



OPEN Genome-wide analysis provides insight into the genetic diversity and adaptability of Kazakhstan local goats

Nelly Kichamu^{1,2,3}, George Wanjala^{4,5}, Kairat Dossybayev^{6,7,8}, Zoltán Bagi¹, Bakhytzhhan Bekmanov^{6,8} & Szilvia Kusza¹✉

To understand the genetic structure of local goat populations in Kazakhstan, we present the first comprehensive genomic analysis using a 70 K SNP array, examining 120 samples from six regions. The populations (Kundyzdy, Darbaza, Shokpar, Ushterek, Kenes and Kosseit) are named based on their breeding regions and are referred to ecotypes. Our results present a high genetic diversity, with observed heterozygosity (H_o) ranging from 0.389 to 0.444. Shokpar and Kundyzdy showed excess heterozygosity, while Kosseit and Darbaza had high inbreeding levels. Most populations recorded weak genetic differentiation, except for Ushterek goats, which were more distinct. For adaptability, Runs of Homozygosity (ROH) and Integrated Haplotype Score (iHS) identified 71 genes under selection, with gene ontology enrichment indicating localization as the most represented biological pathway important for cell response. Additionally, over 60% of these genes are co-expressed, suggesting their significant role in survival of these goats in their local environments. Some identified genes, such as *NLRC4*, *HCLS1*, *IL17D*, *IL17RE*, and *IL17RC*, were linked to immune response, an important adaptation trait in the context of climate change. This study not only emphasizes on high genetic variability of local goats in Kazakhstan but also lays the groundwork for developing conservation and breeding programs for local goats in Kazakhstan and other similar agroecological regions for adaptability and resilience under changing environmental conditions.

Keywords Genetic diversity, Kazakhstan, Local goats, Population structure, Signature of selection

Goats (*Capra hircus*) are said to be descended from the wild bezoar goat (*Capra aegagrus*), which lived within the Fertile Crescent of western Asia (the Iranian region) around 10,000 years ago^{1–4}. As humans moved from one place to another in search of fertile places to settle and trade, they brought the goats with them². This movement (migration) led to the emergence of sub-populations and breeds that can be distinguished by different genetic traits and phenotypic expressions². A prolonged stay in new habitats and environments led to adaptations as they experienced various natural selection pressures, this resulted to the development of adaptive genes which enabled them to thrive in these regions. The adaptation process yielded different breeds, sometimes distinguished phenotypically, while other breeds are phenotypically similar but can only be distinguished by molecular data analysis⁵. These variations can be identified by doing signature selection analysis, which highlights genomic regions under natural or artificial selection and provide insights into adaptive evolution and domestication-related traits⁶. For instance, studies on selection signature have identified genes linked to economically important traits, including *HOXA12*, *PPP1CC*, and *LCORL*, which influence morphology, reproduction, and

¹Centre for Agricultural Genomics and Biotechnology, University of Debrecen, Debrecen 4032, Hungary. ²Doctoral School of Animal Science, University of Debrecen, Debrecen 4032, Hungary. ³Ministry of Agriculture Livestock, Fisheries and Cooperatives, Directorate of Livestock Development, Sheep and Goats Breeding and conservation Station, P.O. Box 254- 20117, Naivasha, Kenya. ⁴Institute of Animal Sciences and Wildlife Management, University of Szeged, Andrassy út 15, Hódmezővásárhely 6800, Hungary. ⁵Livestock, Fisheries, Cooperatives and Irrigation, Directorate of Livestock Production, Ministry of Agriculture, P.O. Box 437-50200, Livestock, Bungoma, Kenya. ⁶Institute of Genetics and Physiology, CS MSHE RK, Almaty 050060, Kazakhstan. ⁷Kazakh Research Institute of Livestock and Fodder Production, Almaty 050035, Kazakhstan. ⁸Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan. ✉email: kusza@agr.unideb.hu

environmental adaptation in some livestock species^{7,8}. Currently, the goats have become special livestock species in terms of their importance and adaptation to the environments in which they are produced⁹.

Indigenous goats specifically are preferred due to their unique adaptive features that can withstand various ecological zones^{9,10}. They are useful domestic livestock species for rural households to exploit because of their body size, productivity levels, food preferences, and low investment costs^{11,12} and hence they contribute to the livelihood improvement of the rural households by acting as a source of food security and income generation^{10,13}. However, the growing human population and the rising demand for livestock products have impacted the genetic diversity of indigenous goats through crossbreeding and the introduction of exotic genotypes¹⁴.

Kazakhstan, a country that covers the major part of Eurasian Steppe and also situated near the early domestication centers, presents a unique opportunity for studying genetic variability in goats¹⁵. Besides this, the varied landscape that contains arid, deep valleys and mountainous regions has likely shaped the unique genetic adaptations in local goat populations. This geographical position at the junction of historical migration routes indicates the possible emergence of unique adaptive features in this goat populations¹⁶. Despite their significance, the local goat populations in Kazakhstan have not been characterized, as they are named after the location or community, they originate from rather than having a specific breed name. Recent population trends in Kazakhstan's goat sector highlight the urgency of genetic characterization and conservation. Between 2020 and 2024, the national goat population showed fluctuations in numbers, initially decreasing from 2.31 million to 2.32 million head (2020–2022), followed by a recovery phase¹⁷. These demographic changes highlight the necessity for thorough genetic monitoring and adaptive management techniques to preserve these genetic diversity and population resilience. To worsen the situation, there is little information on genetic diversity and population structure of these indigenous goats, crossbreeding has gone hand in hand with uncontrolled mating. These practices dilute the locally adapted alleles in indigenous populations and over time, such uncontrolled mating can erode unique adaptive features like disease resistance, heat tolerance and compromise the resilience of these populations to climate change and disease outbreaks all these factors limit improvements in genetics, hence further complicating conservation efforts¹³. Most previous studies have focused on mitochondrial DNA and Y-chromosomal markers^{15,16}, thus leaving a significant knowledge gap in genome-wide variation patterns. While earlier research successfully utilized microsatellite markers to characterize genetic diversity and population structure in various small ruminant breeds in Eastern Europe^{18,19}, recent technological advances have enabled more comprehensive genomic analyses. The use of high-density SNP arrays, such as the 70 K SNP platform, can now provide deeper insights into population structure and genetic diversity patterns which are essential for both conservation efforts and sustainable breeding programs²⁰. But it is important to recognize that a lot of population genetic approaches, including those used with SNP data, are predicated on ideas related to the molecular clock hypothesis and the neutral theory of molecular evolution. As more and more empirical data challenges the idea of mutation rate constancy²¹ and shows the ubiquitous role of selection for instance functional mtDNA variation²², transcription factor binding by STRs²³, and non-neutral diversity patterns²⁴), these frameworks have come under increasing criticism. New hypotheses that highlight the shortcomings of conventional neutral models include the greatest genetic diversity hypothesis²⁵. Therefore, building upon these previous methodological approaches, we used the genome-wide caprine 70 K SNP markers to assess the genetic structure of selected local goat population in Kazakhstan to determine their genetic diversity status using within-population diversity indices including observed and expected heterozygosity (H_o and H_e), inbreeding coefficients. In addition, the relationship between selected populations was assessed using principal component analysis and phylogenetic trees. Further we identified runs of homozygosity patterns and signatures of selection for local adaptation. While our high-density SNP arrays genomic data fill previous gaps in genome-wide analyses, interpretations of diversity and structure are still problematic, as they may reflect both neutral and non-neutral processes considering changing theoretical paradigms. Results from this study will provides an insight into the genetic architecture of Kazakhstan's local goat population for improved conservation and management strategies.

Results

Genetic diversity indices, ROH and inbreeding coefficients

On average, H_o and H_e across all populations were similar, with H_o slightly below the expected value. The mean H_o and H_e for all populations were 0.424 and 0.429, respectively. Population wise, the H_o ranged between 0.389 (Kosseit) and 0.444 (Shokpar), while H_e ranged between 0.441 (Ushterek) and 0.409 (Kosseit) (Supplementary file1, Figure S1, supplementary material online). All goat populations also registered different figures of inbreeding coefficients, Kundyzy and Shokpar registered negative F_{IS} indicating excess heterozygosity in both breeds. Inbreeding coefficients also varied across breeds, for instance, F_{hom} ranged between 0.045 and 0.126, F_{hat1} between -0.002 and 0.135 , F_{hat2} between 0.002 and 0.084 (Table 1).

The correlations between different inbreeding coefficients, F_{hat1} , F_{GRM} or F_{hat3} , F_{hom} and F_{ROH} are shown in Fig. 1. The highest correlation (0.99) was observed between F_{hom} and F_{hat2} while the lowest correlation of 0.07 between F_{hat1} and F_{ROH} . Additionally, the correlation between F_{hat1} and F_{hom} as well as F_{hat1} and F_{hat3} was relatively high 0.86 and 0.79, respectively. We also observed a relatively low correlation between (0.26) F_{hat2} and F_{ROH} . The high level of correlation between F_{hat2} and F_{hom} indicates that the two methods can provide detailed information on the various aspects of inbreeding as compared to other measures.

Analysis of ROH indicated that the count of short runs (0–6 bps) dominated all others with some population lacking long (>48) runs, indicating that, most of the inbreeding in these populations happened in recent generations. Kundyzy and Kenes exhibited highest level count of short runs while Kosseit and Darbaza are the only populations that had longest runs (Table 2).

