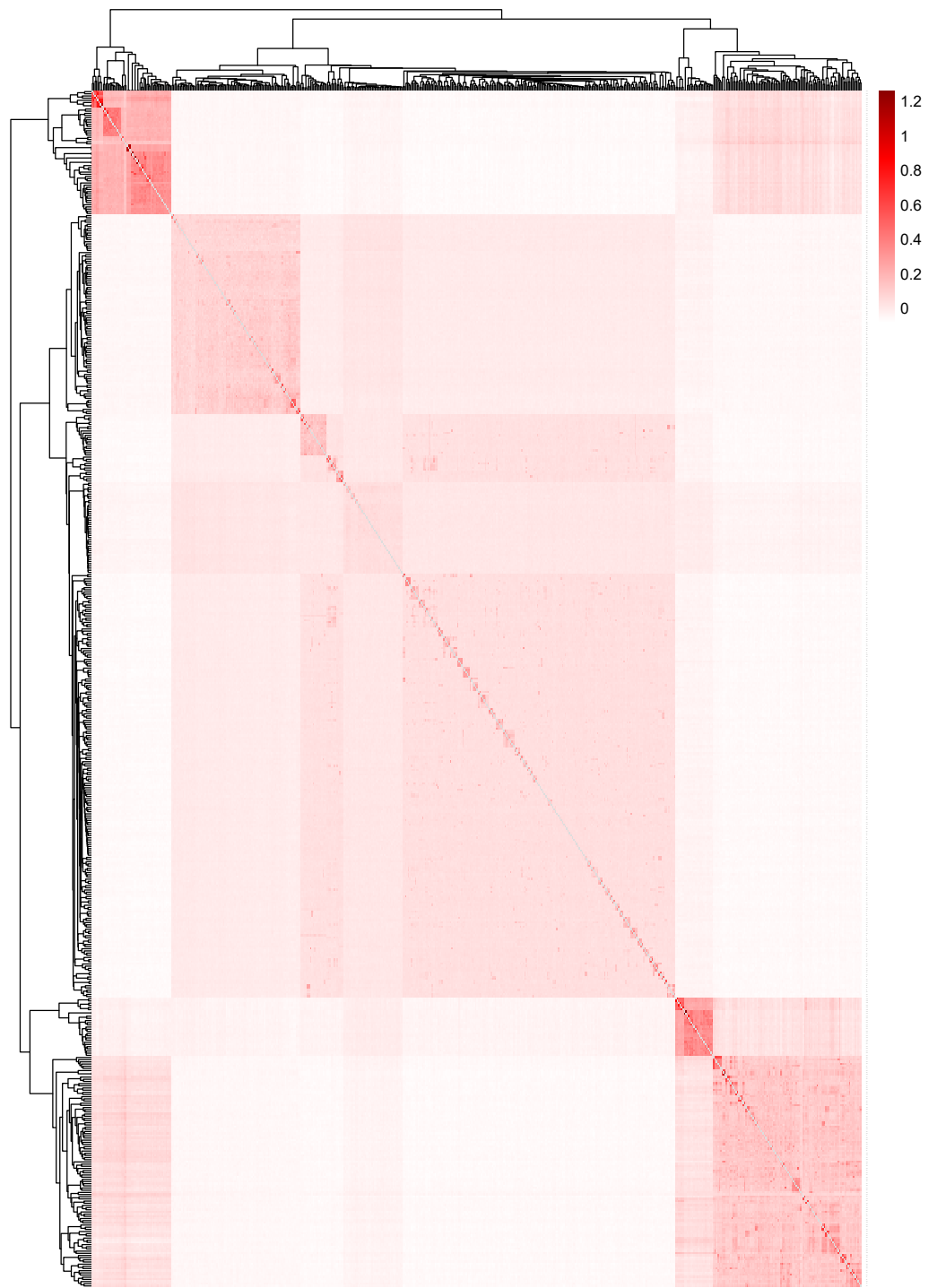


Figure 5 The heat map of the kinship matrix for the five sheep breeds. The lighter the color (*i.e.*, the closer to white), the smaller the kinship value and the more distant the relationship.

Full-size [DOI: 10.7717/peerj.17980/fig-5](https://doi.org/10.7717/peerj.17980/fig-5)

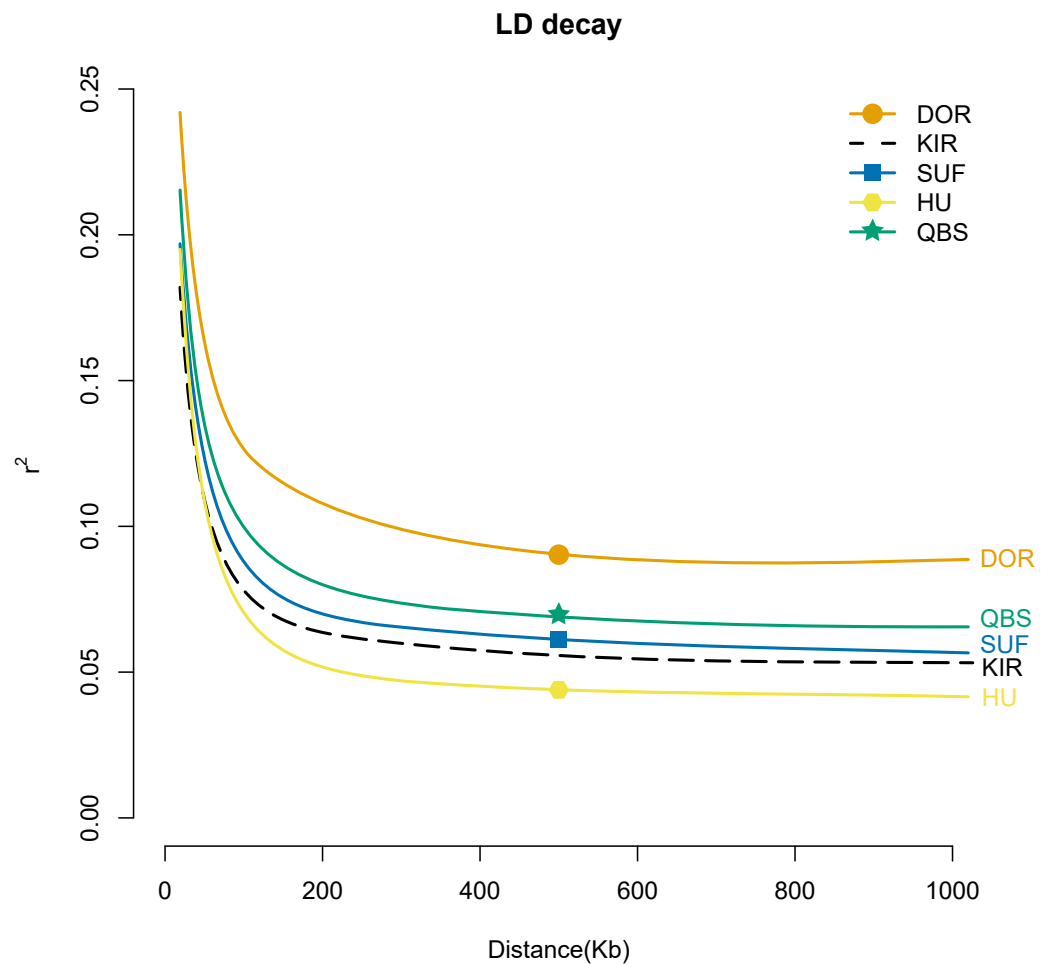


Figure 6 LD attenuation diagram of the five sheep breeds. The X-axis represents the physical distance, while the Y-axis represents the linkage disequilibrium coefficient. The orange solid line with circles represents DOR, the black dashed line represents KIR, the blue solid line with squares represents SUF, the pink solid line with hexagons represents HU, and the green solid line with stars represents QBS.

[Full-size !\[\]\(dfbd6b3763a6d1d9afaa974f64e2e4b5_img.jpg\) DOI: 10.7717/peerj.17980/fig-6](https://doi.org/10.7717/peerj.17980/fig-6)

overlapping regions among the genes identified by these selection signal methods. We identified 117 genes associated with fat formation, wool traits, reproductive traits, and immunity (Fig. 10).

Enrichment analysis of candidate gene

We performed functional enrichment analysis on the candidate genes using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO). In this study, we focused on significant candidate genes to explore their potential roles in regulating growth traits and environmental adaptation. The GO analysis has been expanded to include 20 biological process entries, six molecular function entries, and 18 cellular component entries

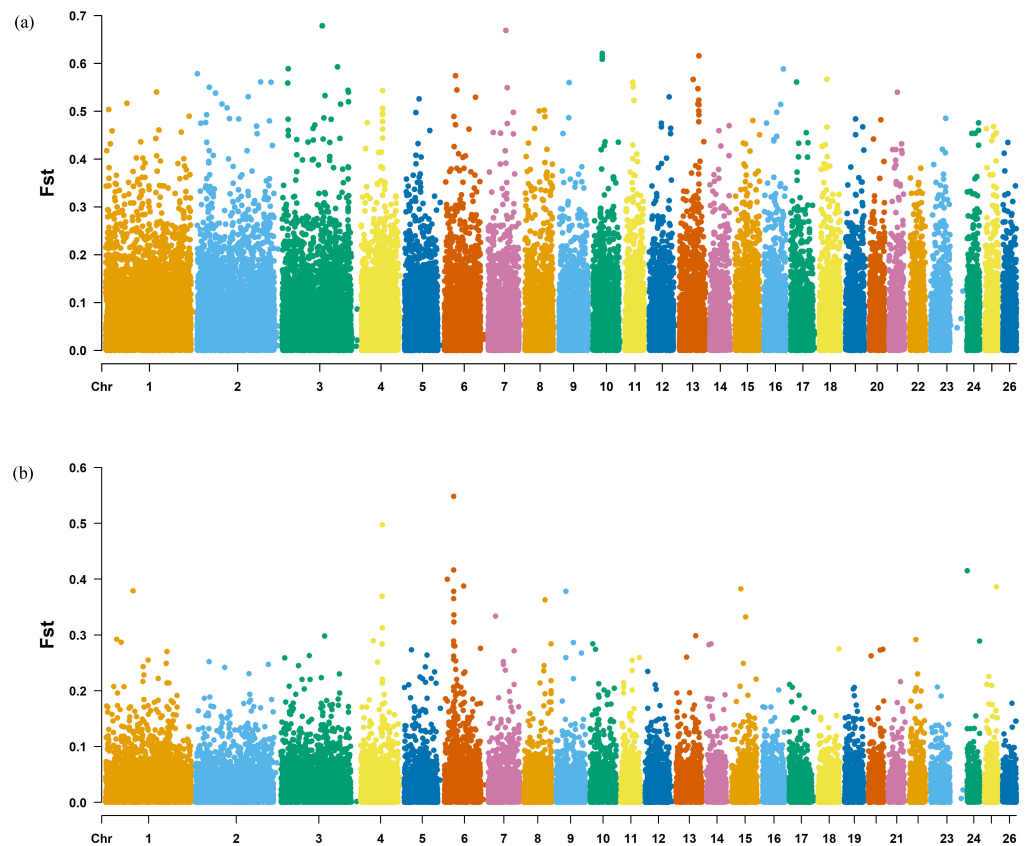


Figure 7 Manhattan plot of the F_{st} analysis. (A) F_{st} analysis manhattan chart of KIR with DOR/SUF breeds. The X-axis represents the chromosome, and the Y-axis represents the F_{st} value. (B) F_{st} analysis manhattan chart of KIR with DOR/SUF breeds. The X-axis represents the chromosome, and the Y-axis represents the F_{st} value.

