



## OPEN ACCESS

## EDITED BY

Sunday O. Peters,  
Berry College, United States

## REVIEWED BY

Herman Revelo,  
Fundación Universitaria San Martín, Colombia  
Berkant İsmail Yıldız,  
Akdeniz University, Türkiye

## \*CORRESPONDENCE

Mohammed Fahad Albeshr,  
✉ albeshr@ksu.edu.sa,  
Abdulrahman K. Aldawish,  
✉ Aldawish.a@outlook.com

†These authors have contributed equally to  
this work

RECEIVED 12 June 2025

ACCEPTED 29 July 2025

PUBLISHED 29 September 2025

## CITATION

Aldawish AK, Ibrahim MA, Alsubaie FM, Saquib Q  
and Albeshr MF (2025) Comparative genome-  
wide analysis of *Ovis aries* in Saudi Arabia  
highlighting inbreeding and genetic isolation of  
the Najdi sheep breed.  
*Front. Genet.* 16:1646127.  
doi: 10.3389/fgene.2025.1646127

## COPYRIGHT

© 2025 Aldawish, Ibrahim, Alsubaie, Saquib and  
Albeshr. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Comparative genome-wide analysis of *Ovis aries* in Saudi Arabia highlighting inbreeding and genetic isolation of the Najdi sheep breed

Abdulrahman K. Aldawish<sup>1,2,3\*†</sup>, Mohanad A. Ibrahim<sup>3,4</sup>,  
Faisal M. Alsubaie<sup>1,2,3</sup>, Quaiser Saquib<sup>1</sup> and  
Mohammed Fahad Albeshr<sup>1\*†</sup>

<sup>1</sup>Zoology Department, College of Science, King Saud University, Riyadh, Saudi Arabia, <sup>2</sup>National Livestock and Fisheries Development Program, Riyadh, Saudi Arabia, <sup>3</sup>Ministry of Environment, Water, and Agriculture (MEWA), Riyadh, Saudi Arabia, <sup>4</sup>Data Science Program, King Abdullah International Medical Research Center, Riyadh, Saudi Arabia

**Introduction:** Sheep (*Ovis aries*) farming through traditional practices plays a vital role in the socio-economic development of the Kingdom of Saudi Arabia (KSA). Genetic relationships among KSA sheep breeds remain poorly characterized.

**Methods:** In this study, we performed a comparative genome-wide single-nucleotide polymorphism (SNP) analysis of three sheep breeds (Najdi, Naemi, and Harri) to evaluate their genetic diversity. Blood samples from 95 individuals were genotyped using OvineSNP50 BeadChip, generating 34,026 high-quality SNPs.

**Result and discussion:** Minor allele frequency (MAF) distribution revealed the highest genetic diversity in Naemi, followed by Harri and Najdi. The run of homozygosity (ROH)-based inbreeding coefficient ( $F_{ROH}$ ) identified Najdi as the most inbred ( $F_{ROH} = 0.053$ ), indicating its historical isolation, while Naemi displayed minimal recent inbreeding ( $F_{ROH} = 0.003$ ). The heterozygosity-based inbreeding coefficient ( $F_{IS}$ ) was highest for Najdi ( $F_{IS} = 0.092$ ), followed by Harri ( $F_{IS} = 0.027$ ), indicating a greater level of inbreeding, while Naemi ( $F_{IS} = -0.055$ ) showed signs of outbreeding.

**Conclusion:** Principal component (PC) and admixture analyses distinctly separated Najdi from the other two breeds, reflecting its unique genetic identity. Wright's fixation index ( $F_{ST}$ ) further affirmed genetic differentiation between Najdi and Naemi. Moreover, 336 high- $F_{ST}$  SNPs were identified, which were linked to breed-specific genetic signatures. Our study reveals genetic diversity among local breeds and highlights the need for conservation efforts for the Najdi sheep.

## KEYWORDS

*Ovis aries*, Najdi, Naemi, Harri, single-nucleotide polymorphism genotyping, genetic diversity, SNP, Saudi Arabia

## Introduction

The livestock segment is a key economic component of the Kingdom of Saudi Arabia (KSA), with prominent reliance on domestic production and imports to meet the high demand of sheep meat in daily food habits. Nearly 80% of the rural population living in the remote villages of KSA is actively engaged in animal husbandry, participating as nomadic pastoralists, agro-pastoral communities, or through involvement in the livestock value chain (Aldosari, 2018). The harsh climatic conditions and low soil quality limit the diversification of the farming system in KSA. Hence, farmers have widely adopted to practice the rearing of sheep and goats, which also provides stability for small farm establishments. The recent livestock data released by the General Authority of Statistics indicate that KSA is home to 20.5 million sheep, of which 19 million were categorized under traditional holdings, playing a vital role in the country's agricultural economy (MEWA, 2023).