Population	N	H _o	He	FIS	Fhom	Fhat1	Fhat2	FGRM
Darbaza	20	0.412 ± 0.16	0.429 ± 0.11	0.052 ± 0.29	0.075 ± 0.09	0.064 ± 0.07	0.076 ± 0.10	0.070 ± 0.08
Kenes	20	0.430 ± 0.15	0.434 ± 0.10	0.004 ± 0.24	0.030 ± 0.03	0.006 ± 0.02	0.034 ± 0.03	0.020 ± 0.02
Kosseit	20	0.389 ± 0.16	0.409 ± 0.12	0.052 ± 0.25	0.126 ± 0.11	0.054 ± 0.08	0.135 ± 0.12	0.094 ± 0.09
Kundyzdy	20	0.426 ± 0.15	0.423 ± 0.11	-0.011 ± 0.23	0.045 ± 0.05	-0.002 ± 0.05	0.053 ± 0.05	0.025 ± 0.05
Shokpar	20	0.444 ± 0.17	0.435 ± 0.11	-0.027 ± 0.28	0.006 ± 0.01	-0.007 ± 0.01	0.010 ± 0.02	0.002 ± 0.01
Ushterek	20	0.441 ± 0.15	0.441 ± 0.10	0.000 ± 0.25	0.007 ± 0.04	0.189 ± 0.07	-0.022 ± 0.05	0.084 ± 0.05

Table 1. The genetic diversity indices for six local goat populations from Kazakhstan. Sample size (N), expected heterozygosity (He), observed heterozygosity (Ho), inbreeding coefficient (FIS), inbreeding coefficient based on homozygosity (Fhom), observed homozygosity against whole population (Fhat1), excess homozygosity based on allele frequencies (Fhat2) and identity by descent (FGRM).

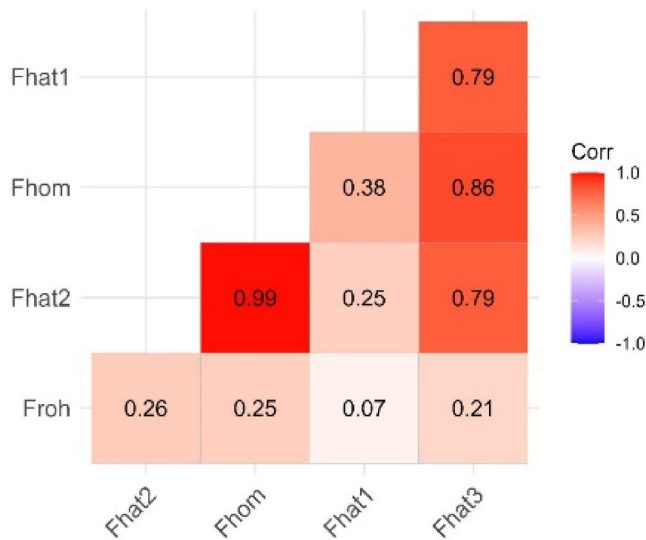


Fig. 1. Correlation between inbreeding coefficients for the six Kazakhstan goat populations.

Runs	Darbaza	Kenes	Kosseit	Kundyzdy	Shokpar	Ushterek
0–6	1331	2005	1937	2399	1239	1832
6–12	61	19	89	34	2	41
12–24	43	15	78	21	1	3
24–48	19	4	51	9	NA	NA
> 48	7	NA	15	NA	NA	NA

Table 2. Count of runs of homozygosity per breed and length of the run (bps).

Intra – and interpopulation relationships and population structure

To assess both intra- and inter-population relationships, several methods were employed. Intra-population relationships, for example, were examined using a neighbor-joining tree (Fig. 2a). This analysis revealed three main groups, each with several branches or subgroups. These results align with the population grouping shown in Fig. 2b, which divides the populations into five major groups. However, the populations of Ushterek and Kosseit show greater differentiation from the others, while Kenes appears to be less differentiated.

Further, interpopulation relationships were evaluated using PCA (Fig. 3). It was observed that similar to the neighbor-joining trees, three major groups emerged with Ushterek and Kosseit exhibiting a tendency of being genomically distant from other groups, although they appear to be genetically correlated based on PC1. All other populations were clustered together indicating close genetic relationship. Both PC1 and PC2 accounted for more than 50% of the genetic variation between populations indicating the existence of population structure.

Co-ancestry analysis was performed using ADMIXTURE software, evaluating various numbers of genetic clusters (K) from K=2 to K=8, (see Supplementary File 1, Figure S2, Supplementary material online). The population structure corresponding to the optimal K value is presented in Fig. 4a, while Fig. 4b displays

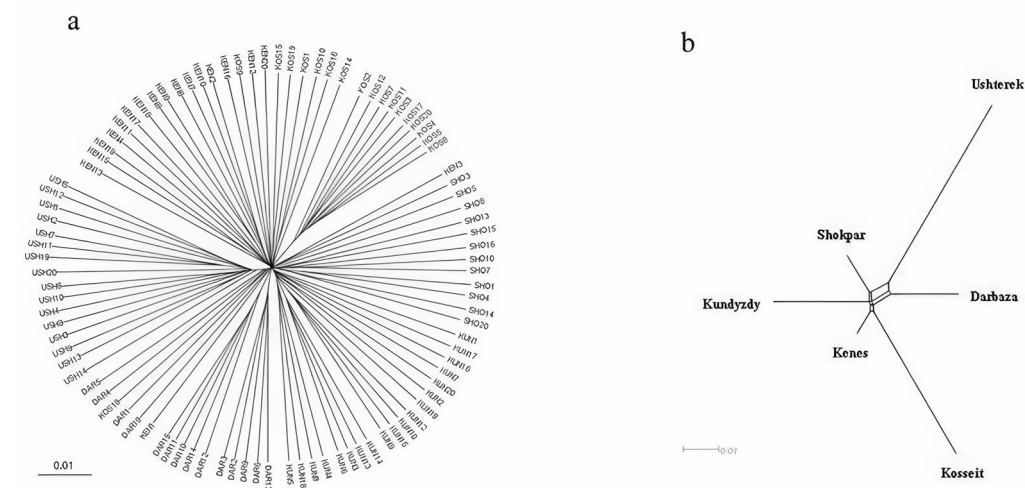


Fig. 2. Neighbor joining tree for the studied populations, (a) unrooted neighbor joining tree and (b) neighbor network based on distance matrix calculated from pairwise F_{st} values.

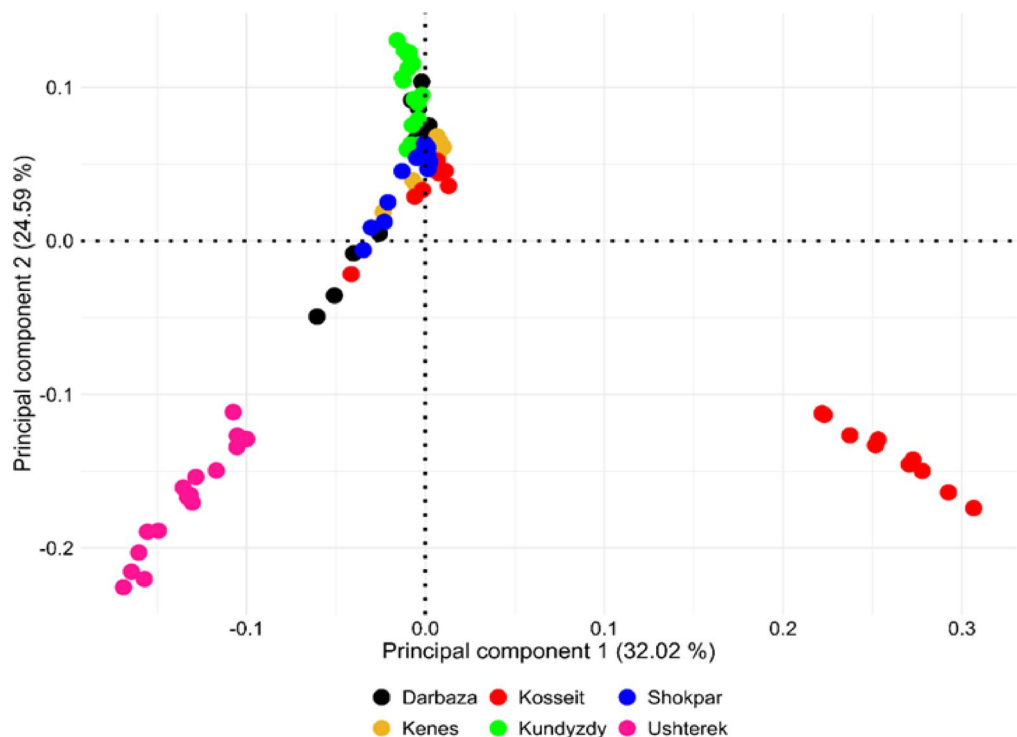


Fig. 3. Population structure of six goat breeds from Kazakhstan as defined by principal component analysis (PCA), PC1 versus PC2.

optimal K with the lowest cross-validation error. At this K value, the analysis revealed high levels of admixture, with Darbaza, Kenes, and Kundyzy showing similar patterns of ancestry, suggesting they represent a single population. In contrast, Kosseit and Ushterek were identified as distinct populations (Fig. 5a, b).

Identification of signatures of selection

Complementary methods were employed to identify selection signatures potentially underlying adaptation to local conditions. ROH identified 98 while iHS identified more than 2000 genes. However, 71 genes were common to both methods, and these were selected for further investigation; (See supplementary File 2, Tables S1 to S3, Supplementary material online). for more details. The distribution of SNPs in ROH and iHS was visualized as shown (Fig. 6a and b).