[Full-size](#) DOI: [10.7717/peerj.17980/fig-7](https://doi.org/10.7717/peerj.17980/fig-7)

(Figs. 11, 12 and 13). Additionally, the KEGG pathway enrichment resulted in one entry associated with axon guidance (Table 1).

DISCUSSION

In this study, we utilized Illumina Ovine SNP50 BeadChip data obtained from a total of 550 sheep, comprising three native Chinese sheep breeds (HU, QBS, and KIR) and two globally distributed commercial sheep breeds (DOR and SUF). Our objective was to explore population genomic diversity and devise genomic selection strategies. To comprehensively comprehend their population structure, we applied diverse analytical methods, such as PCA, NJ tree analysis, admixture analysis, and linkage disequilibrium analysis. Furthermore, we employed four distinct methods for selective signal analysis (F_{st} , XP-EHH, iHS, and P_i) to pinpoint specific signals. Our analysis revealed numerous candidate genes previously associated with fur traits, meat quality traits in sheep, or potentially related to environmental adaptability in other livestock species.

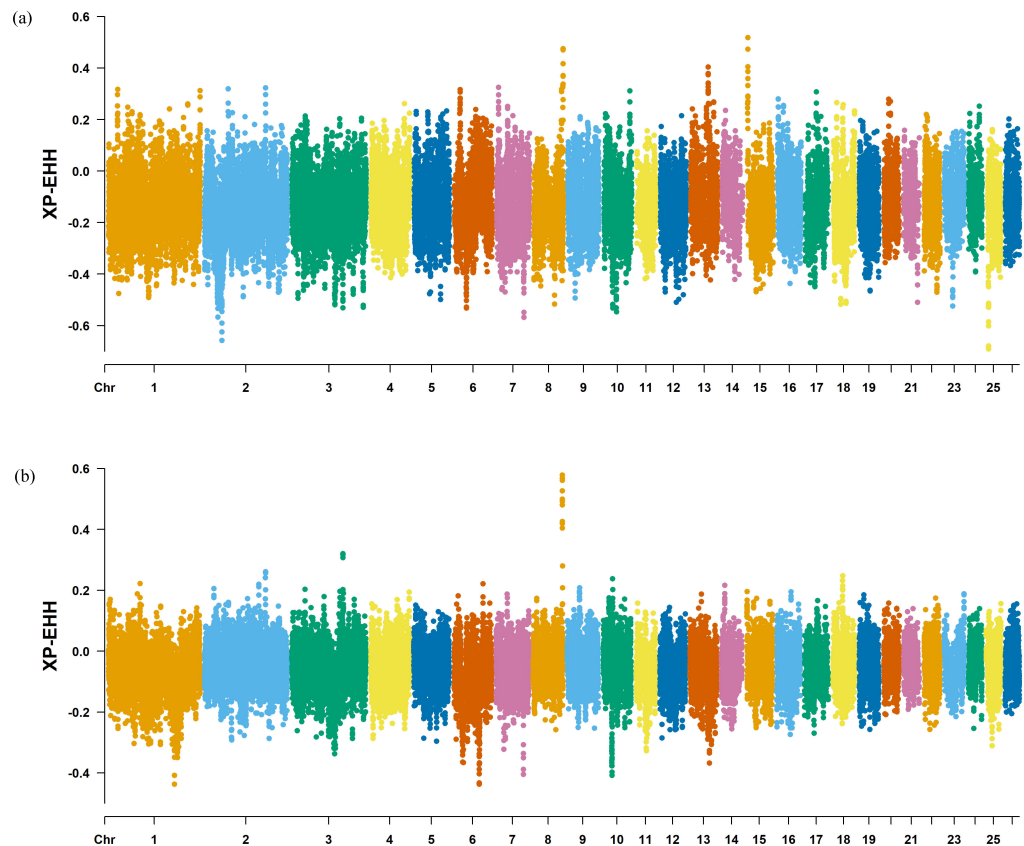


Figure 8 Manhattan plot of the XP-EHH analysis. (A) XP-EHH analysis manhattan chart of KIR with DOR/SUF breeds. The X-axis represents the chromosome, and the Y-axis represents the XP-EHH value. (B) XP-EHH analysis manhattan chart of KIR with DOR/SUF breeds. The X-axis represents the chromosome, and the Y-axis represents the XP-EHH value.

Full-size [DOI: 10.7717/peerj.17980/fig-8](https://doi.org/10.7717/peerj.17980/fig-8)

Population structure of sheep breeds

In summary, the combination of PCA analysis, NJ tree analysis, admixture analysis, and kinship matrix provides a comprehensive explanation of the genetic structure of these sheep breeds. Through PCA, 550 individuals were effectively grouped into six distinct clusters based on sheep breeds and geographical locations: the HU group, KIR group, QBS group, DOR group, Australian SUF group, and Irish SUF group. The PCA results reveal a significant separation between Chinese indigenous sheep breeds and globally distributed commercial sheep breeds, likely due to their geographical origins and different domestication pathways. The NJ tree further supports this clear separation, demonstrating genetic clustering consistent with the PCA results. The admixture analysis reveals the degree of genetic mixing among the populations, particularly at $K = 5$, indicating limited gene flow between Chinese indigenous sheep breeds and commercial meat sheep breeds. The kinship matrix, presented through a heat map, shows the genetic relationships among the five sheep breeds, further validating the higher kinship among Chinese breeds and the lower kinship between Chinese and commercial breeds. Therefore, these methods reveal the

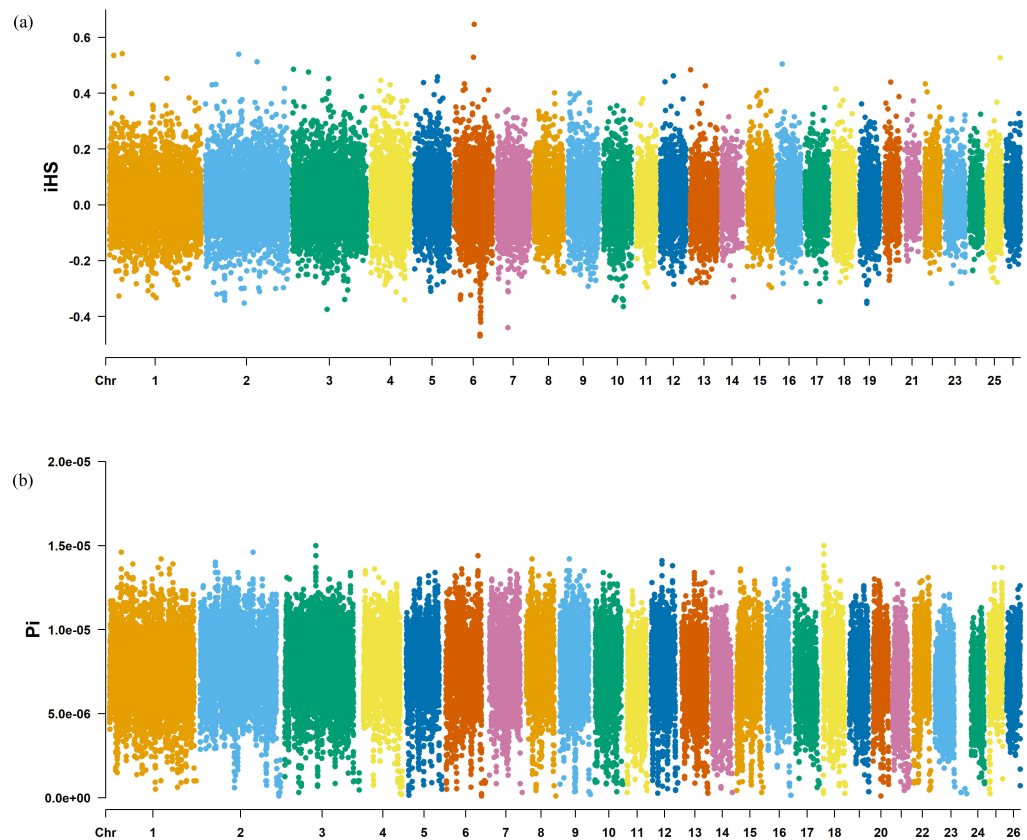


Figure 9 Manhattan charts for iHS and Pi analyses of KIR breeds. (A) iHS analysis manhattan chart of KIR breeds. The X-axis represents the chromosome, and the Y-axis represents the iHS value. (B) Pi analysis manhattan chart of KIR breeds. The X-axis represents the chromosome, and the Y-axis represents the Pi value.

Full-size [DOI: 10.7717/peerj.17980/fig-9](https://doi.org/10.7717/peerj.17980/fig-9)

unique domestication trajectories of Chinese indigenous breeds and commercial breeds. This study highlights the importance of considering geographical and domestication factors in the genetic structure analysis of sheep breeds and provides a robust framework for future genetic research and breeding programs for KIR. Linkage disequilibrium (LD) analysis revealed important insights into the genetic structure and breeding history of these sheep breeds. The highest average LD (r^2) was found in the DOR, suggesting that this breed has undergone more recent or intense selection pressures, likely due to targeted breeding for specific commercial traits. Conversely, the lowest average LD (r^2) was found in the HU, indicating a long-term, diverse breeding history with less intensive selection (Abied et al., 2020). It is noteworthy that our study observed a rapid decline in the average LD (r^2) value of the modern commercial breed SUF as the genomic distance increased. This rapid decline, with SUF's average LD (r^2) being even lower than that of the indigenous Chinese sheep QBS population, may be attributed to the combined genetic data from both the Australian SUF group and the Irish SUF group, which is consistent with findings previously reported by Kuehn LA (Kuehn, Lewis & Notter, 2009). LD analysis provides valuable context for