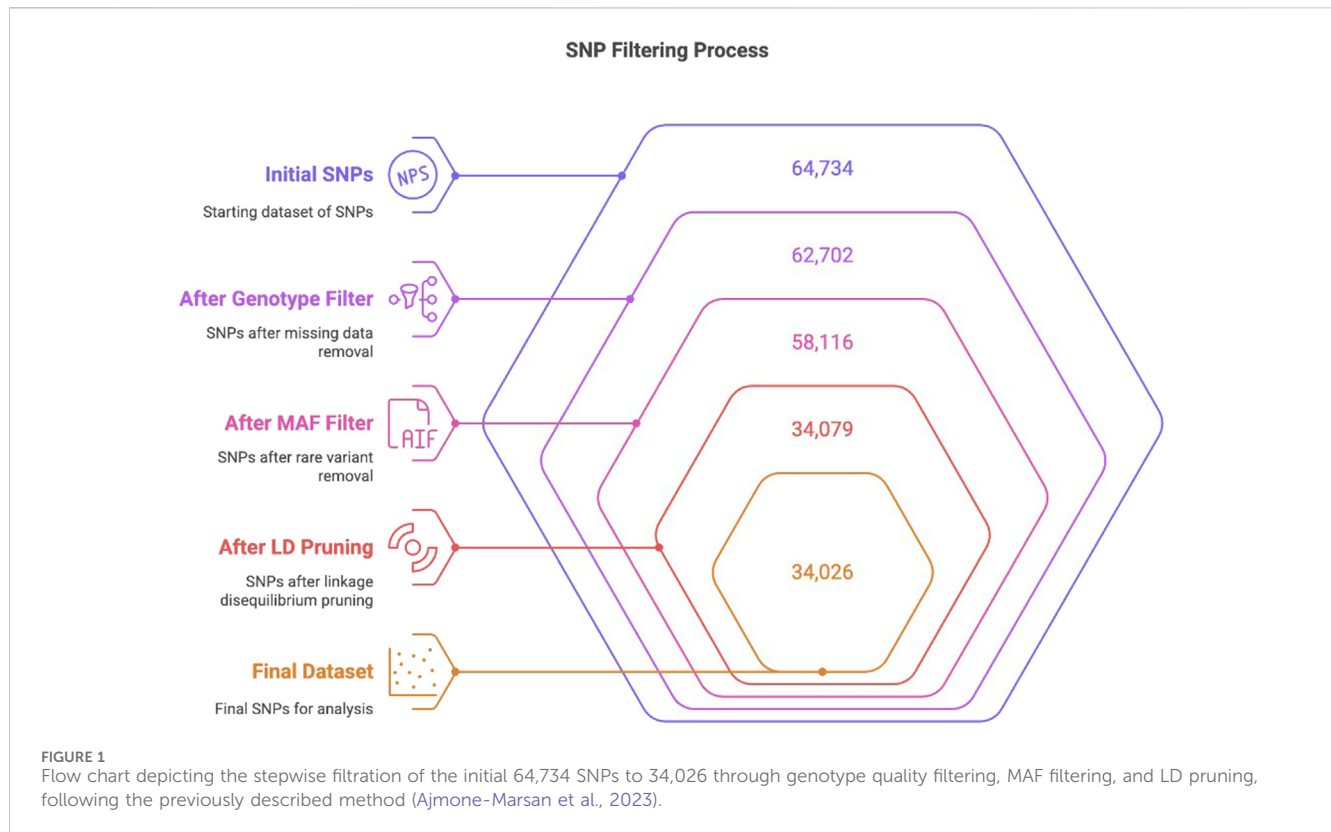
KSA has state-of-the-art facilities implemented for the commercial production of poultry and dairy farms. Sheep rearing is predominately practiced using traditional methods due to the adverse agro-climatic conditions. Despite these challenges, it provides direct employment to many people and helps sustain Saudi Arabia's sociocultural heritage (MEWA, 2021; FAO, 2023b; Alhafi-Alotaibi et al., 2025). Sheep production through organic projects constantly decreased from 16.07 million in 2017 to 12 million in 2021. This decline has been attributed to limited access to pastures, high feed costs, traditional farming practices, and the impacts of climate change (Alhafi-Alotaibi et al., 2025). Despite ongoing efforts, the acceptance of modern breeding systems by traditional sheep breeders remains a significant challenge in KSA. To fill the production gap, under the umbrella of VISION 2030 of KSA, the Sustainable Rural Agricultural Development Program (Reef) (2019–2025) was launched, aiming for transformation in animal husbandry practices (MEWA, 2021). As a result of these ongoing national initiatives, Saudi Arabia achieved a 61% level of national self-sufficiency in red meat production in 2023, reflecting growing domestic capacity to meet food security targets (MEWA, 2023). In this context, understanding and conserving the genetic and productive traits of local breeds becomes increasingly important. Among the different indigenous sheep breeds, the Najdi is believed to have originated in central KSA. It is now also found in the eastern regions of the country and along the border with Kuwait. The breed is known for its distinctive characteristics, economic value, and historical significance (Alamer and Al-hozab, 2004). Phenotypically, Najdi sheep can be identified by their black bodies with white heads, black and white rings around the neck, and their legs and tails have white spots on their legs and tails. They have long legs, fatty tails, convex noses, and pendulous ears. Their average weight and height are 62 kg and 86 cm for male sheep and 45.5 kg and 76 cm for female sheep. Najdi is well adapted to tolerate water shortages and drought and is highly suited to living under desert conditions. Traditionally, the Najdi was reared for its milk and wool but is now more in demand for its meat (MEWA, 2021; FAO, 2022).

Naemi is another well-known breed of sheep that is fully adapted to the semi-arid or arid conditions of KSA and the surrounding Gulf countries. Naemi sheep are fat-tailed and are distinguished by their unique dark brown heads, ears, and necks. They were chiefly bred

for meat; however, they also possess greater potential for milk production and exhibit outstanding drought resistance (Alamer and Al-hozab, 2004; FAO, 2022). The height of rams and ewes of Naemi varies between 60 and 80 cm and 65 and 70 cm, respectively, while body weight ranges between 60 and 90 kg for rams and 30 and 50 kg for ewes. In contrast, Harri is another well-regarded breed of sheep in KSA, primarily raised for meat production, and it is a fat-tailed and coarse-wooled breed named after the volcanic Harat region in the northwest of the country. Being well adapted for high temperatures and poor diet, Harri is most prevalent in regions like the Sarawat Mountains, Tehama, the Hejaz plains, and Qassim and is now also found in Jeddah, Makkah, and Madinah (FAO, 2022). The Harri breed stands out from other regional breeds due to its mostly white body and head, the lack of horns in male sheep, and a high twinning rate. Its average body size is medium; height and weight range from 60 to 70 cm and from 40 to 60 kg, respectively. Among the abovementioned three breeds, Najdi is large-sized, with body measurements approximately 10.9% larger than Naemi and 11.4% larger than Harri sheep, making them the largest among local breeds and reinforcing their economic importance in meat and wool production (El-Zarei et al., 2023).