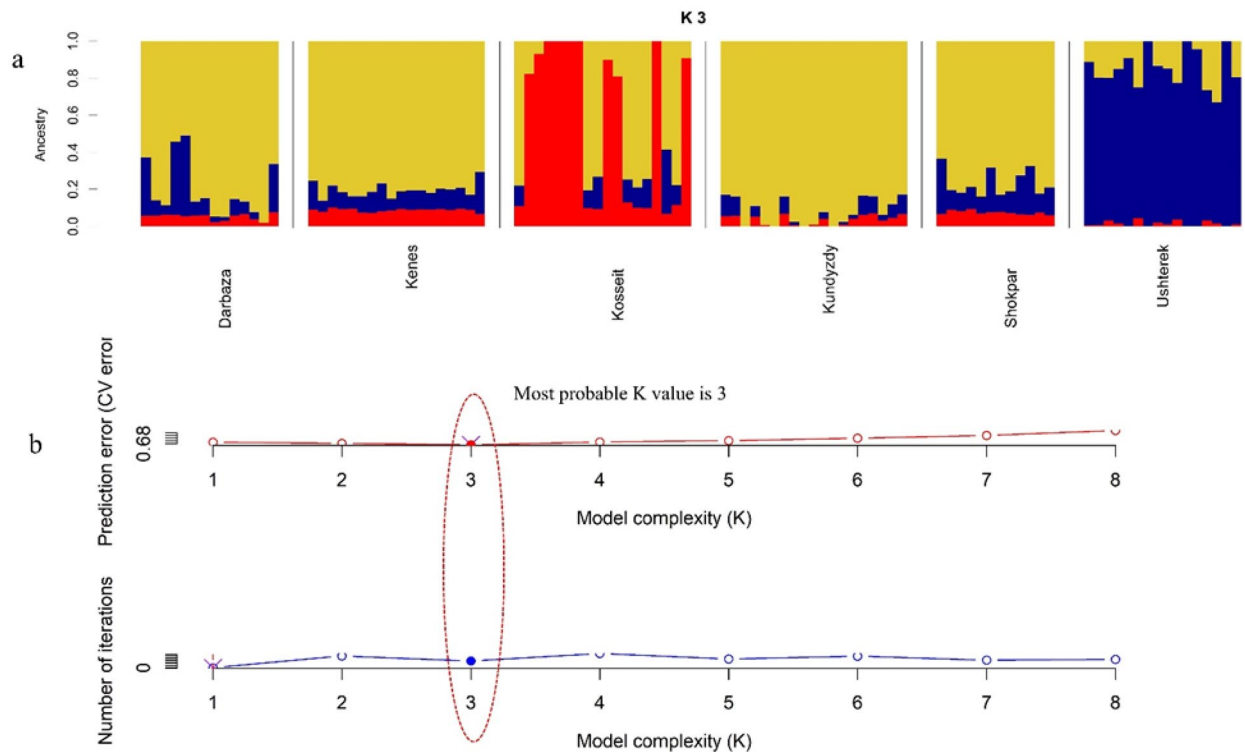


Fig. 4. Map of the blood collection points for the local goat population in Kazakhstan generated using QGIS⁸² (a) location of the country and (b) localization of the sampling area.

Gene ontology and enrichment terms and gene interaction network was performed on common genes to have a deeper understanding into the biological functions of identified genes. The most enriched biological function was localization followed by cellular process, developmental process and metabolism which are important functions for species adaptation. On the other hand, gene interaction network revealed that more than 67% of the interaction is co-expression indicating a strong functional relationship between these genes (See supplementary File 1, Figures S3a and 3b, Supplementary material online).

Discussion

From a total of 120 samples and over 70,000 SNPs, quality control measures yielded a dataset of 97 samples and 49,838 SNPs, achieving an average call rate of 0.97. With a genotyping success rate of 97%, we are confident that our dataset is of high quality, making the conclusions drawn from these results reliable. Genetic diversity studies provide an important insight into the genetic architecture of the population²⁶. In our study the mean heterozygosity values ($H_o=0.424$; $H_e=0.429$) of Kazakhstan goat populations showed a moderate genetic diversity. Although these measurements of genetic diversity initially seem to be in line with neutral theory, attributes variation to random mutation and genetic drift, new empirical findings cast doubt on this fundamental paradigm. The Neutral theory, for instance, assumes that non-selective processes account for most of the genetic variability. Even in genomic regions that are typically categorized as neutral, there is growing evidence that selection, not drift, significantly shapes genetic diversity in species²². For example, genome-wide examinations of humans and water fleas (*Daphnia*) demonstrate that slow changes in gene versions over generations are caused by natural selection rather than chance. This calls into question the long-held belief that selective pressures do not influence the evolution of DNA sections that do not directly code for proteins²². The genetic variation from how findings therefore could be due to the following potential factors: either exposure to intensive selection pressures for economically important traits, or the inheritance of diverse genetic material from their ancestral populations. However, considering that local breeds are rarely subjected to systematic and objective selection programs in traditional management systems, we posit that the observed heterozygosity patterns are more likely to reflect diverse ancestral origins rather than artificial selection effects. This interpretation has historically assumed a neutral baseline, but growing evidence suggests that even SNPs and STRs may overlap with regions under selection^{22,23}. For instance, mitochondrial DNA (mtDNA), which was formerly thought to be a neutral marker but is now understood to be functionally significant²². The mitochondrial local constraint (MLC) metric further quantifies functional conservation in mtDNA, thereby underscoring how selection shapes even non-recombining genomes²². The values obtained in this investigation are within those observed in a study on global goat populations. For instance, the study on the worldwide goat population where indigenous goats from Italy recorded ($H_o=0.47$; $H_e=0.397$), Mongolian indigenous goats ($H_e=0.4$; $H_o=0.36$), Ethiopian Begait goats ($H_e=0.35$; $H_o=0.32$), Ugandan indigenous ($H_e=0.387$; $H_o=0.378$), and South African indigenous goats ($H_e=0.310$; $H_o=0.330$)^{27–30}. It is opined that these heterozygosity ranges fall within the sustainable genetic

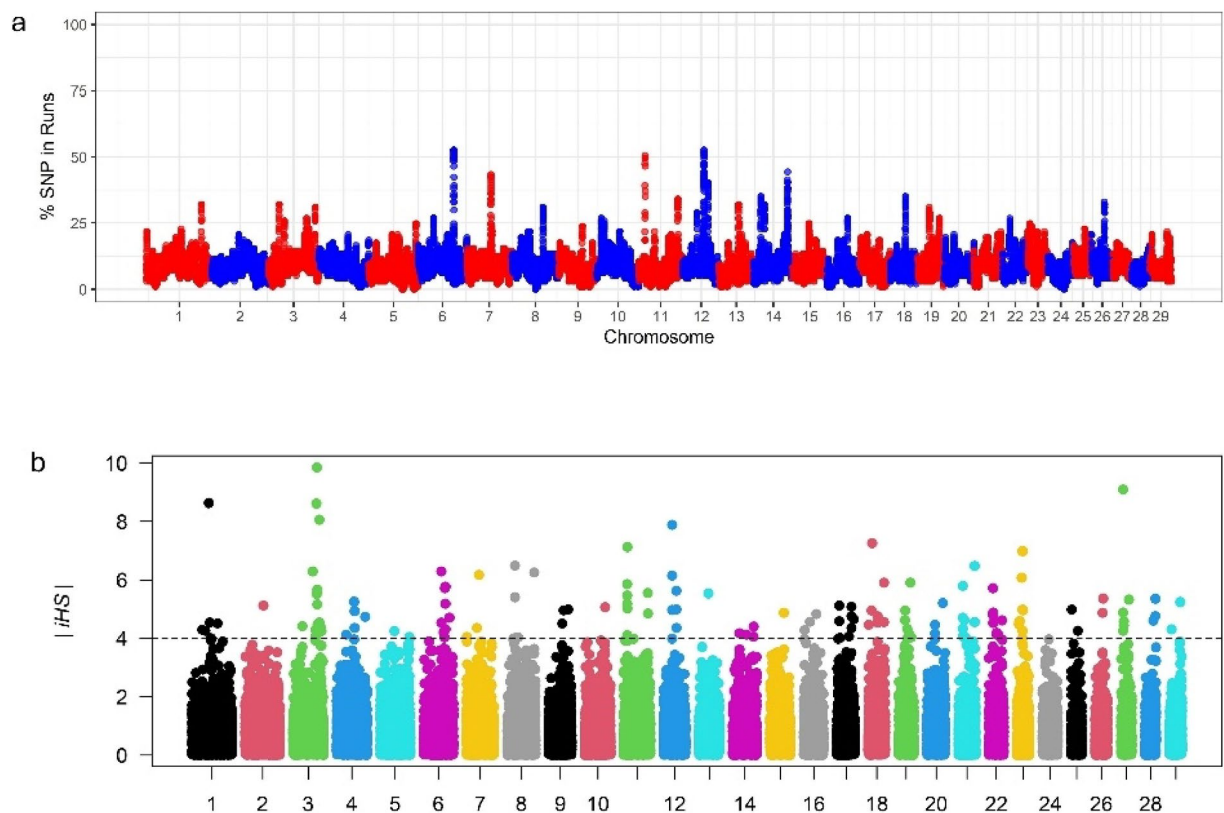


Fig. 5. Admixture analysis showing the proportions of ancestral populations for $K = 3$; (a) Bar chart showing the proportion of admixture with each vertical bar representing an individual. (b) Cross validation error plot showing the lowest value of CV which corresponds to $K = 3$.

breed management, and thus we conclude that the studied populations have a considerable within-population genetic diversity.

The current F_{ST} values of the population (0.014–0.052) show moderate genetic differentiation. This is similar to the findings of the VarGoats project, which reported generally low F_{ST} values (0.017) in Asian goat populations³¹, 0.024 for Bulgarian native goats³², 0.016 in Ugandan Indigenous Goats³³. These values suggest that even though there is some genetic differentiation, there still exists some levels of gene flow among these populations that may have a raised due to pastoralist mobility nature. However, the assumption that low F_{ST} reflects neutrality is increasingly questioned. Adaptive introgression, or balanced selection, can preserve functional diversity across populations while reducing differentiation, as recent research shows that gene flow and selection are not mutually exclusive²¹. The analysis of F_{IS} also revealed a concerning pattern in the Kosseit population (−0.011 to 0.052). While some subpopulations showed a healthy genetic mixing through negative F_{IS} values, the Kosseit group showed a clear sign of isolation with reduced genetic variability ($H_o = 0.389 \pm 0.16$), high levels of inbreeding based on homozygosity values ($H_{om} = 0.126 \pm 0.11$) and positive inbreeding coefficient ($F_{IS} = 0.052$). Although drift and isolation are reasonable explanations, it is impossible to deny the influence of selection on inbreeding patterns. For instance, inbreeding may purge harmful recessive alleles, a process that cannot be detected under rigorous neutral models²¹. From our study, the isolation appears to result from multiple factors including geographical barriers which tends to limit the gene flow into the population, reduction in effective population size due to selective breeding practices, and the use of related breeding bucks within the community. Similar patterns of inbreeding have been documented in other goat populations for instance, Ugandan indigenous goats ($F_{IS} = 0.014$), Chinese indigenous goat populations ($F_{IS} = 0.03$) and Sudanese Nubian goat breed ($F_{IS} = 0.073$)^{33–35}, study on four cattle breeds using 15 STR loci also reported an ($F_{IS} = 0.013$)³⁸, where small population size, traditional management practices, cultural preference to specific trait or colours, intensive selection and geographical barriers were identified as reasons for the levels of inbreeding observed. For our study, the impact is particularly concerning given the challenging environmental terrains where these goats have to thrive for survival.