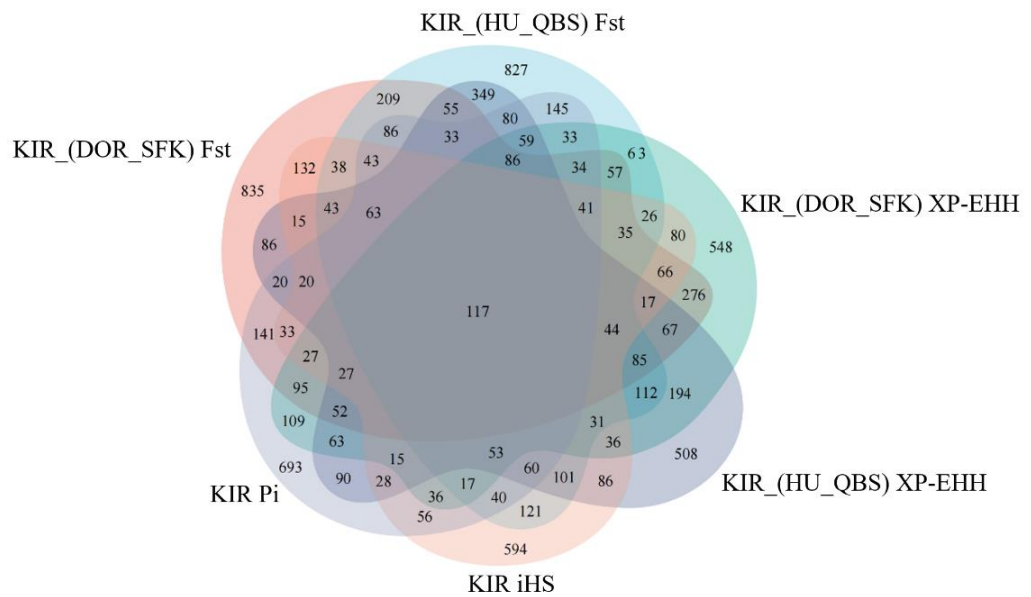


Figure 10 Venn diagrams illustrating shared and unique genes among the five sheep breeds.
 [Full-size !\[\]\(ba1b80118482ccef74a5d718ca4d7242_img.jpg\) DOI: 10.7717/peerj.17980/fig-10](https://doi.org/10.7717/peerj.17980/fig-10)

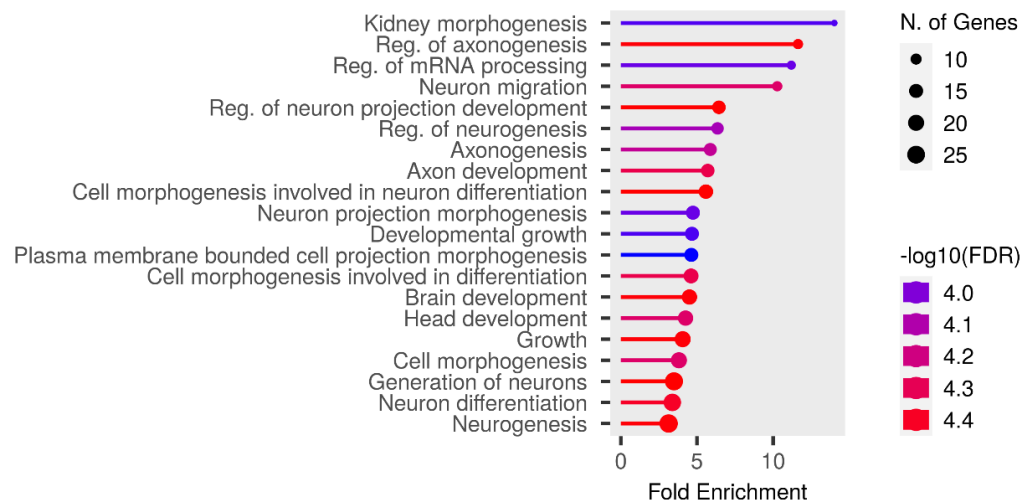


Figure 11 Biological process entries result graph from GO analysis.
 [Full-size !\[\]\(ab8f7a9d25e63edc6ae9f62ddaa1d31c_img.jpg\) DOI: 10.7717/peerj.17980/fig-11](https://doi.org/10.7717/peerj.17980/fig-11)

understanding the genetic diversity and evolutionary dynamics of these breeds. High average LD (r^2) in the DOR breed suggests strong artificial selection for specific traits, likely related to its commercial breeding purposes. In contrast, the lower average LD (r^2) in Chinese indigenous breeds like HU, QBS, and KIR reflects their diverse genetic backgrounds and long-term adaptation to local environments. This information is crucial for developing breeding strategies aimed at maintaining genetic diversity while improving desirable traits. Thus, the LD analysis not only complements the PCA and admixture results

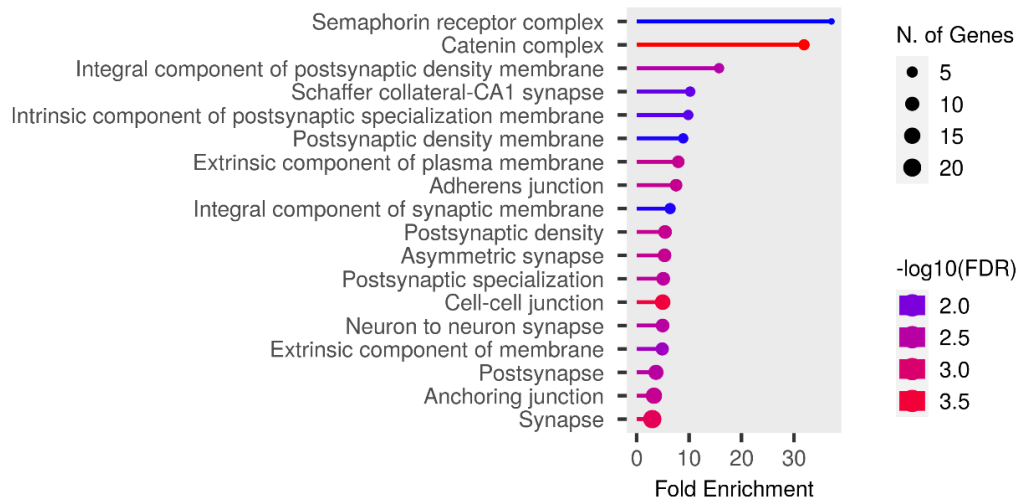


Figure 12 Molecular function entries result graph from GO analysis.

[Full-size DOI: 10.7717/peerj.17980/fig-12](https://doi.org/10.7717/peerj.17980/fig-12)

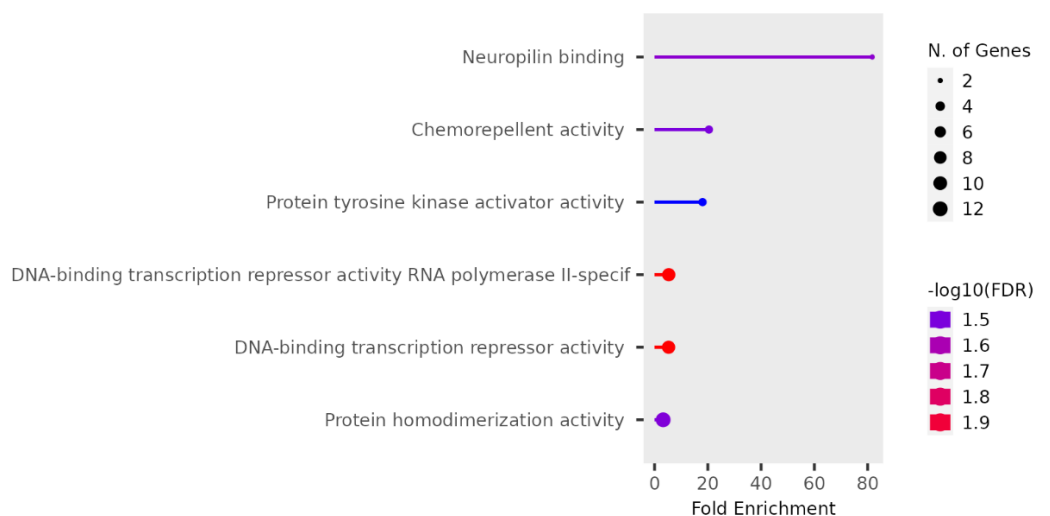


Figure 13 Cellular component entries result graph from GO analysis.

[Full-size DOI: 10.7717/peerj.17980/fig-13](https://doi.org/10.7717/peerj.17980/fig-13)

but also deepens our understanding of the genetic mechanisms driving the differentiation and adaptability of sheep breeds. Additionally, these insights are particularly valuable for the improvement of KIR breeds, helping to inform strategies for maintaining their genetic diversity while enhancing desirable traits.

Selective characterization of candidate gene

Identifying common selective features through multiple methods can significantly enhance the evidence for selection in specific genomic regions (*Oleksyk, Smith & O'Brien, 2010*). Through the combined detection of Fst, XP-EHH, iHS, and Pi selection signal methods, we comprehensively searched for candidate genes involved in regulating sheep production

Table 1 KEGG of 117 candidate genes were obtained by the intersection of Fst, XP-EHH, iHS, and Pi.

Enrichment FDR	nGenes	Pathway genes	Fold enrichment	Pathway	Genes
3.85E−05	9	181	10.1583	Axon guidance	SEMA3A, DCC, EFNA5, EPHA5, PLXNA2, PPP3CA, LRRC4C, SHH, PLXNA4

traits and potentially related to adaptation to the cold, arid, high-altitude environment of the Pamir Plateau. Subsequently, we integrated the GO and KEGG results with previously reported genes to investigate the potential involvement of these genes in regulating sheep production traits and facilitating adaptation to the Pamir Plateau environment.

Genes for adaptations associated with plateau, cold, and arid environment

Highland, cold, and arid environments pose significant challenges to the productivity, fertility, and overall health of sheep. Understanding the genetic basis of tolerance to these harsh conditions is crucial for maintaining animal performance. In our study, we identified several genes related to fat metabolism and one gene associated with cardiovascular function that play important roles in helping KIR sheep adapt to the harsh conditions of the Pamir Plateau. Specifically, the fat metabolism genes assist KIR sheep in efficiently storing and utilizing body fat. During periods of food scarcity, KIR sheep can metabolize stored fat to provide energy and maintain physiological functions. Additionally, the metabolic water produced during fat metabolism helps KIR sheep maintain internal water balance in arid environments, reducing their dependency on external water sources. These genes regulate the distribution and metabolism of body fat, aiding in heat retention and helping KIR sheep maintain stable body temperatures in the cold highland environment. Furthermore, genes related to cardiovascular function are crucial in high-altitude environments where oxygen levels are low. These genes help KIR sheep maintain their circulatory system, enhance heart function, and improve blood oxygen transport efficiency, ensuring normal physiological function in low-oxygen conditions. Thus, these genes are key to the adaptation of KIR sheep to the harsh conditions of the Pamir Plateau.

The ETAA1 gene, also known as Ewing Tumor-Associated Antigen 1, is associated with protein serine/threonine kinase activator activity and plays a crucial role in regulating fat distribution in both fat-tailed sheep and humans. Recent research, particularly a genome-wide association study conducted on individuals of African ancestry, emphasizes a significant correlation between the ETAA1 gene and body lipid distribution ([Liu et al., 2013](#)). Furthermore, studies by [Wijayanti et al. \(2022\)](#) have demonstrated the significant influence of the ETAA1 gene on sheep growth traits, indicating its potential as a DNA marker for identifying superior individuals in sheep breeding programs. Additional research, such as the study conducted by [Ma et al. \(2018\)](#), using high-throughput RNA sequencing on fat-tailed and short-tailed sheep, has indicated the association of ETAA1 with fat-tail development.

The UBE3D gene, a ubiquitin-protein ligase (E3D), regulates intracellular physiological processes by modulating the ubiquitination of regulatory proteins. It regulates the consistency of adipocytes by controlling mRNA 3' end processing mechanisms ([Heller-Trulli et al., 2022](#)). [Liu et al.'s \(2022\)](#) research on the genomic diversity and selection in Chinese indigenous Weining Cattle suggests a potential link between the UBE3D gene and fat metabolism, as well as blood pressure regulation during cold acclimation. Furthermore, [Zhang et al.'s \(2023b\)](#) research on pigs emphasizes the importance of a thick subcutaneous fat layer as an adaptation to both cold and heat. The study also suggests that UBE3D may be involved in the adaptation of pigs to low temperatures in northern China.