The recent data released by FAO on the risk status of the world's mammalian breeds indicated an accelerated decrease in global biodiversity, including 30% of sheep breeds at the largest proportion at risk and 107 breeds of sheep already extinct globally (FAO, 2023a). Indigenous sheep breeds thriving under harsh conditions, tolerate nutritional fluctuations, and resist diseases and parasites are valuable resources—not only for understanding their genetic adaptation and biodiversity but also for supporting their biological conservation (Adam et al., 2015). Despite the prominence of Najdi, Naemi, and Harri in KSA, studies encompassing their genetic diversity and conservation are less studied. Microsatellite analysis of Najdi, Naemi, and Harri breeds exhibited considerable genetic polymorphism among them (Musthafa et al., 2012; Al-Atiyat et al., 2018). RAPD analysis has attributed Najdi as a distinct breed from Naemi and Harri (Adam et al., 2015). A recent study on the molecular profiling between Najdi and Naemi demonstrated significant activation of heat tolerance genes (*HSP90AB1*, *HSPB6*, *HSF1*, *STIP1*, *HSP60*, *HSP90*, and *HSPB1*) in Najdi sheep (Samara et al., 2024). Meanwhile, Naemi showed lower within-breed homogeneity than Harri, indicating higher genetic variability (Sabir et al., 2013). Overall, Najdi, Naemi, and Harri breeds in KSA showed high diversity. Najdi possessed genetically distinct characteristics, and no bottleneck was observed. Naemi exhibited low homogeneity, but it showed high polymorphism. On the other hand, Harri showed high homogeneity and low inbreeding, highlighting the need for comparative genomic analysis to better understand their population structure and evolutionary history (Mahmoud et al., 2020). To date, no genome-wide SNP analyses have been conducted on these local breeds. A comprehensive review of regional sheep genetics summarized microsatellite-based diversity in Awassi sheep, underscoring the need for genome-wide analyses in KSA breeds (Meydan et al., 2024).

Advancements in genomic technology, particularly single-nucleotide polymorphism (SNP) genotyping, have revolutionized the area of livestock genetics, viz., population genetics, genetic diversity, inbreeding coefficients, and selection patterns (Kanaka



et al., 2023). SNP-based genome-wide association studies (GWASs) have successfully identified genes associated with economically important traits such as growth rate, carcass quality, and reproductive efficiency in various sheep breeds (Abied et al., 2020). SNP genotyping provides high-resolution insights into genetic variation, offering a more precise approach to selective breeding and conservation efforts (Deniskova et al., 2019). In this study, we aim to perform a comparative genomic analysis of local sheep breeds (Najdi, Naemi, and Harri) through genome-wide SNP genotyping to comprehensively analyze their genetic diversity and population structure. We also investigated their genomic characteristics, inbreeding coefficients, and effective population size to provide essential insights into their adaptation and resilience. The novelty of this work relies on the fact that local sheep breeds in Saudi Arabia have not yet been investigated from the perspective of SNP genotyping.

## Materials and methods

### Collection of blood samples

Blood samples were collected from Najdi ( $n = 33$ ), Naemi ( $n = 24$ ), and Harri ( $n = 39$ ) across different regions in Saudi Arabia. Representative images of the three breeds are shown in [Supplementary Figure S8](#). Under the guidance of a trained veterinarian, ~5 mL blood was drawn aseptically from the jugular vein and collected in the vacutainer tubes. Ethical clearance (KSU-SE-24-58) for this work was approved by the Research Ethics Committee of King Saud University. Samples were transported to

the laboratory at 4 °C, processed for initial DNA isolation within 24 h, subsequently processed for genomic DNA isolation, and stored at -20 °C until further processing.

### Extraction of DNA

DNA was extracted using the Maxwell® RSC Blood DNA Kit (Promega, United States; Cat. AS1400) and the Quantus™ Fluorometer (Promega, United States; Model E6150), following the manufacturer's protocol. DNA concentration was quantified, and a concentration of 50 ng/μL was ensured to be suitable for downstream genotyping.

### SNP genotyping and quality control

Genotyping was performed using the OvineSNP50 whole-genome genotyping kit (Catalog No. WG-420-1001, Illumina, United States), which captured 64,734 SNP markers distributed across the ovine genome. Genotyping was performed at the Genome Laboratory in the Ministry of Environment, Water, and Agriculture, following the Infinium HTS Assay protocol by Illumina. Data analysis was performed using GenomeStudio software version 2.0.5 (Illumina, United States). Post-implementation of genotype quality filtration, including the exclusion of SNPs with significant missing data, provided 62,702 SNPs. Subsequently, minor allele frequency (MAF) filtration was performed at a threshold of 0.01 to remove rare variants. SNPs that passed QC filters ( $MAF \geq 0.01$ ) generated 58,116 SNPs. Subsequently, linkage disequilibrium (LD)

analysis and pruning were performed, resulting in the filtration of 34,079 SNPs. Extreme heterozygosity filtering was performed using PLINK's (`--het`) function. One Harri individual (H16) that exhibited extreme observed heterozygosity ( $H_o$ ) was removed, and the final dataset included 34,026 SNPs across 95 individuals. This final collection of SNPs was preserved for all ensuing analysis (Ajmone-Marsan et al., 2023) (Figure 1).

To compute MAFs for Harri, Naemi, and Najdi breeds, PLINK v1.9 (`--freq`) was used. Individuals were grouped using breed-specific sample files. MAF values were categorized into four bins, viz., fixed (0.00), rare ( $>0$  to  $<0.05$ ), intermediate ( $\geq 0.05$  to  $<0.10$ ), and common ( $\geq 0.10$  to  $\leq 0.50$ ). The distribution was visualized using Python 3.10.12 to compare MAF patterns across breeds.

## Run of homozygosity (ROH)

The run of homozygosity (ROH) was identified using PLINK v1.9 (Chang et al., 2015) with the (`--homozyg`) command, applying customized thresholds to capture both short and long ROH segments. ROH was defined as continuous homozygous regions spanning at least 1 Mb (`--homozyg-kb 1000`) and containing a minimum of 50 SNPs per segment (`--homozyg-snp 50`). Restricted gaps between consecutive SNPs to 100 kb (`--homozyg-gap 100`) were identified to prevent artificial inflation of ROH segments. Additionally, two missing genotype calls per segment (`--homozyg-window-missing 2`) were allowed to account for possible genotyping errors. The inbreeding coefficient ( $F_{ROH}$ ) based on the ROH was further determined following the earlier method (Ceballos et al., 2018). To assess the robustness of the results and explore the effect of altering ROH length criteria, the analysis was also utilized using alternate minimum ROH lengths of 300 kb and 500 kb.