ROH islands identified in certain genomic regions indicate that some specific alleles have been fixed in the populations, presumably due to selection or genetic drift²⁶, thus suggesting adaptation to local conditions. The length and distribution of ROHs offers an important insight into population structure and selection processes. Long ROHs indicate the presence of a more recent inbreeding, while shorter ROHs suggest the presence of ancient inbreeding within a given population^{36,37}. The mean sum of ROH segment coverage was generally higher for short ROHs than for long ROHs. However, Kundzydy populations showed a higher average sum of short run

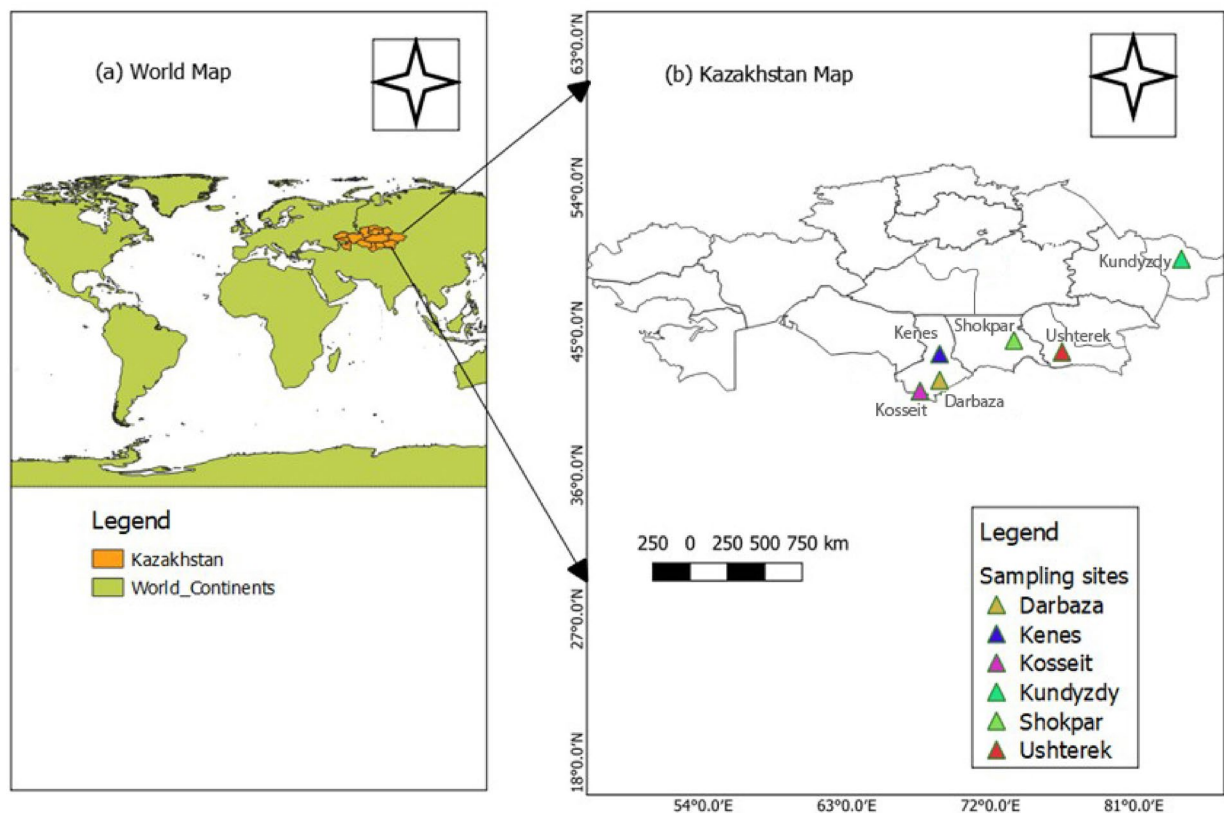


Fig. 6. Manhattan plots showing (a) percentage of SNPs in runs per chromosome and (b) distribution of iHS per chromosome.

at $ROH > 6\text{mb}$ ^{38,39}. The distribution of ROH coverage reported in this study agrees with findings from other studies in goats. For instance³¹, on their study in various goats in Switzerland observed more of short run ROH segments ($< 1\text{mb}$) than those that were ($> 6\text{mb}$), Egyptian goats breed more of $< 2\text{mb}$ ⁴⁰, Italian goat breeds that were raised in the isolated northern Italy⁴¹, and Yunan goat breeds in China⁴², indication that the studied goat breed had more ancient inbreeding event probably due to a more distance common ancestor.

This highest count of short ROH in the Kazakhstan studied goat populations indicates that they could have been initially established by small founding populations but were not particularly highly affected by recent inbreeding. Selection also plays a role in the frequency of ROH in the genome. This finding aligns with the notion that certain populations can maintain genetic diversity despite geographic isolation and that ROHs do not only arise because of inbreeding, but also can provide information about demographic events and selective sweep⁴³.

The use of Neighbor-net joining (Fig. 2b) to assess interpopulation relationship revealed a complex pattern of genetic relationship among the studied goat populations. There was a close relationship between five populations except for Ushterek goats which had a distinct cluster, indicating a unique evolutionary history. This observation is also enhanced in the admixture analysis results (Fig. 5) where high levels of admixture between the studied populations were observed, with Darbaza, Kenes and Kundyzydy showing similar patterns of admixture. This suggests the presence of a single population which may have occurred due to historical crossbreeding amongst them. The country's traditional transhumance routes, which have historically made it easier for animals to travel between different regions, may have contributed to this degree of genetic admixture between populations⁴⁴. Additionally, the sale of breeding bucks, the cultural practices of different ethnic groups in the region⁴⁵, especially their traditional nomadic way of livestock breeding and husbandry techniques and the modern breeding initiatives that have brought about novel genetic exchange patterns, though these are often more structured and documented than historical movements. Together, these factors create a complex web of genetic relationship which has shaped the current population structure in the region.

The historical background of genetic connectivity between populations can be traced through multiple lines of evidence. The findings are similar to those observed in other regions, like the admixture recorded in Ugandan indigenous goats³³. Djallonke and Sahelian goat breeds in Benin⁴⁶, and in new world goat populations where Asian, African and European goats showed evidence of admixture⁴⁷. The VarGoats project has further illuminated these historical admixtures by demonstrating how Y-chromosomal haplogroups reflect ancient migration patterns and more recent breeding practices³¹. The Kosseit population showed a high level of single ancestry component, indicating significant geographic isolation, which is supported by high F_{hat2} and FGRM values. These findings are in line with other goat population studies such as Bohuai goats in China⁴⁸ and Mongolian

cashmere goats⁴⁹. Where their genetic structure was shaped by geographical isolation, breeding and husbandry practices. Various factors may have contributed to results observed in Kosseit goat population, this involves the Kazakhstan geographical terrain (mountains and deep valleys) for example the presence of the Tien Shan and Altai mountains may have created barrier that restricted the free movement of animals in the region, lack of access to breeding centers might have also caused a challenge in the introduction of new breeding material in the population⁵⁰. Besides, the Traditional livestock production systems employed which are passed down through generations might have encouraged inbreeding within the same populations. Additionally, selection for a trait of importance within the locality has further focused on a specific genetic characteristic⁵¹. The findings are in line with other goat studies such as Ethiopian Begait goat²⁸, and Bulgaria native goats³², which observed the mentioned factors as the influence on their single co-ancestry structure.

For this, the sustainability and adaptability of any population can only be maintained by understanding its genetic diversity^{52,53}. The different patterns between populations with lower genetic diversity (Kosseit and Darbaza) and those with higher genetic diversity (Shokpar and Ushterek) underscore the significance of balanced conservation strategies. Some populations have low genetic differentiation because of long-term selective pressures and management systems. This is brought about by the limited gene flow and traditional breeding methods used. Understanding these complex relationships and their historical context is essential to developing effective conservation strategies. As it is important for long-term survival and adaptability to ever changing climatic conditions. Therefore, it is important to prioritize those populations that showed some signs of genetic isolation or low genetic diversity by maintaining its genetic variability.

In developing countries such as Kazakhstan, productivity and reproduction in livestock population is generally affected by the environmental conditions, high temperatures, rocky and mountainous topography, disease incidences, low- and poor-quality feeds and pastures and inadequate information in livestock management⁵⁴. To withstand these challenges, species must develop some adaptive mechanisms that should enhance their immune response, growth, and overall livestock welfare^{55–57}. The existence of mountainous terrain and varied ecological zones in Kazakhstan may have helped in shaping the genetic diversity and enhancing the adaptability mechanism of their local goat populations¹⁶. Besides this, the physical barriers created by high-altitude ranges and deep valleys have led to the restriction of livestock movement thus affecting the genetic diversity of goat population in the region¹⁵, making geographical isolation become a possible key factor to the genetic differentiation observed among the studied goat populations. Additionally, traditional pastoral farming, including seasonal migration along interconnected routes, have further facilitated the gene flow and the exchange of breeding males, thus influencing the genetic landscape of the goat population⁴⁵.