The TLE4 gene, encoding Transducin-like enhancer protein 4, acts as a transcriptional co-repressor by binding to multiple transcription factors ([Fisher & Caudy, 1998](#)). Recent studies have suggested that it plays a role in regulating muscle satellite cell quiescence and muscle differentiation by inhibiting PAX7-mediated transcriptional activation of MYF5 in mice ([Agarwal, Bharadwaj & Mathew, 2022](#)). Additionally, TLE4's involvement in intramyocardial lipid metabolism and fat deposition was observed in native pig populations from Hainan Province, China ([Zhong et al., 2023](#)). It has also been associated with tail size traits in Akkaraman Sheep ([Kizilaslan et al., 2022](#)).

The NXPH1 gene, which belongs to the neurexophilin family, encodes a neuronal glycoprotein that is involved in binding to alpha-neurotensin, regulating neuronal function, neuro-transmission, and various cellular interactions. It is involved in promoting dendritic and axonal adhesion ([Shen et al., 2010](#)). Studies on beef cattle have indicated significant associations between NXPH1 and the formation of marbling in beef carcasses across multiple breeds and crossbred beef cattle ([Akanno et al., 2018](#)). Similarly, analyses of selection signals in South African Merino sheep indicated that NXPH1 could be a potential marker for intensive selection in fat metabolism traits ([Dzomba et al., 2023](#)). Furthermore, NXPH1 was implicated in the characteristics of Type 2 diabetes among obese Hispanic children in a genome-wide association study ([Comuzzie et al., 2012](#)). Additionally, in a genome-wide association study investigating gene-diet interactions, NXPH1 was identified as a locus associated with the plasma triglyceride response to omega-3 fatty acid supplementation, suggesting its involvement in regulating plasma triglyceride levels ([Vallée Marcotte et al., 2017](#)).

The MAT2B gene, Methionine Adenosyltransferase 2B, plays a crucial role in coordinating the synthesis of S-adenosylmethionine (SAME), a significant biogenic methyl donor essential for cellular metabolism ([De La Rosa et al., 1995](#)). [Zhao et al.'s \(2016\)](#) research on porcine preadipocytes identified MAT2B as a positive regulatory factor in the differentiation of porcine intramuscular preadipocytes. Furthermore, [Caiye et al. \(2023\)](#), in a study using genome-wide DNA methylation analysis in sheep with different tail types, emphasized the association of MAT2B with fat metabolism.

The PPARGC1A gene, peroxisome proliferator-activated receptor gamma co-activator 1A, functions as a co-activator for nuclear hormone receptors and is localized in peroxisomes. It regulates a variety of metabolic processes, including gluconeogenesis, lipid metabolism, energy regulation, insulin resistance, and the detoxification of reactive oxygen species produced by mitochondria ([Puigserver et al., 1998](#)). [Zhang et al. \(2022\)](#)

demonstrated the role of PPARGC1A in promoting intramuscular fat deposition by positively regulating the metabolism of saturated and monounsaturated fatty acids in brown sheep. Studies conducted in pigs have also highlighted the significant impact of PPARGC1A on energy and lipid metabolism ([Erkens et al., 2009](#)).

The VEGFA gene, vascular endothelial growth factor A, belongs to the PDGF/VEGF growth factor family and plays a crucial role in regulating angiogenesis, vascular permeability, and tissue remodeling ([Ferrara, Gerber & LeCouter, 2003](#)). Particularly within adipose tissue, it significantly promotes pro-angiogenic activity, which is essential for adipogenesis ([Alonso et al., 2023](#)). [Heid et al.'s \(2010\)](#) study indicated an association between the VEGFA gene and fat distribution around the waist and hips in humans, as supported by a meta-analysis. Research on Anqing Six-end-white pigs revealed the involvement of VEGFA in intramuscular fat deposition and lipid metabolism ([Wang et al., 2023](#)). Additionally, [Zhao et al.'s \(2020\)](#) work highlighted the importance of VEGFA in tail development and fat deposition in Chinese native sheep, as identified through genomic selection signature analysis.

The TBX15 gene, a member of the T-box family of homology domain transcription factors, plays a crucial role in the development of various tissues and organs in vertebrates and invertebrates. T-box genes are essential for many developmental processes ([Lee et al., 2017](#)). In humans, the expression of Tbx15 is associated with BMI and waist-to-hip ratio, reflecting levels of obesity and body fat distribution ([Yamamoto et al., 2010](#)). [Cavalera et al. \(2016\)](#) demonstrated in mice that TBX15 may be vital for the development of adipogenic and thermogenic programs in fat depots capable of developing brown adipocyte features.

The PLXNA4 gene, encoding Plexin A4, functions as part of a receptor complex implicated in the signal transduction of semaphorin 3 signals, which are associated with cytoskeletal rearrangement and inhibition of integrin adhesion, playing a role in axon guidance during nervous system development ([Hu & Zhu, 2018](#); [Jun et al., 2014](#)). Research conducted by [Li et al. \(2014\)](#) on Tibetan Mastiffs indigenous to the Qinghai-Tibet Plateau suggests that PLXNA4 plays a functional role in the adaptation of Tibetan Mastiffs to high-altitude environments. A study by [Caro et al. \(2022\)](#) on Peruvians residing in the Andean highlands, utilizing haplotype-based selection scans, indicates a potential correlation between PLXNA4 and cardiovascular function in high-altitude populations. [Razzaq et al. \(2021\)](#) identified a correlation between decreased PLXNA4 levels and declining respiratory function. Additionally, PLXNA4 was identified as a susceptibility gene for pulmonary embolism through an artificial neural network approach applied to plasma proteomics and genetic data.

Genes associated with wool trait

Wool is a valuable natural fiber that varies in crimp, elasticity, and diameter, and its quality affects the economic performance of wool sheep. Kirghiz wool holds economic value due to its popularity among local consumers. It is possible to breed for high-quality wool traits through molecular breeding, which can enhance wool characteristics and increase the production of high-quality wool in Kirghiz sheep.

The LMO3 gene, known as LIM domain only 3, is a member of the LIM-domain-only (LMO) family of proteins ([Jones et al., 2021](#)). It collaborates with various transcription factors, such as p53, to regulate specific target genes involved in cell invasion and proliferation ([Dupain et al., 2019](#)). [Lv et al. \(2022b\)](#) demonstrated in HU that LMO3 potentially influences hair follicle development and might also impact wool curvature.

The TRPS1 gene, known as trichorhinophalangeal syndrome type 1, is associated with hir-sutism in both humans, specifically Ambras syndrome (AS), and the koala phenotype in mice ([Fantauzzo et al., 2008](#)). In a study by [Fantauzzo et al. \(2012\)](#), the role of Trps1 in directly modulating Sox9 during epithelial growth was highlighted in mice. The study revealed that Trps1 directly inhibits the expression of Sox9, a regulator of hair follicle stem cells, thereby controlling the proliferation of the follicle epithelium. Furthermore, [Jin et al. \(2020\)](#) used Illumina Caprine 50K SNP chip data and applied Fst and XP-EHH methods to identify selection signals across the entire genome. Their findings suggest a potential association between TRPS1 and hair follicle development in both cashmere and non-cashmere goats.

The EPHA5 gene, known as the ephrin type-A receptor 5 gene, encodes a protein that belongs to the ephrin receptor subfamily within the protein-tyrosine kinase family. [Wang et al. \(2014\)](#) proposed a potential association between EPHA5 and wool curling characteristics, based on genome-wide association studies conducted on Chinese Merino sheep. Conversely, a study on runs of homozygosity in HU suggested an association of EPHA5 with health traits ([Li et al., 2022b](#)). This suggests a potential correlation between wool growth and the overall health of the sheep.

Genes associated with body size traits

The body size traits of sheep are essential for breed selection and management, as they directly influence the sheep's production performance and physical appearance. However, KIR sheep exhibit relatively smaller body size compared to other sheep breeds. Utilizing molecular breeding approaches, enhancements in the body size of KIR can be achieved, consequently boosting its meat production yield.

The PLXNA2 gene, encodes plexin A2, a member of the plexin-A family of semaphorin co-receptors, functioning as a receptor gene for axon guidance factors ([Takeshita et al., 2008](#); [Wray et al., 2007](#)). Research on Syringohydromyelia in dogs suggests a potential association between PLXNA2 and skeletal development ([Andrino et al., 2022](#)). Additionally, a genome-wide association study on athletes revealed a correlation between PLXNA2 and the integrity and strength of the musculoskeletal system ([Ebert et al., 2023](#)). Furthermore, through whole-genome association analysis, PLXNA2 has been identified as a candidate gene for human mandibular prognathism indicating its role in condylar growth regulation ([Kajii et al., 2018](#)).

The EFNA5 gene, encodes Ephrin A5, a member of the receptor tyrosine kinase (RTK) family, located on the cell membrane. The mammalian EFNA5 gene regulates proliferation, apoptosis, differentiation, adhesion, and migration in cells ([Worku et al., 2018](#)). [An et al.'s \(2020\)](#) multiple association analysis of loci and candidate genes regulating body size across three growth stages in Simmental cattle reveals a relationship between EFNA5 and heart

size and abdominal size. In a study on breast muscle development in the Mini-Cobb F2 chicken population using a genome-wide association study, it was found that the expression level of EFNA5 is associated with breast muscle weight, potentially indirectly or directly regulating breast muscle development ([He et al., 2022](#)).

Genes associated with reproductive trait

The reproductive traits in sheep arise from complex interactions among various genetic and physiological systems. Currently, Kirghiz sheep display a reproduction rate of 86.9%, with a notably low occurrence of multiparity. Efforts are underway to enhance multiparity by utilizing molecular breeding techniques to establish lines of multiparous sheep.

The PPP3CA gene, identified as protein phosphatase 3 catalytic subunit alpha, was initially discovered in brain tissue in 1979 and is known to play multiple essential roles in biological functions ([Tash et al., 1988](#)). Extensive research has investigated its contributions to cell growth, development, immune response, neurodevelopmental disorders, and spermatogenesis ([Rydzanicz et al., 2019](#)). In the field of live-stock production, PPP3CA stands out as a pivotal candidate gene associated with reproductive traits. Research on goats has firmly established the connection of the PPP3CA gene to litter size and semen quality, highlighting its significant role in governing reproductive traits in goats ([Bai et al., 2022](#)). Additionally, [Wang et al.'s \(2008\)](#) research demonstrated the active modulation of placental endothelial cell proliferation and signaling by PPP3CA. This regulatory function potentially influences the placental vascular system and blood flow, thereby affecting various reproductive processes.