## Genetic diversity determination

To assess within-population genetic diversity, observed heterozygosity ( $H_o$ ) and unbiased expected heterozygosity ( $H_e$ ) were determined using the (`--het`) command in PLINK (Chang et al., 2015). Additionally, the inbreeding coefficient ( $F_{IS}$ ) based on the unbiased expected heterozygosity was also estimated (Ajmone-Marsan et al., 2023). The results were analyzed and visualized using Python 3.10.12 to evaluate deviations in heterozygosity across individuals.  $F_{IS}$  was calculated as follows:

$$F_{IS} = \frac{H_e - H_o}{H_e},$$

where  $H_e$  represents the expected heterozygosity and  $H_o$  represents the observed heterozygosity.

To estimate the effective population size ( $N_e$ ), LD was calculated using PLINK v1.9 (Purcell et al., 2007; Ajmone-Marsan et al., 2023), based on the  $r^2$  statistics (Hill and Robertson, 1966). Pairwise LD ( $r^2$ ) was calculated within a 1 Mb window using the PLINK flags (`--r2 --ld-window-r2 0 --ld-w`). The  $N_e$  trends for each breed were visualized in R 4.3.2 using the ggplot2 package, with initial errors in PLINK flags (`--ne`) resolved by refining input parameters and modifying the model fitting script.

## Population structure analysis

Principal component analysis (PCA) was performed to assess the genetic relationships among individuals from the three sheep breeds. The genomic relationship matrix (GRM) was constructed, and PCs were extracted to capture population structure patterns. The genomic relationship matrix was calculated following the earlier described approaches (Purcell et al., 2007; Yang et al., 2011). Visualization of PCA results was performed in R 4.3.2 (R-CORE Team, 2021) using statistical and bioinformatics packages, including ggplot2 for graphical representation.

Population structure and individual ancestry proportions were analyzed using sparse nonnegative matrix factorization (sNMF) in the LEA package (Frichot and François, 2015; R-CORE Team, 2021). Genotype data were processed to test  $K = 1$  to 10, with cross-entropy used to identify the optimal  $K$ . The best runs were selected based on the lowest cross-entropy values across multiple repetitions. Ancestry proportions (Q-matrices) were extracted for the optimal  $K$ -values and visualized using ggplot2. Stacked bar plots illustrate admixture patterns, aiding in the interpretation of genetic relationships and population structure. Additionally, a phylogenetic network was constructed using SplitsTree CE 6.0.0 to visualize relationships among individuals based on genome-wide SNPs. The resulting tree is presented in Supplementary Figure S7.

Wright's fixation index ( $F_{ST}$ ) was estimated to quantify the genetic differentiation between populations (Weir and Cockerham, 1984). The calculations were performed using the GRM following the previously described method (Yang et al., 2011). Visualization of  $F_{ST}$  values across populations was conducted in R 4.3.2 (R-CORE Team, 2021) using ggplot2 to generate a heatmap and a Manhattan plot, facilitating the interpretation of genetic differentiation patterns.

## Analysis of shared haplotypes

To assess haplotype sharing, we conducted identity-by-descent (IBD) segment detection using Refined IBD v4.1 (Browning and Browning, 2013; Revelo et al., 2022). Genotypes were phased to infer haplotype structure using Beagle v4.1, and IBD identification was carried out with a 40-SNP window, a minimum segment length of 1.5 cm, and a minimum LOD score of 3. Results were summarized across breed pairs, and Circos visualization was performed using the circlize package in R (Gu et al., 2014).

## Results

### MAF distribution revealed genetic variation among sheep breeds

The MAF distribution exhibited distinct patterns of genetic variations among Harri, Naemi, and Najdi. The majority of SNPs congregated within the common MAF range ( $\geq 0.10$ – $\leq 0.50$ ), with Naemi exhibiting the highest proportion (87.2%), followed by Harri (78.2%) and Najdi (71.2%) (Figure 2; Table 1). Najdi retained the highest proportions of fixed (9.9%) and rare alleles (10.4%). Relatively, Harri showed 3.96% and 10.4% of fixed and rare