To withstand these challenging conditions, these goats may have developed special mechanisms to boost their adaptation. In our study, we identified various candidate genes that were suspected to be underpinning adaptation to local environmental conditions. Here we only discuss genes that were detected by both iHS and ROH methods. The results from the gene ontology enrichment analysis revealed that the most over-represented biological term was localization (see Supplementary File 1, Figure S3a, Supplementary material online). Besides, gene-gene network analysis showed that 67% of these genes interact by co-expression (see Supplementary File 1, Figure S3b, Supplementary material online,). This suggests that these goat populations may have undergone evolutionary adaptations, leading to the development of proteins that localize to various cellular compartments, potentially evolving new functions that enhance their survival in harsh environmental conditions. In general, the most over-represented biological pathways identified in our study, including localization, cellular processes, and developmental processes, have also been reported in other research works. Notably, these pathways were observed in studies by⁵⁸ in Ethiopian goat populations, and¹⁸ in Swiss Alpine goat breeds. This consistency confirms the generality of these adaptation patterns across different ecological zones.

Some of the genes that we identified and are suspected to be enhancing these goats' adaptability include *BIRC6*, *IFT88*, *IL17D*, *IL17RE*, *IL17RC*, *NRN1L*, *NLR4*, *HCLS1*, *TSHR*, *FBXO40* and *GJB2* (See supplementary File 2, Tables S1 to S3, Supplementary material online), which have been known to play a key role in immune response, metabolic regulation, skeletal and muscle development and reproductive efficiency. For instance, *NLR4* and *HCLS1* genes play a key role in immunity function by acting as a sensor which recognizes some bacteria components and activates inflammasome pathways to produce pro-inflammatory cytokine which is an essential response for goat populations that are produced in regions that are susceptible to bacterial infections and diseases infestation⁵⁹. Additionally, *HCLS1* has been associated with resistance to *E.coli* infections in sheep⁵⁹ and thus highlighting its importance in livestock immunity. Further, we also identify *IL17D*, *IL17RE* and *IL17RC*, gene which is part of interleukin-17 group, this gene act by regulating the cytokine production and influence the recruitment of innate immune cells to infection areas⁶⁰. For instance In humans, the *IL17RE* gene have been seen to work in conjunction with *IL17RA* to drive IL-17 C-mediated mucosal and TH17 immune responses⁶¹. These interactions are important as they play a role of enhanced immunity⁶². Similar mechanisms have also been observed in identified in teleost fish like turbot⁶³. Besides, studies have also reported the importance of this gene in regulating immune system in including the Ethiopian indigenous goats populations⁶⁴, East African goats in Uganda³² and Garfagnina goats breeds in Italy⁶⁵ and have also been associated with gastrointestinal parasite resistance in sheep⁵⁹. Maintaining the health of the animals is one of the important economic aspects in livestock production, especially those produced in regions prone to disease pathogens⁵⁴. Hence the presence of these immune-related genes suggests that enhanced immunity is an important trait under selection in Kazakhstan goat populations.

BIRC6 and *TSHR* genes are said to be responsible for maintaining pregnancy health in bovine⁶⁶. Some studies have linked *BIRC6* to the increased litter sizes, hair follicle development and reproductive efficiency for example in the Mongolian cashmere and Markhoz goats^{67,68}. The role that this gene plays in the reproductive trait makes it a valuable genetic marker to be involved in the breeding program especially for the regions whose fertility is the key factor to livestock breeding. Besides, the *TSHR* gene identified in our study is responsible for signal

transduction, specifically in seasonal reproduction and cell survival which are vital in adapting to reproductive strategies in various environmental signals⁶⁹. The genes have been observed to influence seasonal reproduction in livestock, for instance, the presence of *TSHR* was seen to have influence on reproductive seasonality in Swiss goats⁷⁰. Additionally, this gene has been linked to metabolic regulation and photoperiodic control of reproduction, enabling all year-round egg production in chicken⁷¹. In sheep, *TSHR* has been associated with environmental adaptability and litter size^{72–74}.

The *IFT88* gene is essential for the formation of cilia and flagella, making it important for various cellular processes, including signaling and movement. As a result, it plays a significant role in development and cellular homeostasis⁷⁵. In goats, the *IFT88* gene has been associated with cashmere and fiber production, hence we propose that this gene may have undergone unconscious selection pressure. This productivity is a key economic trait valued by many livestock keepers in relevant regions, making it an important focus in breeding programs. Breeders aiming to enhance fiber and cashmere yields should prioritize this trait in their selection strategies^{67,76}. Another gene identified in our analysis was *FBXO40* gene which is vital in cell survival and proliferation acts by controlling the pathways related to cell growth and cell death mechanisms⁷⁷. In livestock, the *FBXO40* gene has been related to muscle formation and control⁷⁸. The *FBXO40* gene has been observed to enhance muscle development in pigs⁷⁹. For Kazakhstan goats, we propose that this gene is crucial for physical endurance thus important in moving across the steep valleys and mountainous regions of Kazakhstan. On the other hand, the genes *GJB2* and *GJB6*, which encode gap junction proteins associated with cell communication, were also identified. These genes influence body size, skeletal and embryonic development, as well as indirectly affecting growth traits⁸⁰.

Management implications and future conservation perspectives

Our results have important implications for conservation and breeding programs and indicate the need for further studies. The Kosseit population emerges as priority for conservation efforts owing to their low levels of genetic diversity and high inbreeding coefficients. However, maximum genetic diversity theory (MGD) cautions against simplistic interventions thus maintaining genetic diversity is not merely about avoiding inbreeding but also conserving the adaptive traits through selection driven variation²⁴. Therefore, Important interventions like adoption of a controlled breeding program and periodic introduction of unrelated breeding bucks should be considered to curb further genetic erosion in this population. On the other hand, the high genetic diversity exhibited in the Shokpar and Ushterek goat populations act as a major genetic reservoir that needs to be carefully monitored and maintained for their genetic variation. Genetic polymorphism studies in Eastern European breeds have revealed important insights into breed characterization and genetic diversity patterns⁸¹, highlighting the value of molecular markers for understanding breed relationships and developing conservation strategies. The selection signatures identified in this study will serve as important features in target breeding programs that will seek to improve adaptation while maintaining genetic diversity. However, there are several limitations in our current study that need to be addressed through future research. The relatively small sample size per population might be responsible for the inability to discern some rare variants and fine-scaled population structure. Also, the absence of phenotypic data at this time hinders any attempts to make such relations between the genetic variants detected and adaptive traits. A deeper understanding could be gained by incorporating some details as the correlations of environmental data and genetic patterns which can explain the local adaptation mechanisms in detail. To rule out any ambiguities between neutral and selective processes, future studies should consider incorporating temporal genomic data such as tracking allele frequencies across generations and environmental covariates. Whole genome sequencing can be done to identify functional variants overlapping within the neutral SNPs while including a detailed phenotypic characterization, and environmental data. Such an inclusive approach will allow for a better understanding of the potential adaptive traits of these goat populations and inform on better strategies on how conservation measures can be put in place. The absence of phenotypic data at this time hinders any attempts to correlate genetic variants with adaptive traits. Additionally, monitoring and evaluation measures should be put in place to help in assessing the changes in genetic diversity and population structures especially for those goat populations that have been identified as being at risk of extinction. This will facilitate utilization in decision making about the verification of breeding outcomes by contributing value to the generation and preservation of genetic and phenotypic information in a central database. Furthermore, successful implementation of conservation strategies requires a strong stakeholder engagement and coordination programs. Which includes establishing networks between breeding centers, research institutions, and local farmers organizations, developing training programs for genetic management, and creating a regional conservation action plan with clear timelines and responsibilities. Regular monitoring and evaluation of conservation outcomes will be essential to assess the effectiveness of these interventions and make necessary adjustments. These evidence-based conservation strategies, derived from our genomic findings, provide a framework for maintaining the genetic integrity of Kazakhstan's local goat populations while preserving their adaptive potential for future breeding programs.

Conclusion

Based on the high-density 70 K SNP data, this study therefore presents a comprehensive genomic characterization of the local goats of Kazakhstan, which not only unveils important patterns of genetic diversity, population differentiation but also potential adaptive traits. While this advanced methodology provides high-resolution insights, our interpretations remain constrained by reliance on neutrality-based metrics such as *F_{ST}* and *ROH*, that is rooted in the flawed assumptions of neutral theory and the molecular clock hypothesis. The results of the molecular assessment revealed a high variability in levels of genetic diversity among populations, with Kosseit and Darbaza exhibiting alarming levels of low genetic diversity and high-risk rates of inbreeding which increases susceptibility to disease and low reproductive performance, indicating an urgent need for conservation efforts



Fig. 7. Local goats sampled in Kazakhstan; (a) Darbaza, (b) Ushterek, (c) Kosseit, (d) Shokpar, (e) Kundyzdy, and (f) Kenes (Photos taken by authors).

targeting these populations. On the other hand, the genetic diversity of the Shokpar and Ushterek populations is higher, making them reservoirs of rare alleles for the breeding program aimed at enhancing resilience while preserving genetic variation. The identification of selection signatures and adaptive genes (NLRC4, TSHR, IFT88) linked to immunity, reproduction, and environmental stress underscores the pervasive role of selection in shaping genomic architecture of Kazakhstan's indigenous goats, although our finding is incompatible with the neutral theory of drift-dominated evolution, but it provides an overview of population structure of this population. Gene Ontology (GO) enrichment highlighted important pathways including localization and cellular processes, suggesting that the protein localization mechanism is critical for survival in harsh environmental conditions which aligns with growing evidence that functional constraints, even in non-coding regions, challenge neutrality-based frameworks. While the use of high-density 70 K SNPs in our study provides high-resolution insights, our interpretations remain constrained by reliance on neutrality-based measurements which are rooted in the flawed assumptions of neutral theory and the molecular clock hypothesis, which recent studies has mostly rejected due to their predictive nature. Therefore, the use of Maximum Genetic Diversity (MGD) theory, which posits that diversity is actively maintained to buffer environmental change, has the potential to provide a more robust lens for interpreting genomic patterns in various species. Basing on this argument, this study therefore contributes significantly to understanding the indigenous goats genetic resource in Central Asia, yet its broader implications focused on adopting selection-aware models. Future work should therefore consider integrating whole-genome sequencing, phenotypic data, and Maximum Genetic Diversity (MGD) theory to disentangle adaptive processes from outdated neutral assumptions. By bridging these methodological advancements with theoretical innovation, we can develop strategies that will conserve both genetic diversity and adaptive potential of this indigenous populations in various agroecological zones worldwide.