The PDHA2 gene, known as alpha-2 pyruvate dehydrogenase, is responsible for synthesizing pyruvate dehydrogenase A2, a vital component of the pyruvate dehydrogenase complex. Studies across various mammalian species have revealed the involvement of PDHA2 in spermatogenesis. PDHA2 encodes the mitochondrial matrix enzyme exclusively expressed in the testis and prominently found in ejaculated spermatozoa ([Dahl et al., 1990](#)). In humans, the expression of the PDHA2 gene is specific to post-meiotic germ cells ([Pinheiro et al., 2012a](#)). Research on human spermatogenesis indicates that the expression of the PDHA2 gene plays a critical role in ensuring sustained protein expression, thereby supporting germ cell viability and function ([Pinheiro et al., 2012b](#)). In a study of hamster spermatozoa by [Kumar, Rangaraj & Shivaji \(2006\)](#), it was successfully identified that pyruvate dehydrogenase A2 undergoes tyrosine phosphorylation during hamster spermatogenesis. Reports suggest that PDHA2, a testis-specific phosphotyrosine in hamster sperm, is associated with the fibrous sheath and the energy acquisition process in hamster spermatozoa.

The NTRK2 gene, known as neurotrophic receptor tyrosine kinase 2, is a transmembrane receptor that exhibits a high affinity for brain-derived neurotrophic factor (BDNF) and neurotrophin-4/5 (NTF-4/5) ([Esmaeili-Fard et al., 2021](#)). In mice, Ntrk2 receptors play a crucial role in promoting follicular development by facilitating the formation of a functional FSH receptor. Research involving Ntrk2-null mice has highlighted the crucial role of NTRK2 in follicle assembly, maintaining their structural integrity, and supporting early follicle growth ([Kerr et al., 2009](#)). [Mirshokraei et al. \(2013\)](#) found significant expression of NTRK2 in the

oviduct and uterus of non-pregnant sheep, with elevated levels observed in the uterus of pregnant sheep. Moreover, [Chen et al. \(2012\)](#) detected distinct NTRK2 expression patterns in the ovaries of HU, indicating a potential correlation with off-season reproduction.

Genes associated with immune characteristics

This study also screened for a gene associated with immunity, and immunomodulation is essential in the response to resistance to pathogen invasion.

The GATA3 gene, also known as GATA binding protein 3, primarily regulates the development, proliferation, and maintenance of T cells. Variants and deletion mutants of this gene have been associated with various inflammatory diseases ([Wan, 2014](#)). Studies on blackface lambs infected with *Cyclothrrix circulans* have highlighted the regulatory role of GATA3 in T cell polarization and the development of immune responses ([Wilkie et al., 2016](#)). Furthermore, in chickens infected with the Infectious Bursal Disease Virus (IBDV), the GATA3 transcription factor inhibits viral replication by directly binding to the virus's promoter, effectively combating the IBDV infection ([Li et al., 2022a](#)).

CONCLUSIONS

Our study results have significant implications for the improvement plans of the KIR breed. By utilizing the Illumina Ovine SNP50 BeadChip genotype data, we identified selection signals and population structure in KIR sheep and other breeds known for their excellent wool characteristics and high prolificacy in harsh environments (HU and QBS), as well as sheep breeds known for their rapid growth and high-quality meat traits (DOR and SUF). These selection signals are closely related to adaptation to the Pamir Plateau environment, meat production characteristics, wool traits, and immune functions.

To confirm the specific roles of these genes in these important traits, further research is needed. These studies could include molecular biology functional validation and genome-wide association studies to evaluate the specific roles of these genes in sheep phenotypic traits.

These findings provide valuable insights for improving the KIR breed and may help local herders breed hybrid sheep that are well-suited to the Pamir Plateau environment and have excellent production characteristics. When using identified genetic markers for selective breeding, we can implement these strategies through marker-assisted selection (MAS). First, based on the identified genetic markers, we can define target traits such as meat quality, wool quality, and environmental adaptability. Second, using high-throughput genotyping technology, we can genotype candidate breeding individuals to identify those carrying favorable genetic markers. Finally, we can prioritize these individuals for breeding to increase the frequency of desirable traits in the next generation.

In future research, we plan to expand the sample size, use higher-density marker arrays, and further optimize selection signal methods to more comprehensively identify gene regions related to adaptability and production traits. Additionally, we will explore how environmental factors interact with genetic markers to influence adaptability and production traits, providing more scientific evidence for the improvement and diversity conservation of KIR sheep or other sheep breeds.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was funded by grants from the Natural Science Foundation of China (NO: 32060743), Bintuan Science and Technology Program (NO: 2022CB001-09). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Natural Science Foundation of China: 32060743.

Bintuan Science and Technology Program: 2022CB001-09.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Ruizhi Yang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Zhipeng Han conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Wen Zhou conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Xuejiao Li performed the experiments, prepared figures and/or tables, and approved the final draft.
- Xuechen Zhang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Lijun Zhu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Jieru Wang analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Xiaopeng Li conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Cheng-long Zhang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yahui Han performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Lianrui Li conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Shudong Liu conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This work has been reviewed and approved by the Tarim University Science and Technology Ethics Committee (Application Number: 2023039)

Data Availability

The following information was supplied regarding data availability:

The data is available at figshare: yang, ruizhi; Liu, Shu-dong (2024). The Illumina Ovine SNP50 BeadChip data for Kirghiz sheep and Qira Black sheep.. figshare. Thesis. <https://doi.org/10.6084/m9.figshare.25664511.v5>.

REFERENCES

- Abied A, Bagadi A, Bordbar F, Pu Y, Augustino SMA, Xue X, Xing F, Gebreselassie G, Han JL, Mwacharo JM, Ma Y, Zhao Q. 2020. Genomic diversity, population structure, and signature of selection in five Chinese native sheep breeds adapted to extreme environments. *Genes* 11(5):494 DOI 10.3390/genes11050494.
- Ablondi M, Eriksson S, Tetu S, Sabbioni A, Viklund Å, Mikko S. 2019. Genomic divergence in Swedish warmblood horses selected for Equestrian disciplines. *Genes* 10(12):976 DOI 10.3390/genes10120976.
- Agarwal M, Bharadwaj A, Mathew SJ. 2022. TLE4 regulates muscle stem cell quiescence and skeletal muscle differentiation. *Journal of Cell Science* 135(4):jcs256008 DOI 10.1242/jcs.256008.
- Akanno EC, Chen L, Abo-Ismael MK, Crowley JJ, Wang Z, Li C, Basarab JA, MacNeil MD, Plastow GS. 2018. Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. *Genetics Selection Evolution* 50(1):48 DOI 10.1186/s12711-018-0405-y.
- Alberto FJ, Boyer F, Orozco-terWengel P, Streeter I, Servin B, De Villemereuil P, Benjelloun B, Librado P, Biscarini F, Colli L, Barbato M, Zamani W, Alberti A, Engelen S, Stella A, Joost S, Ajmone-Marsan P, Negrini R, Orlando L, Rezaei HR, Naderi S, Clarke L, Flicek P, Wincker P, Coissac E, Kijas J, Tosser-Klopp G, Chikhi A, Bruford MW, Taberlet P, Pompanon F. 2018. Convergent genomic signatures of domestication in sheep and goats. *Nature Communications* 9(1):813 DOI 10.1038/s41467-018-03206-y.
- Alonso-García M, Suárez-Vega A, Fonseca PAS, Marina H, Pelayo R, Mateo J, Arranz JJ, Gutiérrez-Gil B. 2023. Transcriptome analysis of perirenal fat from Spanish Assaf suckling lamb carcasses showing different levels of kidney knob and channel fat. *Frontiers in Veterinary Science* 10:1150996 DOI 10.3389/fvets.2023.1150996.
- An B, Xu L, Xia J, Wang X, Miao J, Chang T, Song M, Ni J, Xu L, Zhang L, Li J, Gao H. 2020. Multiple association analysis of loci and candidate genes that regulate body size at three growth stages in Simmental beef cattle. *BMC Genetics* 21(1):32 DOI 10.1186/s12863-020-0837-6.

- Andrino S, Lorenzo V, Dunner S, Contreras E, Cañón J, Sevane N. 2022. Syringohydromyelia in dogs: the genomic component underlying a complex neurological disease. *Animals* 12(19):2622 DOI 10.3390/ani12192622.
- Animal Genetic Resources in China: Sheep and Goats. 2012. China National Commission of Animal Genetic Resources. 451.
- Bai Y, Zhang T, Liu N, Wang C, Guo Z, Pan C, Zhu H, Lan X. 2022. Investigation of copy number variations (CNVs) of the goat PPP3CA gene and their effect on litter size and semen quality. *Animals* 12(4):445 DOI 10.3390/ani12040445.
- Caiye Z, Song S, Li M, Huang X, Luo Y, Fang S. 2023. Genome-wide DNA methylation analysis reveals different methylation patterns in Chinese indigenous sheep with different type of tail. *Frontiers in Veterinary Science* 10:1125262 DOI 10.3389/fvets.2023.1125262.
- Cao Y, Song X, Shan H, Jiang J, Xiong P, Wu J, Shi F, Jiang Y. 2020. Genome-wide association study of body weights in hu sheep and population verification of related single-nucleotide polymorphisms. *Frontiers in Genetics* 11:588 DOI 10.3389/fgene.2020.00588.
- Cao YH, Xu SS, Shen M, Chen ZH, Gao L, Lv FH, Xie XL, Wang XH, Yang H, Liu CB, Zhou P, Wan PC, Zhang YS, Yang JQ, Pi WH, Hehua E, Berry DP, Barbato M, Esmailzadeh A, Nosrati M, Salehian-Dehkordi H, Dehghani-Qanatqestani M, Dotsev AV, Deniskova TE, Zinovieva NA, Brem G, Štěpánek O, Ciani E, Weimann C, Erhardt G, Mwacharo JM, Ahbara A, Han JL, Hanotte O, Miller JM, Sim Z, Coltman D, Kantanen J, Bruford MW, Lenstra JA, Kijas J, Li MH. 2021. Historical introgression from wild relatives enhanced climatic adaptation and resistance to pneumonia in sheep. *Molecular Biology and Evolution* 38(3):838–855 DOI 10.1093/molbev/msaa236.
- Caro CR, Nieves-Colón MA, Rawls E, Rubin-de Celis V, Lizárraga B, Vidaurre T, Sandoval K, Fejerman L, Stone AC, Moreno EA, Bosch E. 2022. Uncovering signals of positive selection in Peruvian populations from three ecological regions. *Molecular Biology and Evolution* 39(8):msac158 DOI 10.1093/molbev/msac158.
- Cavalera M, Axling U, Berger K, Holm C. 2016. Rose hip supplementation increases energy expenditure and induces browning of white adipose tissue. *Nutrition & Metabolism* 13:91 DOI 10.1186/s12986-016-0151-5.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:s13742-015-0047-8 DOI 10.1186/s13742-015-0047-8.
- Chen L, Liu K, Zhao Z, Blair HT, Zhang P, Li D, Ma RZ. 2012. Identification of sheep ovary genes potentially associated with off-season reproduction. *Journal of Genetics and Genomics* 39(4):181–190 DOI 10.1016/j.jgg.2012.03.002.
- Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, Butte NF. 2012. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLOS ONE* 7(12):e51954 DOI 10.1371/journal.pone.0051954.
- Dahl HHM, Brown RM, Hutchison WM, Maragos C, Brown GK. 1990. A testis-specific form of the human pyruvate dehydrogenase E1 alpha subunit is coded for by an