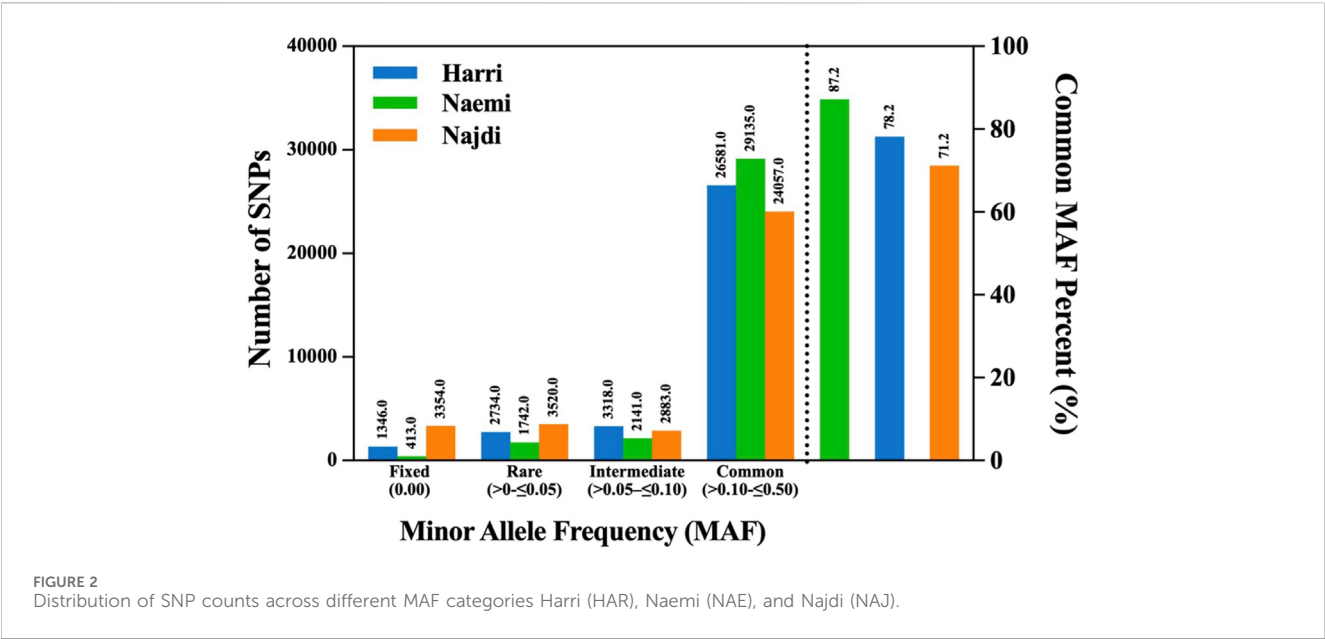


TABLE 1 Distribution of MAF in three breeds of sheep in KSA.

| MAF                                    | Harri (%) | Naemi (%) | Najdi (%) |
|----------------------------------------|-----------|-----------|-----------|
| Fixed (0.00)                           | 3.96      | 1.24      | 9.92      |
| Rare ( $\geq 0-\leq 0.05$ )            | 8.05      | 5.21      | 10.41     |
| Intermediate ( $\geq 0.05-\leq 0.10$ ) | 9.76      | 6.40      | 8.53      |
| Common ( $\geq 0.10-\leq 0.50$ )       | 78.20     | 87.20     | 71.20     |

alleles, respectively. However, the lowest proportions of fixed alleles (1.24%) and rare alleles (5.21%) were found for Naemi. Harri showed the highest intermediate MAF profile (9.76%) among all breeds (Figure 2; Table 1).

ROH exhibited inbreeding patterns in sheep breeds

The total length of ROH and the number of ROH segments were evaluated to determine the genetic diversity and inbreeding patterns in Harri, Najdi, and Naemi (Table 2). The ROH-based inbreeding coefficient ( $F_{ROH}$ ) was minimal (0.003) in Naemi. Historical inbreeding was more evident in Najdi (0.053) and Harri (0.023) breeds.

Najdi exhibited the longest total ROH length (143.37 Mb) and the highest mean ROH count per individual (21.18). Harri displayed a moderate ROH profile with a total ROH length of 61.78 Mb. Naemi showed the lowest ROH accumulation (7.99 Mb). Moreover, the data from ROH analysis at lower thresholds of 300 and 500 kb are presented in Supplementary Table S1.

Heterozygosity analysis indicated the genetic diversity in sheep breeds

The distribution of heterogeneity-based inbreeding coefficient ( $F_{IS}$ ) across the population followed a bimodal pattern and high-frequency peaks spanning from -0.05 to 0.05. The majority of  $F_{IS}$  values range from -0.05 to 0.20, with a skew toward positive values (Figure 3A). Najdi exhibited the highest inbreeding coefficient ( $F_{IS} = 0.092$ ), followed by Harri ( $F_{IS} = 0.027$ ) and Naemi ( $F_{IS} = -0.055$ ) (Figure 3B). Genetic diversity among Harri, Najdi, and Naemi based on expected heterozygosity ( $He$ ), observed heterozygosity ( $Ho$ ), and

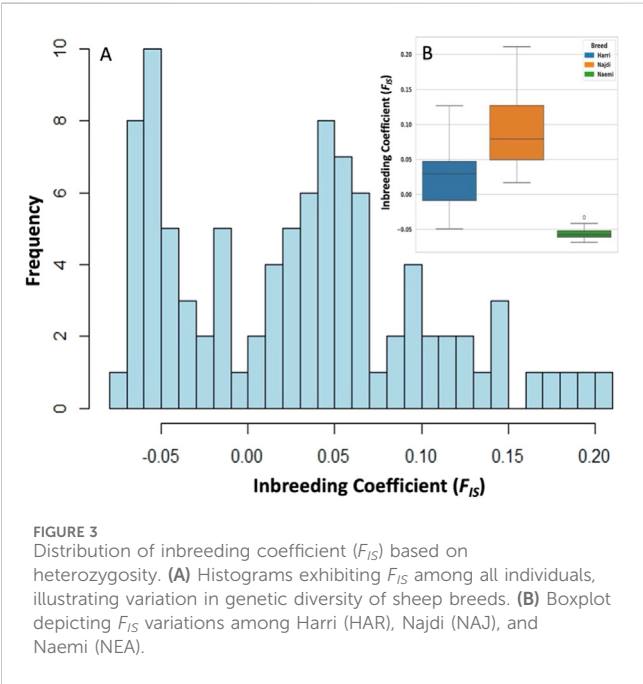


TABLE 2 ROH and inbreeding coefficient determined in three breeds of sheep in KSA.

| Breed | n  | Total ROH length (Mb) | Mean ROH count | $F_{ROH}$ |
|-------|----|-----------------------|----------------|-----------|
| Harri | 38 | 61.78                 | 11.23          | 0.023     |
| Najdi | 33 | 143.37                | 21.18          | 0.053     |
| Naemi | 24 | 7.99                  | 1.96           | 0.003     |

Number of individuals (n) and inbreeding coefficient ( $F_{ROH}$ ) are based on 1 Mb threshold of ROH.

TABLE 3 Genetic diversity analysis among sheep breeds.

| Breed | n  | $H_o$ | $H_e$ | $N_e$ |
|-------|----|-------|-------|-------|
| Harri | 38 | 0.342 | 0.352 | 6.08  |
| Najdi | 33 | 0.319 | 0.352 | 3.3   |
| Naemi | 24 | 0.371 | 0.352 | 6.16  |

$H_o$ , observed heterozygosity;  $H_e$ , expected heterozygosity;  $N_e$ , effective population size.

effective population size ( $N_e$ ) is shown in Table 3. The lowest effective population size was found in Najdi ( $N_e = 3.30$ ), with a subsequent increase observed in Harri ( $N_e = 6.08$ ) and Naemi ( $N_e = 6.16$ ).