Materials and methods

Sample collection and genotyping

A total of 120 blood samples were collected from different villages which located in South (Darbaza, Shokpar, Ushterek, Kenes and Kosseit) and East (Kundyzdy) parts of Kazakhstan. To study 20 individuals were randomly selected from each village (Fig. 4).

The studied goat populations are not formally characterized as distinct breeds but rather represent ecotypes of local aboriginal breeds, here named after the villages where they are raised. These goats exhibit several phenotypic manifestations such as ear set, horn shape, and coat colors, reflecting the lack of systematic breed classification in Kazakhstan and demonstrates phenotypic diversity (Fig. 7).

Samples were collected by experienced veterinarians, following all requirements. Blood samples were collected from the jugular vein of the goats by using vacutainer tubes⁸³. The samples were then stored in the laboratory freezer at temperatures below -20°C , awaiting to be processed to FTA cards⁸⁴ and sent to the NEOGEN Company in Scotland for genotyping using CaprineSNP70K Bead Chip array (Illumina, USA). The 70 K bead chip contains more than 70,000 evenly spaced SNP-targeting probes, which allows adequate SNP density for assessment of genetic diversity and signatures of selection. The whole pipeline from DNA isolation to variant call was performed by the genotyping company. Post-genotyping data management was performed interchangeably using PLINK v1.9 and v2⁸⁵. The data used in this analysis before quality control comprised 120 samples, each with 64,264 autosomes. Quality control was done using the following filtering thresholds:

- Exclusion of samples with more than 10% of missing data,

- SNPs with minor allele frequencies of less than 5%,
- SNPs with more than 10% missing genotypes,
- SNPs which violated Hardy-Weinberg equilibrium (HWE) using “--hwe 1e-6 keep-fewhet” flag.
- We excluded first-line relations (parent- offspring and sibling-sibling) using --king-cutoff 0.177.
- All SNPs at linkage disequilibrium were pruned using the “--indep-pairwise 200 kb 1 0.5” flag and options.

During the quality control process, the following exclusions were made:

- One sample was removed due to missing genotypes,
- 2,059 variants were excluded because of missing genotype data,
- 8,588 variants were discarded for not meeting minor allele frequency thresholds.
- 22 samples were removed due to close relationships,
- 3,711 variants were eliminated based on LD pruning thresholds,
- 58 variants were excluded due to HWE threshold violations.

Quality control measures resulted in a dataset comprising 97 samples and 49,838 SNPs, with an average call rate of 0.97 for further analysis.

Ethics declarations

Sampling in this study were approved by Local Ethics Commission Institute of Genetics and Physiology CS MSHE RK, approval number #3, 19.09.2024. All methods were carried out in accordance with relevant guidelines and regulations established by the institutional committee and in compliance with national legislation on animal experimentation. Furthermore, all methods are reported in accordance with ARRIVE guidelines to ensure comprehensive and transparent reporting of animal research. The welfare of the animals was monitored with appropriate measures implemented to minimize discomfort and distress.

Bioinformatic analysis

Within and between population genetic diversity

Basic diversity indices, including observed and expected homozygosity (H_e and H_o) were calculated using Arlequin software⁸⁶, a standard tool for estimating genetic diversity indices. while the inbreeding coefficient (method-of-moments F coefficient estimates calculated as (observed homozygote count - expected count) / (total observations - expected count) was calculated by Plink v 1.9, this allows for a straightforward estimation of excess homozygosity in a given population. It is important to mention here that this F is later referred to as F_{hom} . To identify the genetic structure and relationships between populations, several methods were applied. Firstly, multidimensional scaling was performed by principal component analysis (PCA) using Plink v2, PCA was chosen as it is appropriate for reducing dimensionality and visualizing genetic variation, under the assumption of linear independence between principal components. Visualization of the PCA was done using tidy verse package⁸⁷ in R software⁸⁸. Secondly, pairwise fixation index was performed using StAMPP package⁸⁹ implemented in R⁸⁸, this allows for biallelic SNP data and allows for estimation of population differentiation under the assumption of random mating within populations, and it was performed to determine the level of differentiation between ecotypes. Neighbor joining network was constructed using Split Tree program⁹⁰, These methods rely on genetic distance matrices derived from SNP data and do not assume Hardy-Weinberg equilibrium, making them suitable for diverse population structures. Further, the relationship between individuals within the population was calculated and visualized using Tassel software⁹¹. Admixture proportion or ancestry analysis of each sampled individual was performed using algorithm implemented in ADMIXTURE 1.3 software⁹², This method assumes Hardy-Weinberg equilibrium and linkage equilibrium within clusters. The algorithm was performed in unsupervised clustering with number of clusters (K) set between 1 and 8. The optimal K value was identified as the K value with the lowest cross-validation error. Visualization of Admixture results was performed using BITE v2⁹³ package implemented in R software⁸⁸.

Detection of runs of homozygosity (ROH) and inbreeding

Data used to detect ROHs was not trimmed for linkage disequilibrium (LD) since LD is part of the genomic variability that defines phenotypic variations between individuals and populations. ROH was done in R package – detectRUNS⁹⁴ using sliding window method by applying the following thresholds: window size of sliding window of 15, the threshold of overlapping windows of the same state (homozygous/heterozygous) to call a SNP in a RUN-0.05, minimum number of SNPs in the run-3, minimum number of homozygous/heterozygous SNP in the sliding window-1, minimum length of run in base pairs (bps) -1000 and minimum density of SNPs per 1000 bps = 1000. The detected runs were grouped into five depending on the length; 1–6 bps, 6–12, 12–24, 24–48 and > 48bps.

Identification of signatures of selection

The identification of ROH islands was performed using the detectRUNS R package⁸⁸. ROH islands (top runs) were defined as region where SNPs are inside runs in more than 70% of the individuals in the population. And it is believed that genes identified within these ROH islands underpins adaptation to local environments. Further, a complimentary method to ROH islands - Integrated haplotype score (IHS) method was used to detect signatures of selection using Rehh package⁹⁵ implemented in R software. Candidate regions were calculated and identified using a sliding window method by applying thresholds similar to ones applied in the identification of ROH islands. BioMart package, developed under Bioconductor, and implemented in R software was used to hunt for protein coding genes at or near the genomic regions identified as signatures of selection. Gene annotation was

done based on ARS1.2 annotated genome⁹⁶ on both + and – strands. Only genes that were found in both ROH Islands iHS method were considered for further gene ontology terms and network analysis. Gene ontology terms were performed using Metaphase platform⁹⁷, while gene-gene network analysis was performed by GeneMania⁹⁸. Description of gene functions was done using the already published research works.

Data availability

Data supporting this manuscript is available upon request from the corresponding author.