- intronless gene on chromosome 4. *Genomics* **8**(2):225–232
DOI [10.1016/0888-7543\(90\)90275-y](https://doi.org/10.1016/0888-7543(90)90275-y).
- De La Rosa J, Ostrowski J, Hryniewicz MM, Kredich NM, Kotb M, LeGros HL, Valentine M, Geller AM. 1995.** Chromosomal localization and catalytic properties of the recombinant alpha subunit of human lymphocyte methionine adenosyltransferase. *Journal of Biological Chemistry* **270**(37):21860–21868 DOI [10.1074/jbc.270.37.21860](https://doi.org/10.1074/jbc.270.37.21860).
- Dupain C, Gracia C, Harttrampf AC, Rivière J, Geoerger B, Massaad ML. 2019.** Newly identified LMO3-BORCS5 fusion oncogene in Ewing sarcoma at relapse is a driver of tumor progression. *Oncogene* **38**(47):7200–7215 DOI [10.1038/s41388-019-0914-3](https://doi.org/10.1038/s41388-019-0914-3).
- Dzomba EF, Van Der Nest MA, Mthembu JN, Soma P, Snyman MA, Chimonyo M, Muchadeyi FC. 2023.** Selection signature analysis and genome-wide divergence of South African Merino breeds from their founders. *Frontiers in Genetics* **13**:932272 DOI [10.3389/fgene.2022.932272](https://doi.org/10.3389/fgene.2022.932272).
- Ebert JR, Magi A, Unt E, Prans E, Wood DJ, Koks S. 2023.** Genome-wide association study identifying variants related to performance and injury in high-performance athletes. *Experimental Biology and Medicine* **248**(20):1799–1805 DOI [10.1177/15353702231198068](https://doi.org/10.1177/15353702231198068).
- Edea Z, Dadi H, Dessie T, Kim KS. 2019.** Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes Genomics* **41**(8):973–981 DOI [10.1007/s13258-019-00820-y](https://doi.org/10.1007/s13258-019-00820-y).
- Erkens T, Vandesompele J, Van Zeveren A, Peelman LJ. 2009.** Correlation between porcine PPARGC1A mRNA expression and its downstream target genes in backfat and longissimus dorsi muscle. *Journal of Applied Genetics* **50**(4):361–369 DOI [10.1007/BF03195694](https://doi.org/10.1007/BF03195694).
- Esmaili-Fard SM, Gholizadeh M, Hafezian SH, Abdollahi-Arpanahi R. 2021.** Genome-wide association study and pathway analysis identify NTRK2 as a novel candidate gene for litter size in sheep. *PLOS ONE* **16**(1):e0244408 DOI [10.1371/journal.pone.0244408](https://doi.org/10.1371/journal.pone.0244408).
- Eydivandi S, Roudbar MA, Ardestani SS, Momen M, Sahana G. 2021.** A selection signatures study among Middle Eastern and European sheep breeds. *Journal of Animal Breeding and Genetics* **138**(5):574–588 DOI [10.1111/jbg.12536](https://doi.org/10.1111/jbg.12536).
- Fantauzzo KA, Kurban M, Levy B, Christiano AM. 2012.** Trps1 and its target gene Sox9 regulate epithelial proliferation in the developing hair follicle and are associated with hypertrichosis. *PLOS Genetics* **8**(11):e1003002 DOI [10.1371/journal.pgen.1003002](https://doi.org/10.1371/journal.pgen.1003002).
- Fantauzzo KA, Tadin-Strapps M, You Y, Mentzer SE, Baumeister FA, Cianfarani S, Van Maldergem L, Warburton D, Sundberg JP, Christiano AM. 2008.** A position effect on TRPS1 is associated with Ambras syndrome in humans and the Koala phenotype in mice. *Human Molecular Genetics* **17**(22):3539–3551 DOI [10.1093/hmg/ddn247](https://doi.org/10.1093/hmg/ddn247).
- Fariello MI, Servin B, Tosser-Klopp G, Rupp R, Moreno C, International Sheep Genomics Consortium, San Cristobal M, Boitard S. 2014.** Selection signatures in worldwide sheep populations. *PLOS ONE* **9**(8):e103813 DOI [10.1371/journal.pone.0103813](https://doi.org/10.1371/journal.pone.0103813).

- Ferrara N, Gerber HP, LeCouter J. 2003. The biology of VEGF and its receptors. *Nature Medicine* 9(6):669–676 DOI 10.1038/nm0603-669.
- Fisher AL, Caudy M. 1998. Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. *Genes & Development* 12(13):1931–1940 DOI 10.1101/gad.12.13.1931.
- Gascuel O, Steel M. 2006. Neighbor-joining revealed. *Molecular Biology and Evolution* 23(11):1997–2000 DOI 10.1093/molbev/msl072.
- Granot Y, Tal O, Rosset S, Skorecki K. 2016. On the apportionment of population structure. *PLOS ONE* 11(8):e0160413 DOI 10.1371/journal.pone.0160413.
- He Y, Shi H, Li Z, Kang J, Li M, Liu M, Liu Y, Zhao J, Dou T, Jia J, Duan Y, Wang K, Ge C. 2022. Identification of new genes and genetic variant loci associated with breast muscle development in the Mini-Cobb F2 chicken population using a genome-wide association study. *Genes* 13(11):2153 DOI 10.3390/genes13112153.
- Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, Thorleifsson G, Zillikens MC, Speliotes EK, Mägi R, Workalemahu T, White CC, Bouatia-Naji N, Harris TB, Berndt SI, Ingelsson E, Willer CJ, Weedon MN, Luan J, Vedantam S, Esko T, Kilpeläinen TO, Kutalik Z, Li S, Monda KL, Dixon AL, Holmes CC, Kaplan LM, Liang L, Min JL, Moffatt MF, Molony C, Nicholson G, Schadt EE, Zondervan KT, Feitosa MF, Ferreira T, Allen HL, Weyant RJ, Wheeler E, Wood AR, MAGIC, Estrada K, Goddard ME, Lettre G, Mangino M, Nyholt DR, Purcell S, Smith AV, Visscher PM, Yang J, McCarroll SA, Nemesh J, Voight BF, Absher D, Amin N, Aspelund T, Coin L, Glazer NL, Hayward C, Heard-Costa NL, Hottenga JJ, Johansson Å, Johnson T, Kaakinen M, Kapur K, Ketkar S, Knowles JW, Kraft P, Kraja AT, Lamina C, Leitzmann MF, McKnight B, Morris AP, Ong KK, Perry JRB, Peters MJ, Polasek O, Prokopenko I, Rayner NW, Ripatti S, Rivadeneira F, Robertson NR, Sanna S, Sovio U, Surakka I, Teumer A, van Wingerden S, Vitart V, Zhao JH, Cavalcanti-Proença C, Chines PS, Fisher E, Kulzer JR, Lecoeur C, Narisu N, Sandholt C, Scott LJ, Silander K, Stark K, Tammesoo M-L, Teslovich TM, Timpson NJ, Watanabe RM, Welch R, Chasman DI, Cooper MN, Jansson J-O, Kettunen J, Lawrence RW, Pellikka N, Perola M, Vandenput L, Alavere H, Almgren P, Atwood LD, Bennett AJ, Biffar R, Bonnycastle LL, Bornstein SR, Buchanan TA, Campbell H, Day INM, Dei M, Dörr M, Elliott P, Erdos MR, Eriksson JG, Freimer NB, Fu M, Gagel S, Geus EJC, Gjesing AP, Grallert H, Gräßler J, Groves CJ, Guiducci C, Hartikainen A-L, Hassanali N, Havulinna AS, Herzig K-H, Hicks AA, Hui J, Igl W, Jousilahti P, Jula A, Kajantie E, Kinnunen L, Kolcic I, Koskinen S, Kovacs P, Kroemer HK, Krzely V, Kuusisto J, Kvaloy K, Laitinen J, Lantieri O, Lathrop GM, Lokki M-J, Luben RN, Ludwig B, McArdle WL, McCarthy A, Morken MA, Nelis M, Neville MJ, Paré G, Parker AN, Peden JF, Pichler I, Pietiläinen KH, Platou CGP, Pouta A, Ridderstråle M, Samani NJ, Saramies J, Sinisalo J, Smit JH, Strawbridge RJ, Stringham HM, Swift AJ, Teder-Laving M, Thomson B, Usala G, van Meurs JBJ, van Ommen G-J, Vatin V, Volpato CB, Wallaschofski H, Walters GB, Widen E, Wild SH, Willemsen G, Witte DR, Zgaga L, Zitting P, Beilby JP, James AL, Kähönen M, Lehtimäki T, Nieminen MS, Ohlsson C, Palmer

- LJ, Raitakari O, Ridker PM, Stumvoll M, Tönjes A, Viikari J, Balkau B, Ben-Shlomo Y, Bergman RN, Boeing H, Smith GD, Ebrahim S, Froguel P, Hansen T, Hengstenberg C, Hveem K, Isomaa B, Jørgensen T, Karpe F, Khaw K-T, Laakso M, Lawlor DA, Marre M, Meitinger T, Metspalu A, Midthjell K, Pedersen O, Salomaa V, Schwarz PEH, Tuomi T, Tuomilehto J, Valle TT, Wareham NJ, Arnold AM, Beckmann JS, Bergmann S, Boerwinkle E, Boomsma DI, Caulfield MJ, Collins FS, Eiriksdottir G, Gudnason V, Gyllenstein U, Hamsten A, Hattersley AT, Hofman A, Hu FB, Illig T, Iribarren C, Jarvelin M-R, Linda Kao WH, Kaprio J, Launer LJ, Munroe PB, Oostra B, Penninx BW, Pramstaller PP, Psaty BM, Quertermous T, Rissanen A, Rudan I, Shuldiner AR, Soranzo N, Spector TD, Syvanen A-C, Uda M, Uitterlinden A, Völzke H, Vollenweider P, Wilson JF, Witteman JC, Wright AF, Abecasis GR, Boehnke M, Borecki IB, Deloukas P, Frayling TM, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, North KE, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, Hirschhorn JN, Assimes TL, Wichmann H-E, Thorsteinsdottir U, van Duijn CM, Stefansson K, Cupples LA, Loos RJF, Barroso I, McCarthy MI, Fox CS, Mohlke KL, Lindgren CM. 2010. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics* 42(11):949–960 DOI 10.1038/ng.685.
- Heller-Trulli D, Liu H, Mukherjee S, Moore CL. 2022. UBE3D regulates mRNA 3'-end processing and maintains adipogenic potential in 3T3-L1 cells. *Molecular and Cellular Biology* 42(12):e0017422 DOI 10.1128/mcb.00174-22.
- Hill WG. 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33(2):229–239 DOI 10.1038/hdy.1974.89.
- Hu S, Zhu L. 2018. Semaphorins and their receptors: from axonal guidance to atherosclerosis. *Frontiers in Physiology* 9:1236 DOI 10.3389/fphys.2018.01236.
- Hui Z, Di R, Wang F-Y, Wang X-Yu, He X-Y, Liu Y-F, Chu M-X. 2022. Association analysis of GDF 9 and BMP15 gene polymorphisms and litter size in Qira black sheep. *Journal of Anhui Agricultural University* 49(2):265–271 (In Chinese).
- Jin M, Lu J, Fei X, Lu Z, Quan K, Liu Y, Chu M, Di R, Wang H, Wei C. 2020. Genetic signatures of selection for cashmere traits in Chinese Goats. *Animals* 10(10):1905 DOI 10.3390/ani10101905.
- Jolliffe IT, Cadima J. 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A* 374(2065):20150202 DOI 10.1098/rsta.2015.0202.
- Jones B, Li C, Park MS, Lerch A, Jacob V, Johnson N, Kuang JQ, Dhall S, Sathyaamoorthy M. 2021. Comprehensive comparison of amnion stromal cells and chorion stromal cells by RNA-Seq. *International Journal of Molecular Sciences* 22(4):1901 DOI 10.3390/ijms22041901.
- Jun G, Asai H, Zeldich E, Drapeau E, Chen C, Chung J, Park JH, Kim S, Haroutunian V, Foroud T, Kuwano R, Haines JL, Pericak-Vance MA, Schellenberg GD, Lunetta KL, Kim JW, Buxbaum JD, Mayeux R, Ikezu T, Abraham CR, Farrer LA. 2014. PLXNA4 is associated with Alzheimer disease and modulates tau phosphorylation. *Annals of Neurology* 76(3):379–392 DOI 10.1002/ana.24219.