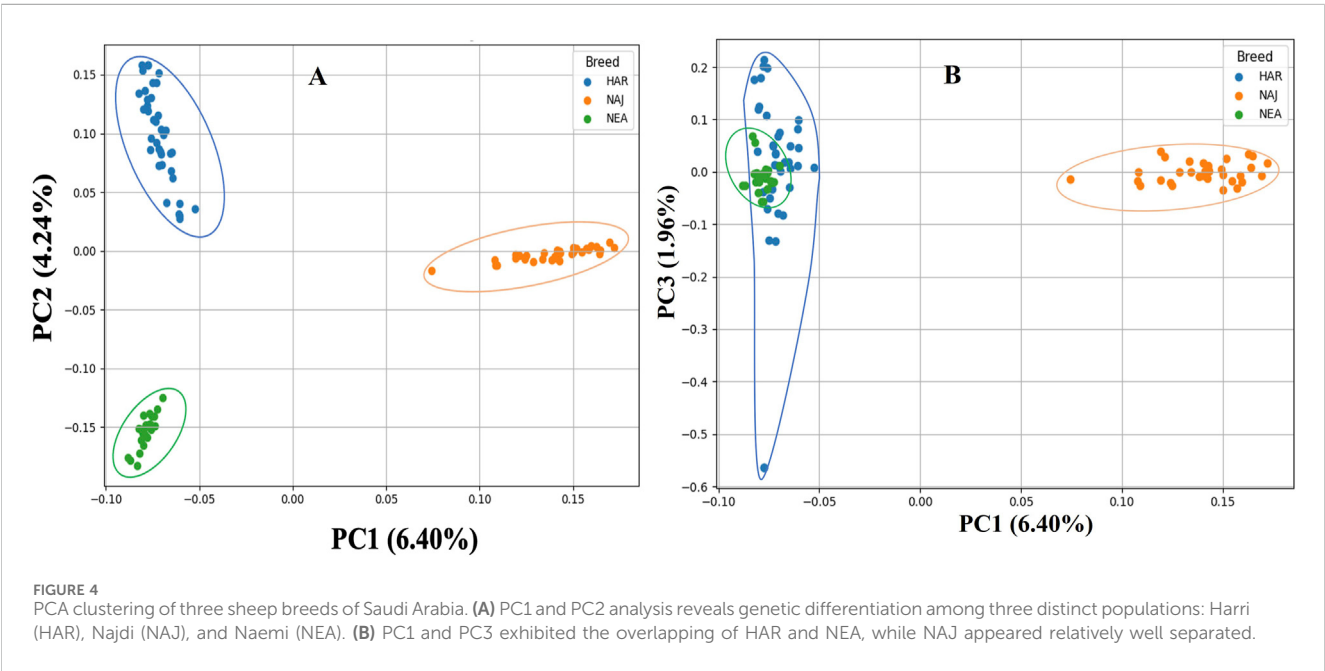
Population structure of Saudi sheep breeds

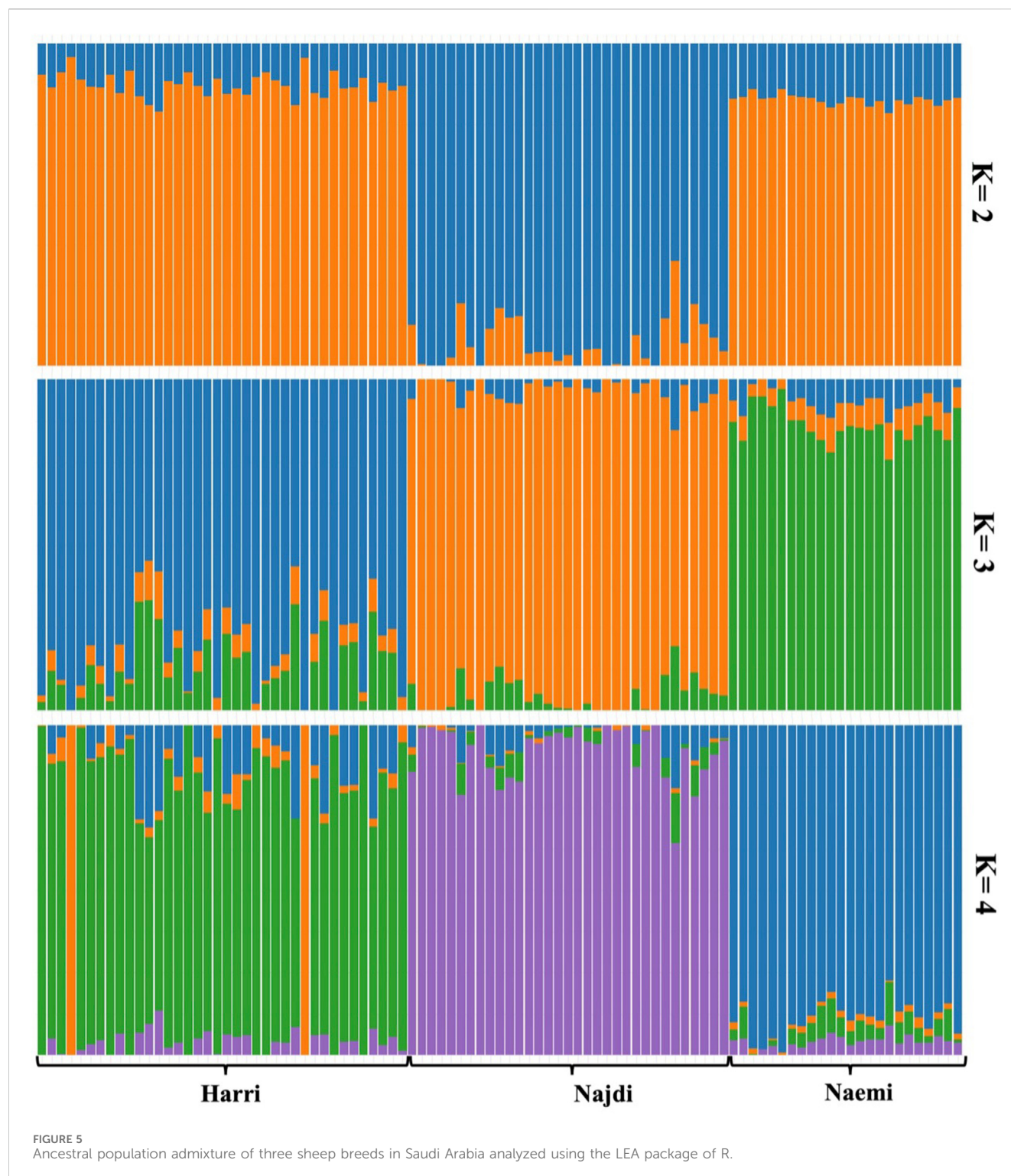
PCA was performed on genome-wide SNP data from 95 individuals representing three distinct breeds of sheep populations. The first 10 principal components explained decreasing proportions of genetic variance, with PC1 (6.40%), PC2 (4.24%), and PC3 (1.96%) collectively accounting for 12.6% of total variation. All components relatively exhibited a small part of the total variance (Supplementary Figure S1,S2). The scatter plot of