Received: 26 February 2025; Accepted: 13 May 2025

Published online: 02 June 2025

References

1. Zeder, M. A. & Hesse, B. The initial domestication of goats (*Capra hircus*) in the Zagros mountains 10,000 years ago. *Science* **287**, 2254–2257. <https://doi.org/10.1126/science.287.5461.2254> (2000).
2. Zeder, M. A. Animal domestication in the Zagros: an update and directions for future research. *MOM Éditions*. **49** (1), 243–277 (2008).
3. Naderi, S. et al. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences*, **105** (46), 17659–17664. (2008). <https://doi.org/10.1073/pnas.0804782105>.
4. Daly, K. G. et al. Ancient goat genomes reveal mosaic domestication in the fertile crescent. *Science* **361** (6397), 85–88. <https://doi.org/10.1126/science.aas9411>. (2018).
5. Manunza, A. et al. Discovering novel clues of natural selection on four worldwide goat breeds. *Sci. Rep.* **13** (1), 2110. <https://doi.org/10.1038/s41598-023-27490-x>. (2023).
6. Panigrahi, M. et al. Landmarks in the history of selective sweeps. *Anim. Genet.* **54** (6), 667–688. <https://doi.org/10.1111/age.13355>. (2023).
7. Panigrahi, M. et al. Trajectory of livestock genomics in South Asia: a comprehensive review. *Gene* **843** <https://doi.org/10.1016/j.gene.2022.146808> (2022).
8. Nayak, S. S. et al. A comprehensive review of livestock development: insights into domestication, phylogenetics, diversity, and genomic advances. *Mamm. Genome*. **35**, 577–599. <https://doi.org/10.1007/s00335-024-10075-y> (2024).
9. Magoro, A. M., Mtileni, B., Hadebe, K. & Zwane, A. Assessment of genetic diversity and conservation in South African Indigenous goat ecotypes: A review. *Animals* **12** (23), 3353. <https://doi.org/10.3390/ani12233353>. (2022).
10. Omondi, I. A., Baltenweck, I., Drucker, A. G., Obare, G. A. & Zander, K. K. Valuing goat genetic resources: a pro-poor growth strategy in the Kenyan semi-arid tropics. *Trop. Anim. Health Prod.* **40**, 583–596. <https://doi.org/10.1007/s11250-008-9137-2> (2008).
11. Nair, M. R. R. et al. Bhatta R. Goat as the ideal climate-resilient animal model in tropical environment: revisiting advantages over other livestock species. *Int. J. Biometeorol.* **65**, 2229–2240. <https://doi.org/10.1007/s00484-021-02179-w> (2021).
12. Joy, A. et al. Resilience of small ruminants to climate change and increased environmental temperature: A review. *Animals* **10**, 867. <https://doi.org/10.3390/ani10050867> (2020).
13. Iñiguez, L. et al. D. Characterization of Mohair and cashmere in regions of Kazakhstan, Kyrgyzstan and Uzbekistan. *Small Ruminant Res.* **120** (2–3), 209–218. <https://doi.org/10.1016/j.smallrumres.2014.05.004> (2014).
14. Food and Agriculture Organization of the United Nations (FAO). *World Livestock: Livestock in food security*. Rome, Italy. (2011). Available at: <https://www.fao.org/4/i2373e/i2373e.pdf>.
15. Tabata, R. et al. The Eurasian steppe is an important goat propagation route: A phylogeographic analysis using mitochondrial DNA and Y-chromosome sequences of Kazakhstani goats. *Anim. Sci. J. = Nihon Chikusan Gakkaiho*. **90** (3), 317–322. <https://doi.org/10.1111/asj.13144> (2019).
16. Hermes, T. R. et al. High mitochondrial diversity of domesticated goats persisted among bronze and Iron age pastoralists in the inner Asian mountain corridor. *PLoS One*. **15** (5). <https://doi.org/10.1371/journal.pone.0233333>. (2020).
17. Kassenbayev, G., Kerimova, U., Rakhimzhanova, G. & Shalgimbayeva, K. Animal husbandry market in Kazakhstan: dynamics and prognosis. (2024). <https://doi.org/10.48077/scihor4.2024.176>.
18. Kusza Sz et al. Study of genetic differences among Slovak Tsigai populations using microsatellite markers. *Czech J. Anim. Sci.* **54** (10), 468–474. <https://doi.org/10.17221/1670-CJAS> (2009).
19. Kusza Sz et al. The genetic variability of Hungarian Tsigai sheep. *Archiv Tierzucht* **53** (3), 309–317. (2010). <https://doi.org/10.5194/aab-53-309-2010>. (2010).
20. Yirgu, M. et al. Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm. *BMC Genomic Data*. **24** (1), 7. <https://doi.org/10.1186/s12863-023-01109-6> (2023).
21. Lynch, M., Wei, W., Ye, Z. & Pfreder, M. The genome-wide signature of short-term Temporal selection. *PNAS* **121** (28), e2307107121 (2024).
22. Lake, N. J. et al. Quantifying constraint in the human mitochondrial genome. *Nature* **635**, 390–397 (2024).
23. Horton, C. A. et al. Short tandem repeats bind transcription factors to tune eukaryotic gene expression. *Sci* **381** (6664), eadd1250. <https://doi.org/10.1126/science.add1250> (2023).
24. Hang, M. et al. Global genomic diversity and conservation priorities for domestic animals are associated with the economies of their regions of origin. *Sci. Rep.* **8**, 11677. <https://doi.org/10.1038/s41598-018-30061-0> (2018).
25. Huang, S. The overlap feature of the genetic equidistance result—a fundamental biological phenomenon overlooked for nearly half of a century. *Biol. Theory* **5**, 1–9. https://doi.org/10.1162/BIOT_a_00021 (2010).
26. Colli, L. et al. Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genet. Selection Evol.* **50**, 1–20. <https://doi.org/10.1186/s12711-018-0422-x>. (2018).
27. Mukhina, V., Svishcheva, G., Voronkova, V., Stolpovsky, Y. & Piskunov, A. Genetic diversity, population structure and phylogeny of Indigenous goats of Mongolia revealed by SNP genotyping. *Animals* **12** (3), 1–15. <https://doi.org/10.3390/ani12030221>. (2022).
28. Gebreselase, H. B., Nigussie, H., Wang, C., Luo, C. G. & Diversity Population structure and selection signature in begait goats revealed by Whole-Genome sequencing. *Animals* **14**, 307. <https://doi.org/10.3390/ani14020307>. (2024).
29. Nantongo, Z. et al. Genetic diversity, population structure and kinship relationships highlight the environmental influence on Uganda's Indigenous goat populations. *Front. Genet.* **15**, 1385611. <https://doi.org/10.3389/fgene.2024.1385611>. (2024).
30. Chokoe, T. C. et al. Genetic diversity of South African indigenous goat population from four provinces using genome-wide SNP data. *Sustainability* **12**, 10361. (2020). <https://doi.org/10.3390/su122410361>. (2020).
31. Denoyelle, L. et al. VarGoats project: a dataset of 1159 whole-genome sequences to dissect *Capra hircus* global diversity. *Genet. Selection Evol.* **53** (1), 86. <https://doi.org/10.1186/s12711-021-00659-6> (2021).
32. Yordanov, G. et al. Genetic differentiation and population structure of two Bulgarian local goat breeds using microsatellite markers. *Diversity* (14242818), 16(7). doi.10.3390/d16070388. (2024).

33. Onzima, R. B. et al. Genome-wide characterization of selection signatures and runs of homozygosity in Ugandan goat breeds. *Front. Genet.* **318**. <https://doi.org/10.3389/fgene.2018.00318>. (2018).
34. Li, M. H. et al. Genetic relationships among twelve Chinese Indigenous goat populations based on microsatellite analysis. *Genet. Selection Evol.* **34**, 1–16. <https://doi.org/10.1186/1297-9686-34-6-729>. (2002).
35. Osman, M., Nadia, J. K., Ghada, H. A. E. & Rahman, A. H. A. Susceptibility of sudanese Nubian goats, Nilotic Dwarf goats and Garag Ewes to experimental infection with a mechanically transmitted trypanosoma Vivax stock. *Pak J. Biol. Sci.* **11**, 472–475. <https://doi.org/10.3923/pjbs.2008.472.475> (2008).
36. Ciampolini, R. et al. Statistical analysis of individual assignment tests among four cattle breeds using fifteen STR loci. *J. Anim. Sci.* **Vol. 84** (1), 11–19. <https://doi.org/10.2527/2006.84111x> (2006).
37. Browning, S. R. & Browning, B. L. Identity by descent between distant relatives: detection and applications. *Annu. Rev. Genet.* **46** (1), 617–633. <https://doi.org/10.1146/annurev-genet-110711-155534> (2012).
38. Peripolli, E. et al. Runs of homozygosity: current knowledge and applications in livestock. *Anim. Genet.* **47** (6), 637–646. <https://doi.org/10.1111/age.12526>. (2016).
39. Brito, L. F. et al. Characterization of linkage disequilibrium, consistency of gametic phase and admixture in Australian and Canadian goats. *BMC Genet.* **16**, 1–15. <https://doi.org/10.1186/s12863-015-0220-1> (2015).
40. Sallam, A. M. et al. Genome-wide landscape of runs of homozygosity and differentiation across Egyptian goat breeds. *BMC Genom.* **24** (1), 573. <https://doi.org/10.1186/s12864-023-09679-6> (2023).
41. Cortellari, M. et al. Runs of homozygosity in the Italian goat breeds: impact of management practices in low-input systems. *Genet. Selection Evol.* **53**, 1–11. <https://doi.org/10.1186/s12711-021-00685-4> (2021).
42. Huang, C. et al. Runs of homozygosity detection and selection signature analysis for local goat breeds in Yunnan, China. *Genes* **15** (3), 313. <https://doi.org/10.3390/genes15030313>. (2024).
43. Gorssen, W., Meyermans, R., Janssens, S. & Buys, N. A publicly available repository of ROH Islands reveals signatures of selection in different livestock and pet species. *Genet. Selection Evol.* **53**, 1–10. <https://doi.org/10.1186/s12711-020-00599-7> (2021).
44. Deniskova, T. et al. Population structure and genetic diversity of sheep breeds in Kyrgyzstan. *Front. Genet.* **10** <https://doi.org/10.3389/fgene.2019.01311> (2019).
45. Marshall, F. B., Dobney, K., Denham, T. & Capriles, J. M. Evaluating the roles of directed breeding and gene flow in animal domestication. *Proc. Natl. Acad. Sci. U.S.A.* **111** (17), 6153–6158. <https://doi.org/10.1073/pnas.1312984110> (2014).
46. Whannou, H. R. V. et al. Genetic diversity assessment of the Indigenous goat population of Benin using microsatellite markers. *Front. Genet.* **14**, 1079048. <https://doi.org/10.3389/fgene.2023.1079048> (2023).
47. Paim, T. D. P. et al. New world goat populations are a genetically diverse reservoir for future use. *Sci. Rep.* **9** (1), 1476. <https://doi.org/10.1038/s41598-019-38812-3> (2019).
48. Yao, Z. et al. Genetic diversity and signatures of selection in BoHuai goat revealed by whole-genome sequencing. *BMC Genom.* **24** (1), 116. <https://doi.org/10.1186/s12864-023-09204-9> (2023).
49. Baldan, S. et al. Genetic characterization of cashmere goat (*Capra hircus*) populations in Mongolia. *Front. Genet.* **15**, 1421529. <https://doi.org/10.3389/fgene.2024.1421529> (2024).
50. Robinson, S., Bozayeva, Z., Mukhamedova, N., Djanibekov, N. & Petrick, M. Ranchers or pastoralists? Farm size, specialisation and production strategy amongst cattle farmers in south-eastern Kazakhstan. *Pastoralism* **11** (1), 31. <https://doi.org/10.1186/s13570-021-00217-1> (2021).
51. Kerven, C., Robinson, S. & Behnke, R. Pastoralism at scale on the Kazakh rangelands: from clans to workers to ranchers. *Front. Sustainable Food Syst.* **4**, 590401. <https://doi.org/10.3389/fsufs.2020.590401>. (2021).
52. Wanjala, G. et al. A review on the potential effects of environmental and economic factors on sheep genetic diversity: consequences of climate change. *Saudi J. Biol. Sci.* **30** (1), 103505. <https://doi.org/10.1016/j.sjbs.2022.103505>. (2023).
53. Villanueva, B. et al. The value of genomic relationship matrices to estimate levels of inbreeding. *Genet. Selection Evol.* **53** (1), 42. <https://doi.org/10.1186/s12711-021-00635-0> (2021).
54. Kichamu, N. et al. A review on Indigenous goats of East Africa: A case for conservation and management. *Biology* **13** (6), 419. <https://doi.org/10.3390/biology13060419>. (2024).
55. Astuti, P. K. et al. Climate change and dairy farming sustainability; a causal loop paradox and its mitigation scenario. *Heliyon* <https://doi.org/10.1016/j.heliyon.2024.e25200> (2024).
56. Sundaram, B. & Kanneganti, T. D. Advances in Understanding activation and function of the NLR4 inflammasome. *Int. J. Mol. Sci.* **22** (3), 1048. <https://doi.org/10.3390/ijms22031048>. (2021).
57. Stella, A. et al. AdaptMap: exploring goat diversity and adaptation. *Genet. Selection Evol.* **50**, 1–7. <https://doi.org/10.1186/s12711-018-0427-5> (2018).
58. Petretto, E. et al. Investigating the footprint of post-domestication dispersal on the diversity of modern European, African and Asian goats. *Genet. Sel. Evol.* **56**, 55. <https://doi.org/10.1186/s12711-024-00923-5> (2024).
59. Duan, Y. et al. A study of the resistance of Hu sheep lambs to *Escherichia coli* F17 based on whole genome sequencing. *Animals* **14** (1), 161. <https://doi.org/10.3390/ani14010161>. (2024).
60. Wang, X. et al. Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. *Scientific Reports*, **6**(1), p.38932. <https://doi.org/10.1038/srep38932>. (2016).
61. Nies, J. F. & Panzer, U. IL-17 C/IL-17RE: emergence of a unique axis in TH17 biology. *Front. Immunol.* **11**, 341. <https://doi.org/10.3389/fimmu.2020.00341>. (2020).
62. Huangfu, L. et al. The IL-17 family in diseases: from bench to bedside. *Sig Transduct. Target. Ther.* **8**, 402 <https://doi.org/10.1038/s41392-023-01620-3> (2023).
63. González-Fernández, C., García-Álvarez, M. A. & Cuesta, A. Identification and functional characterization of fish IL-17 receptors suggest important roles in the response to nodavirus infection. *Mar. Life Sci. Technol.* **6**, 252–265. <https://doi.org/10.1007/s42995-024-00225-1> (2024).
64. Tarekegn, G. M. et al. Ethiopian Indigenous goats offer insights into past and recent demographic dynamics and local adaptation in sub-Saharan African goats. *Evol. Appl.* **14** (7), 1716–1731. <https://doi.org/10.1111/eva.13118> (2021).
65. Dadousis, C. et al. Keep Garfagnina alive. An integrated study on patterns of homozygosity, genomic inbreeding, admixture and breed traceability of the Italian Garfagnina goat breed. *PloS One*. **16** (1), e0232436. <https://doi.org/10.1371/journal.pone.0232436>. (2021).
66. Salilew-Wondim, D. et al. Depletion of BIRC6 leads to retarded bovine early embryonic development and blastocyst formation in vitro. *Reprod. Fertility Dev.* **22** (3), 564–579. <https://doi.org/10.1071/RD09112>. (2010).
67. Zhao, Q. et al. Genomic inbreeding and runs of homozygosity analysis of cashmere goat. *Animals: Open. Access. J. MDPI*. **14** (8), 1246. <https://doi.org/10.3390/ani14081246>. (2024).
68. Mahmoudi, P. et al. Genome-wide association study reveals novel candidate genes for litter size in Markhoz goats. *Front. Veterinary Sci.* **23**, 9:1045589. <https://doi.org/10.3389/fvets.2022.1045589>. (2022).
69. Karlsson, A. C. et al. A domestication-related mutation in the thyroid-stimulating hormone receptor gene (TSHR) modulates photoperiodic response and reproduction in chickens. *J. Mol. Endocrinol.* **56** (2), 1–14. <https://doi.org/10.1016/j.ygcn.2016.02.010> (2016).
70. Signer-Hasler, H. et al. Runs of homozygosity in Swiss goats reveal genetic changes associated with domestication and modern selection. *Genet. Selection Evol.* **54** (1), 6. <https://doi.org/10.1186/s12711-022-00695-w>. (2022).