- Kajii TS, Oka A, Hatta M, Yamazaki J, Yamashita J, Iida J. 2018. PLXNA2 identified as a candidate gene by genome-wide association analysis for mandibular prognathism in human chondrocytes. *Biomedical Reports* 9(3):253–258 DOI 10.3892/br.2018.1128.
- Kalds P, Zhou S, Gao Y, Cai B, Huang S, Chen Y, Wang X. 2022. Genetics of the phenotypic evolution in sheep: a molecular look at diversity-driving genes. *Genetics Selection Evolution* 54(1):61 DOI 10.1186/s12711-022-00753-3.
- Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28(1):27–30 DOI 10.1093/nar/28.1.27.
- Kerr B, Garcia-Rudaz C, Dorfman M, Paredes A, Ojeda SR. 2009. NTRK1 and NTRK2 receptors facilitate follicle assembly and early follicular development in the mouse ovary. *Reproduction* 138(1):131–140 DOI 10.1530/REP-08-0474.
- Kijas JW. 2014. Haplotype-based analysis of selective sweeps in sheep. *Genome* 57(8):433–437 DOI 10.1139/gen-2014-0049.
- Kizilaslan M, Arzik Y, White SN, Piel LMW, Cinar MU. 2022. Genetic parameters and genomic regions underlying growth and linear type traits in Akkaraman sheep. *Genes* 13(8):1414 DOI 10.3390/genes13081414.
- Koshkina O, Deniskova T, Dotsev A, Kunz E, Selionova M, Medugorac I, Zinovieva N. 2023. Phylogenetic analysis of Russian native sheep breeds based on mtDNA sequences. *Genes* 14(9):1701 DOI 10.3390/genes14091701.
- Krivoruchko A, Surov A, Skokova A, Kanibolotskaya A, Saprikina T, Kukharuk M, Yatsyk O. 2022. A genome-wide search for candidate genes of meat production in Jalgin Merino considering known productivity genes. *Genes* 13(8):1337 DOI 10.3390/genes13081337.
- Kuehn LA, Lewis RM, Notter DR. 2009. Connectedness in Targhee and Suffolk flocks participating in the United States National Sheep Improvement Program. *Journal of Animal Science* 87(2):507–515 DOI 10.2527/jas.2008-1092.
- Kumar V, Rangaraj N, Shivaji S. 2006. Activity of pyruvate dehydrogenase A (PDHA) in hamster spermatozoa correlates positively with hyperactivation and is associated with sperm capacitation. *Biology of Reproduction* 75(5):767–777 DOI 10.1095/biolreprod.106.053587.
- Lee KY, Sharma R, Gase G, Ussar S, Li Y, Welch L, Berryman DE, Kispert A, Bluher M, Kahn CR. 2017. Tbx15 defines a glycolytic subpopulation and white adipocyte heterogeneity. *Diabetes* 66(11):2822–2829 DOI 10.2337/db17-0218.
- Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution* 32(10):2798–800 DOI 10.1093/molbev/msv150.
- Lei Z, Sun W, Guo T, Li J, Zhu S, Lu Z, Qiao G, Han M, Zhao H, Yang B, Zhang L, Liu J, Yuan C, Yue Y. 2021. Genome-wide selective signatures reveal candidate genes associated with hair follicle development and wool shedding in sheep. *Genes* 12(12):1924 DOI 10.3390/genes12121924.
- Li Y, Chen Z, Fang Y, Cao C, Zhang Z, Pan Y, Wang Q. 2022b. Runs of homozygosity revealed reproductive traits of hu sheep. *Genes* 13(10):1848 DOI 10.3390/genes13101848.

- Li J, Haiyilati A, Zhou L, Chen J, Wang Y, Gao L, Cao H, Li X, Zheng SJ. 2022a. GATA3 inhibits viral infection by promoting microRNA-155 expression. *Journal of Virology* 96(7):e0188821 DOI 10.1128/jvi.01888-21.
- Li Y, Wu DD, Boyko AR, Wang GD, Wu SF, Irwin DM, Zhang YP. 2014. Population variation revealed high-altitude adaptation of Tibetan mastiffs. *Molecular Biology and Evolution* 31(5):1200–1205 DOI 10.1093/molbev/msu070.
- Liu Y, Cheng H, Wang S, Luo X, Ma X, Sun L, Chen N, Zhang J, Qu K, Wang M, Liu J, Huang B, Lei C. 2022. Genomic diversity and selection signatures for wein-ing cattle on the border of Yunnan-Guizhou. *Frontiers in Genetics* 13:848951 DOI 10.3389/fgene.2022.848951.
- Liu CT, Monda KL, Taylor KC, Lange L, Demerath EW, Palmas W, Wojczynski MK, Ellis JC, Vitolins MZ, Liu S, Papanicolaou GJ, Irvin MR, Xue L, Griffin PJ, Nalls MA, Adeyemo A, Liu J, Li G, Ruiz-Narvaez EA, Chen W-M, Chen F, Henderson BE, Millikan RC, Ambrosone CB, Strom SS, Guo X, Andrews JS, Sun YV, Mosley TH, Yanek LR, Shriner D, Haritunians T, Rotter JI, Speliotes EK, Smith M, Rosenberg L, Mychaleckyj J, Nayak U, Spruill I, Timothy Garvey W, Pettaway C, Nyante S, Bandera EV, Britton AF, Zonderman AB, Rasmussen-Torvik LJ, Ida Chen Y-D, Ding J, Lohman K, Kritchevsky CB, Zhao W, Peyser PA, Kardina SLR, Kabagambe E, Broeckel U, Chen G, Zhou J, Wassertheil-Smoller S, Neuhaus ML, Rumpersaud E, Psaty B, Kooperberg C, Manson JE, Kuller LH, Ochs-Balcom HM, Johnson KC, Sucheston L, Ordovas JM, Palmer JR, Haiman CA, McKnight B, Howard BV, Becker DM, Bielak LF, Liu Y, Allison MA, Grant SFA, Burke GL, Patel SR, Schreiner PJ, Borecki IB, Evans MK, Taylor H, Sale MM, Howard V, Carlson CS, Rotimi CN, Cushman M, Harris TB, Reiner AP, Adrienne Cupples L, North KE, Fox CS. 2013. Genome-wide association of body fat distribution in African ancestry populations suggests new loci. *PLOS Genetics* 9(8):e1003681 DOI 10.1371/journal.pgen.1003681.
- Lv FH, Cao YH, Liu GJ, Luo LY, Lu R, Liu MJ, Li WR, Zhou P, Wang XH, Shen M, Gao L, Yang J-Q, Yang H, Yang Y-L, Liu C-B, Wan P-C, Zhang Y-S, Pi W-H, Ren Y-L, Shen Z-Q, Wang F, Wang Y-T, Li J-Q, Salehian-Dehkordi H, Hehua E, Liu Y-G, Chen J-F, Wang J-K, Deng X-M, Esmailizadeh A, Dehghani-Qanatqestani M, Charati H, Nosrati M, Štěpánek O, Rushdi HE, Olsaker I, Curik I, Gorkhali NA, Paiva SR, Caetano AR, Ciani E, Amills M, Weimann C, Erhardt G, Amane A, Mwacharo JM, Han J-L, Hanotte O, Periasamy K, Johansson AM, Hallsson JH, Kantanen J, Coltman DW, Bruford MW, Lenstra JA, Li M-H. 2022a. Whole-genome resequencing of worldwide wild and domestic sheep elucidates genetic diversity, introgression, and agronomically important loci. *Molecular Biology and Evolution* 39(2):msab353 DOI 10.1093/molbev/msab353.
- Lv X, Chen W, Wang S, Cao X, Yuan Z, Getachew T, Mwacharo JM, Haile A, Sun W. 2022b. Integrated hair follicle profiles of microRNAs and mRNAs to reveal the pattern formation of hu sheep lambskin. *Genes* 13(2):342 DOI 10.3390/genes13020342.

- Ma L, Li Z, Cai Y, Xu H, Yang R, Lan X. 2018. Genetic variants in fat- and short-tailed sheep from high-throughput RNA-sequencing data. *Animal Genetics* 49(5):483–487 DOI 10.1111/age.12699.
- Metrak M, Sulwinski M, Chachulski L, Wilk M, Suska-Malawska M. 2015. Creeping environmental problems in the Pamir Mountains: landscape conditions, climate change, wise use and threats. In: Öztürk M, Hakeem K, Faridah-Hanum I, Efe R, eds. *Climate change impacts on high-altitude ecosystems*. Cham: Springer, 665–694 DOI 10.1007/978-3-319-12859-7_28.
- Mirshokraei P, Hassanpour H, Rahnama A, Foster WG. 2013. Gene expression of BDNF and its receptors, TrkB and p75 in the uterus and oviduct of pregnant and non-pregnant ewes. *Research in Veterinary Science* 95(1):164–168 DOI 10.1016/j.rvsc.2013.03.009.
- Montana G, Hoggart C. 2007. Statistical software for gene mapping by admixture linkage disequilibrium. *Briefings in Bioinformatics* 8(6):393–395 DOI 10.1093/bib/bbm035.
- Moradi MH, Nejati-Javaremi A, Moradi-Shahrbabak M, Dodds KG, Brauning R, McEwan JC. 2022. Hitchhiking mapping of candidate regions associated with fat deposition in Iranian thin and fat tail sheep breeds suggests new insights into molecular aspects of fat tail selection. *Animals* 12(11):1423 DOI 10.3390/ani12111423.
- Morell Miranda P, Soares AER, Günther T. 2023. Demographic reconstruction of the Western sheep expansion from whole-genome sequences. *G3* 13(11):jkad199 DOI 10.1093/g3journal/jkad199.
- Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* 76(10):5269–5273 DOI 10.1073/pnas.76.10.5269.
- Nosrati M, Asadollahpour Nanae H, Amiri Ghanatsaman Z, Esmailizadeh A. 2019. Whole genome sequence analysis to detect signatures of positive selection for high fecundity in sheep. *Reproduction in Domestic Animals* 54(2):358–364 DOI 10.1111/rda.13368.
- Notter DR. 1998. The U.S. National Sheep Improvement Program: across-flock genetic evaluations and new trait development. *Journal of Animal Science* 76(9):2324–2330 DOI 10.2527/1998.7692324x.
- Oleksyk TK, Smith MW, O’Brien SJ. 2010. Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B* 365(1537):185–205 DOI 10.1098/rstb.2009.0219.
- Ombayev A, Parzhanov Z, Azhimetov N, Zhylykibayev A, Abishov M, Issabayeva A. 2024. Increasing the meat productivity of young sheep based on the use of the gene pool of the Dorper and Hissar breeds. *Brazilian Journal of Biology* 83:e278807 DOI 10.1590/1519-6984.278807.
- Pan ZY, He XY, Wang XY, Guo XF, Cao XH, Hu WP, R Di, Liu QY, Chu MX. 2016. Selection signature in domesticated animals. 1069–1080.