PC1 vs. PC2 revealed clear separation among the sheep breeds (Figure 4A). Subsequent analysis of the breeds using PC1 vs. PC3 revealed the overlapping of Harri and Naemi, while Najdi appeared relatively well separated. PC1 vs. PC4, PC1 vs. PC5, and PC2 vs. PC3 analyses further showed overlapping patterns between Harri and Naemi (Supplementary Figures S3–S5).

Admixture analysis was performed using the LEA package at  $K = 2$  to  $K = 4$  levels, which exhibited that  $K = 4$  had the lowest cross-validation error (Supplementary Figure S6). At  $K = 2$ , a clear genetic separation of Najdi was observed relative to Harri and Naemi. At this level, Najdi was predominantly assigned to a single cluster (92.79%), while Harri (87.21%) and Naemi (82.58%) were grouped in different clusters (Figure 5). At  $K = 3$ , Harri individuals clustered mainly in one group (81.5%), while Naemi sheep were strongly associated with another group (86.14%). The Najdi breed remained genetically uniform (92%). However, at  $K = 4$ , each of the three breeds showed a distinct ancestry pattern. In particular, Najdi individuals were assigned to a single unique cluster (92%). Naemi was dominated by a separate cluster (89.03%), while Harri was associated with a distinct cluster (77.31%) but displayed traces of admixture.

Haplotype sharing analysis revealed varying degrees of IBD segment exchange between the three sheep breeds (Figure 6). The total shared IBD lengths were the highest between Harri and Najdi (673.1 cm), followed by Najdi and Naemi (552.6 cm), and lowest between Harri and Naemi (503.2 cm).



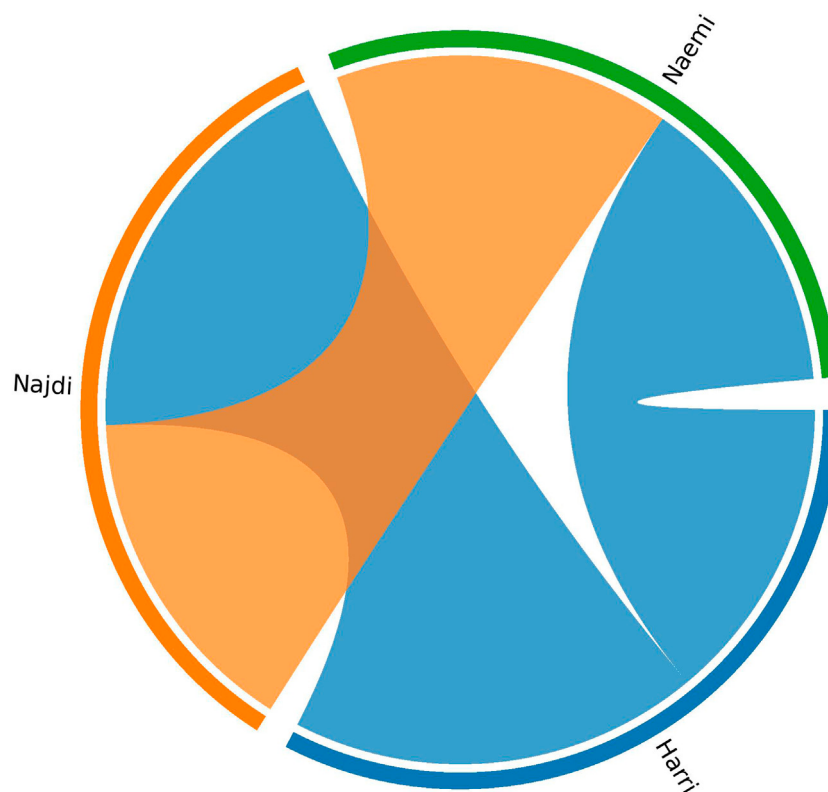


### $F_{ST}$ analysis affirmed genetic differentiation between breeds

To explore the genetic relationships among individual sheep and between breeds, a relatedness heatmap was generated using pairwise genomic similarity values for 95 individuals. Distinct block-wise clustering was observed, and a gradient of blue and red regions based on  $F_{ST}$  values reflects lower and higher similarity across breeds

(Figure 7). Moreover, the pairwise  $F_{ST}$  values reveal moderate levels of genetic differentiation among the three breeds. The highest  $F_{ST}$  value was found between Najdi and Naemi (0.1071), followed by Najdi and Harri (0.0999). The lowest  $F_{ST}$  value was observed between Harri and Naemi (0.0671) (Supplementary Table S2).

To further investigate the genomic regions contributing to  $F_{ST}$  differentiation, the Manhattan plot was developed (Figure 8A). The plot displays  $F_{ST}$  values across autosomes, with several SNPs



**FIGURE 6**  
Circos plot showing haplotype sharing among three Saudi sheep breeds based on IBD analysis.

exceeding the threshold ( $F_{ST} = 0.99$ ), suggesting potential loci under breed-specific selection. A total of 336 high  $F_{ST}$  SNPs were identified. In particular, 119 SNPs for Najdi, 112 SNPs for Naemi, and 105 SNPs for Harri were found. Intronic SNPs were predominant in all breeds, with the highest proportion in Harri (71%), followed by Naemi (61%) and Najdi (57%). Intergenic SNPs accounted for approximately 19%–23% of the total SNPs, with Naemi showing the highest proportion (23%). Upstream gene SNPs were relatively low across all breeds, with Najdi having the highest percentage (7%) compared to Naemi (5%) and Harri (4%) (Figure 8B).