71. Rubin, C. J. et al. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–591. <https://doi.org/10.1038/nature08832> (2010).
72. Liu, Z. et al. Genome-wide analysis reveals signatures of selection for important traits in domestic sheep from different ecoregions. *BMC Genom.* **17**, 863. <https://doi.org/10.1186/s12864-016-3212-2> (2016).
73. Tao, L. et al. Combined approaches identify known and novel genes associated with sheep litter size and non-seasonal breeding. *Anim. Genet.* **52**, 857–867. <https://doi.org/10.1111/age.13138> (2021).
74. Li, M. et al. Screen of small fragment mutations within the sheep thyroid stimulating hormone receptor gene associated with litter size. *Anim. Biotechnol.* **34**, 658–663. <https://doi.org/10.1080/10495398.2021.1992415> (2023).
75. Pazour, G. J. et al. Chlamydomonas IFT 88 and its mouse homologue, polycystic kidney disease gene Tg 737, are required for assembly of cilia and flagella. *J. Cell Biol.* **151** (3), 709–718. <https://doi.org/10.1083/jcb.151.3.709> (2000).
76. Liu, Z. et al. Comprehensive whole-genome resequencing unveils genetic diversity and selective signatures of the Xiangdong black goat. *Front. Genet.* **15**, 1326828. <https://doi.org/10.3389/fgene.2024.1326828> (2024).
77. Feng, C., Yang, F. & Wang, J. FBXO4 inhibits lung cancer cell survival by targeting Mcl-1 for degradation. *Cancer Gene Ther.* **24**, 342–347. <https://doi.org/10.1038/cgt.2017.24> (2017).
78. Ye, J., Zhang, Y., Xu, J., Zhang, Q. & Zhu, D. FBXO40, a gene encoding a novel muscle-specific F-box protein, is upregulated in denervation-related muscle atrophy. *Gene* **404** (1–2), 53–60. <https://doi.org/10.1016/j.gene.2007.08.020> (2007).
79. Zou, Y. et al. An FBXO40 knockout generated by CRISPR/Cas9 causes muscle hypertrophy in pigs without detectable pathological effects. *Biochem. Biophys. Res. Commun.* **498** (4), 940–945. <https://doi.org/10.1016/j.bbrc.2018.03.085> (2018).
80. Kim, J. et al. The genome landscape of Indigenous African cattle. *Genome Biol.* **18**, 1–14. <https://doi.org/10.1186/s13059-017-1153-y> (2017).
81. Kusza Sz, Ilie, D. E., Sauer, M., Nagy, K. & Atanasiu, T. S. Gavojdian D. Study of *LGB* gene polymorphisms of small ruminants reared in Eastern Europe. *Czech J. Anim. Sci.*, 63, 152–159. <https://doi.org/10.17221/93/2017-CJAS> (2018).
82. QGIS development team. QGIS geographic information system. open-source geospatial foundation project. (2024).
83. Thomson, J. M., Easton, A. N. N. C. & Faragher, E. B. The use of vacutainer tubes for collection and storage of blood for coagulation testing. *Clin. Lab. Haematol.* **5** (4), 413–421. <https://doi.org/10.1111/j.1365-2257.1983.tb00514.x> (1983).
84. Dash, H. R., Shrivastava, P. & Das, S. Reliable use of Whatman™ FTATM cards for one-step collection and isolation of DNA. In *Principles and practices of DNA analysis: a laboratory manual for forensic DNA typing* (109–115). Springer. (2020). https://doi.org/10.1007/978-1-0716-0274-4_14 (2020).
85. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81** (3), 559–575. <https://doi.org/10.1086/519795> (2007).
86. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinf. Online.* **1**, 47–50 (2005). <http://cmpg.unibe.ch/software/arlequin3>.
87. Wickham, H. & Tidyverse Easily install and load the 'tidyverse'. *R Package Version.* **1** (1). <https://doi.org/10.21105/joss.01686> (2017).
88. Team, R. C. R: A Language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria: (2020). Available at: <https://www.R-project.org>
89. Pembleton, L. W., Cogan, N. O. I. & Forster, J. W. StAMPP: an R package for calculation of genetic differentiation and population structure from SNP data. *Mol. Ecol. Resour.* **13** (4), 946–952. <https://doi.org/10.1111/1755-0998.12129.9> (2013).
90. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23** (2), 254–267. <https://doi.org/10.1093/molbev/msj030> (2006).
91. Bradbury, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23** (19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308> (2007).
92. Alexander, D. H. & Novembre, J. Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. <https://doi.org/10.1101/gr.094052.109> (2009).
93. Milanese, M. et al. *BITE*, an R package for biodiversity analyses. *bioRxiv* 181610; [online]. (2017).
94. Biscarini, F., Cozzi, P., Gaspa, G. & Marras, G. Detect runs of homozygosity and runs of heterozygosity in diploid genomes. Available online: <https://CRAN.R-project.org/package=detectRUNS> (2019) (accessed on 2 December 2024).
95. Gautier, M. & Vitalis, R. *rehh*: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, 28(8), 1176–1177. (2012). <https://doi.org/10.1093/bioinformatics/bts115>
96. Huang, Y., Chen, L. & Zhang, H. A near complete genome for goat genetic and genomic research. *BMC Genom.* **22**, 703. <https://doi.org/10.1186/s12711-021-00668-5> (2021).
97. The gene ontology consortium. gene ontology knowledgebase in 2023. *Genetics* **224** (1), iyad031. <https://doi.org/10.1093/genetics/iyad031> (2023).
98. Warde-Farley, D. et al. The genemania prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38** (suppl_2), W214–W220. <https://doi.org/10.1093/nar/gkq537> (2010).

Acknowledgements

The authors are grateful to private farmers who agreed to provide materials for the study and to Meirzhan Begmanov, Tilek Kapasuly, Gulfairuz Shaltenbay and Temirlan Kulboldin for their assistance. N.K. was supported by the Tempus Public Foundation within the Stipendium Hungaricum Programme for the studies. Supported by the University of Debrecen Scientific Research Bridging Fund (DETKA) and by the University of Debrecen Program for Scientific Publication.

Author contributions

Conceptualization, B.B. and Sz.K.; writing—original draft preparation, N.K.; sample collection, B.B. and K.D.; statistical analysis, N.K.; G.W.; Sz.K.; writing—reviewing and editing, G.W.; B.B.; K.D.; Z.B.; Sz.K.; supervision, Sz.K. All authors have read and agreed to the published version of the manuscript.

Funding

Open access funding provided by University of Debrecen.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-02427-8>.

Correspondence and requests for materials should be addressed to S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025