- Pinheiro A, Nunes MJ, Milagre I, Rodrigues E, Silva MJ, De Almeida IT, Rivera I. 2012a. Demethylation of the coding region triggers the activation of the human testis-specific PDHA2 gene in somatic tissues. *PLOS ONE* 7(6):e38076 DOI 10.1371/journal.pone.0038076.
- Pinheiro A, Silva MJ, Graça I, Silva J, Sá R, Sousa M, Barros A, De Almeida IT, Rivera I. 2012b. Pyruvate dehydrogenase complex: mRNA and protein expression patterns of E1 α subunit genes in human spermatogenesis. *Gene* 506(1):173–178 DOI 10.1016/j.gene.2012.06.068.
- Puigserver P, Wu Z, Park CW, Graves R, Wright M, Spiegelman BM. 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* 92(6):829–839 DOI 10.1016/s0092-8674(00)81410-5.
- Razzaq M, Iglesias MJ, Ibrahim-Kosta M, Goumide L, Soukariéh O, Proust C, Roux M, Suchon P, Boland A, Daiaín D, Olaso R, Havervall S, Thalin C, Butler L, Deleuze JF, Odeberg J, Morange PE, Trégouët DA. 2021. An artificial neural network approach integrating plasma proteomics and genetic data identifies PLXNA4 as a new susceptibility locus for pulmonary embolism. *Scientific Reports* 11(1):14015 DOI 10.1038/s41598-021-93390-7.
- R Core Team. 2023. R: a language and environment for statistical computing. Vienna, Austria: The R Foundation for Statistical Computing. Available at <https://www.r-project.org/>.
- Rydzanicz M, Wachowska M, Cook EC, Lisowski P, Kuźniewska B, Szymańska K, Diecke S, Prigione A, Szczaluba K, Szybińska A, Koppolu A, Pienkowski VM, Kosińska J, Wiweger M, Kostrzewa G, Brzozowska M, Domańska-Pakieła D, Jurkiewicz E, Stawiński P, Gromadka A, Zielenkiewicz P, Demkow U, Dziembowska M, Kuźnicki J, Creamer TP, Płoski R. 2019. Novel calcineurin A (PPP3CA) variant associated with epilepsy, constitutive enzyme activation and downregulation of protein expression. *European Journal of Human Genetics* 27(1):61–69 DOI 10.1038/s41431-018-0254-8.
- Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419(6909):832–837 DOI 10.1038/nature01140.
- Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, Huentelman MJ, Craig DW, DeChairo BM, Potkin SG, Jack CR, Weiner MW, Saykin AJ, the Alzheimer’s Disease Neuroimaging Initiative. 2010. Alzheimer’s disease neuroimaging initiative, whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. *Neuroimage* 53(3):1051–1063 DOI 10.1016/j.neuroimage.2010.01.042.
- Shriner D. 2023. Overview of admixture mapping. *Current Protocols* 3(2):e677 DOI 10.1002/0471142905.hg0123s76.

- Singer JB. 2009. Candidate gene association analysis. *Methods in Molecular Biology* 573:223–30 DOI 10.1007/978-1-60761-247-6_13.
- Sjöstrand AE, Sjödin P, Jakobsson M. 2014. Private haplotypes can reveal local adaptation. *BMC Genetics* 15:61 DOI 10.1186/1471-2156-15-61.
- Suzuki Y. 2010. Statistical methods for detecting natural selection from genomic data. *Genes & Genetic Systems* 85(6):359–376 DOI 10.1266/ggs.85.359.
- Takeshita M, Yamada K, Hattori E, Iwayama Y, Toyota T, Iwata Y, Tsuchiya KJ, Sugihara G, Hashimoto K, Watanabe H, Iyo M, Kikuchi M, Okazaki Y, Yoshikawa T. 2008. Genetic examination of the PLXNA2 gene in Japanese and Chinese people with schizophrenia. *Schizophrenia Research* 99(1–3):359–364 DOI 10.1016/j.schres.2007.11.002.
- Tang H, Peng J, Wang P, Risch NJ. 2005. Estimation of individual admixture: analytical and study design considerations. *Genetic Epidemiology* 28(4):289–301 DOI 10.1002/gepi.20064.
- Tash JS, Krinks M, Patel J, Means RL, Klee CB, Means AR. 1988. Identification, characterization, and functional correlation of calmodulin-dependent protein phosphatase in sperm. *Journal of Cell Biology* 106(5):1625–1633 DOI 10.1083/jcb.106.5.1625.
- Thomas PD. 2017. The gene ontology and the meaning of biological function. *Methods in Molecular Biology* 1446:15–24 DOI 10.1007/978-1-4939-3743-1_2.
- Vallée Marcotte B, Guénard F, Cormier H, Lemieux S, Couture P, Rudkowska I, Vohl MC. 2017. Plasma triglyceride levels may be modulated by gene expression of IQCJ, NXPH1, PHF17 and MYB in humans. *International Journal of Molecular Sciences* 18(2):257 DOI 10.3390/ijms18020257.
- Wan YY. 2014. GATA3: a master of many trades in immune regulation. *Trends in Immunology* 35(6):233–242 DOI 10.1016/j.it.2014.04.002.
- Wang YL, Hou YH, Ling ZJ, Zhao HL, Zheng XR, Zhang XD, Yin ZJ, Ding YY. 2023. RNA sequencing analysis of the longissimus dorsi to identify candidate genes underlying the intramuscular fat content in Anqing Six-end-white pigs. *Animal Genetics* 54(3):315–327 DOI 10.1111/age.13308.
- Wang K, Song Y, Chen DB, Zheng J. 2008. Protein phosphatase 3 differentially modulates vascular endothelial growth factor and fibroblast growth factor 2-stimulated cell proliferation and signaling in ovine fetoplacental artery endothelial cells. *Biology of Reproduction* 79(4):704–710 DOI 10.1095/biolreprod.108.068957.
- Wang Z, Zhang H, Yang H, Wang S, Rong E, Pei W, Li H, Wang N. 2014. Genome-wide association study for wool production traits in a Chinese Merino sheep population. *PLOS ONE* 9(9):e107101 DOI 10.1371/journal.pone.0107101.
- Wijayanti D, Erdenee S, Akhatayeva Z, Li H, Li J, Cai Y, Jiang F, Xu H, Lan X. 2022. Genetic polymorphisms within the ETAA1 gene associated with growth traits in Chinese sheep breeds. *Animal Genetics* 53(3):460–465 DOI 10.1111/age.13197.
- Wilkie H, Gossner A, Bishop S, Hopkins J. 2016. Variations in T cell transcription factor sequence and expression associated with resistance to the sheep nematode teladorsagia circumcincta. *PLOS ONE* 11(2):e0149644 Erratum in: PLoS One. 11 (3) (2016) e0152396 DOI 10.1371/journal.pone.0149644.

- Worku T, Wang K, Ayers D, Wu D, Ur Rehman Z, Zhou H, Yang L. 2018. Regulatory roles of ephrinA5 and its novel signaling pathway in mouse primary granulosa cell apoptosis and proliferation. *Cell Cycle* 17(7):892–902 DOI 10.1080/15384101.2018.1456297.
- Wray NR, James MR, Mah SP, Nelson M, Andrews G, Sullivan PF, Montgomery GW, Birley AJ, Braun A, Martin NG. 2007. Anxiety and comorbid measures associated with PLXNA2. *Archives Of General Psychiatry* 64(3):318–326 DOI 10.1001/archpsyc.64.3.318.
- Yamamoto Y, Gesta S, Lee KY, Tran TT, Saadatirad P, Kahn CR. 2010. Adipose depots possess unique developmental gene signatures. *Obesity* 18(5):872–878 Erratum in: *Obesity* (Silver Spring). 2010 May;18(5):1064 DOI 10.1038/oby.2009.512.
- Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics* 88(1):76–82 DOI 10.1016/j.ajhg.2010.11.011.
- Young CC. 1913. Strange sheep of Asiatic Russia: broad-tailed and fat-rumped creatures among hardiest domestic animals peculiar results of hybridization-first national sheep congress at Moscow. *Journal of Heredity* 4(4):184–192 DOI 10.1093/oxfordjournals.jhered.a106854.
- Zhang C, Dong SS, Xu JY, He WM, Yang TL. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35(10):1786–1788 DOI 10.1093/bioinformatics/bty875.
- Zhang X, Liu C, Kong Y, Li F, Yue X. 2022. Effects of intramuscular fat on meat quality and its regulation mechanism in Tan sheep. *Frontiers in Nutrition* 9:908355 DOI 10.3389/fnut.2022.908355.
- Zhang CL, Zhang J, Tuersuntuoheti M, Chang Q, Liu S. 2023a. Population structure, genetic diversity and prolificacy in pishan red sheep under an extreme desert environment. *Frontiers in Genetics* 14:1092066 DOI 10.3389/fgene.2023.1092066.
- Zhang L, Zhang S, Zhan F, Song M, Shang P, Zhu F, Li J, Yang F, Li X, Qiao R, Han X, Li X, Liu G, Wang K. 2023b. Population genetic analysis of six Chinese indigenous pig meta-populations based on geographically isolated regions. *Animals* 13(8):1396 DOI 10.3390/ani13081396.
- Zhao C, Chen X, Wu W, Wang W, Pang W, Yang G. 2016. MAT2B promotes adipogenesis by modulating SAME levels and activating AKT/ERK pathway during porcine intramuscular preadipocyte differentiation. *Experimental Cell Research* 344(1):11–21 DOI 10.1016/j.yexcr.2016.02.019.
- Zhao F, Deng T, Shi L, Wang W, Zhang Q, Du L, Wang L. 2020. Genomic scan for selection signature reveals fat deposition in Chinese Indigenous sheep with extreme tail types. *Animals* 10(5):773 DOI 10.3390/ani10050773.
- Zhong Z, Wang Z, Xie X, Tian S, Wang F, Wang Q, Ni S, Pan Y, Xiao Q. 2023. Evaluation of the genetic diversity, population structure and selection signatures of three native Chinese pig populations. *Animals* 13(12):2010 DOI 10.3390/ani13122010.

Zhou T, Xu K, Zhao F, Liu W, Li L, Hua Z, Zhou X. 2023. itol.toolkit accelerates working with iTOL (Interactive Tree of Life) by an automated generation of annotation files. *Bioinformatics* **39**(6):btad339 DOI [10.1093/bioinformatics/btad339](https://doi.org/10.1093/bioinformatics/btad339).