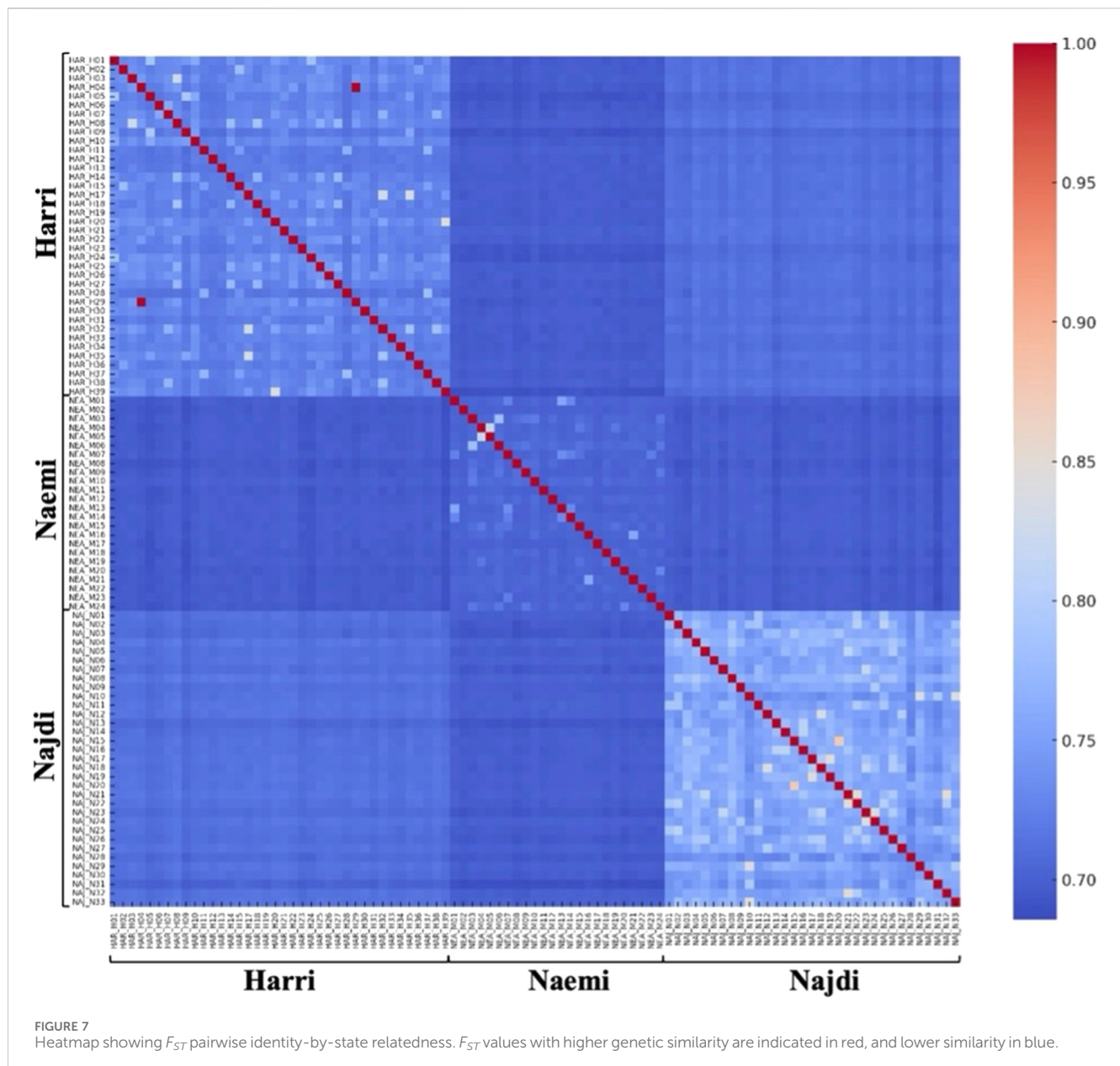
## Discussion

We have performed genomic analysis of three local sheep breeds (Najdi, Naemi, and Harri) of KSA, aiming to evaluate differences in their genetic diversity, population structure, inbreeding coefficients, and effective population sizes to gain key insights into their adaptive potential and genetic resilience. Relative to previous findings using microsatellites, our approach using the genome-wide SNP genotyping has offered novel insights, unraveling a holistic view of the genetic diversity of above the breeds (Al-Atiyat et al., 2018; Mahmoud et al., 2020; El-Zarei et al., 2023).

The distribution of MAF indicated a unique pattern of genetic variations across the three sheep breeds. In particular, Naemi exhibited the highest frequency of common SNPs, which accounted for 87.2%. On the other hand, Harri and Najdi

showed 78.2% and 71.2%, specifying a balanced allele distribution and enhanced genetic diversity in Naemi. Although Najdi exhibited greater frequencies of rare and fixed (10.41% and 9.92%) allele distribution, Harri showed 8.1% and 3.96% frequencies, respectively, indicating decreased heterozygosity and putative bottlenecks in Najdi. In addition, Harri showed intermediate patterns of MAF SNPs pointing toward a moderate level of genetic diversity. Nonetheless, our study exhibited an overall mean of 5.04% fixed alleles across populations, which is lower than the average 8.10% of fixed alleles in five sheep breeds of Ethiopia/Africa (Adilo, Arsi-Bale, Blackhead Somali, Horro, and Menz) (Edea et al., 2017) but higher than Merino (3%) and Corriedale (4%) sheep breeds (Grasso et al., 2014). Hence, the accumulation of fixed alleles typically indicates reduced gene flow and historical isolation. The low fixation level in Naemi indicates potential genetic introgression or outbreeding.

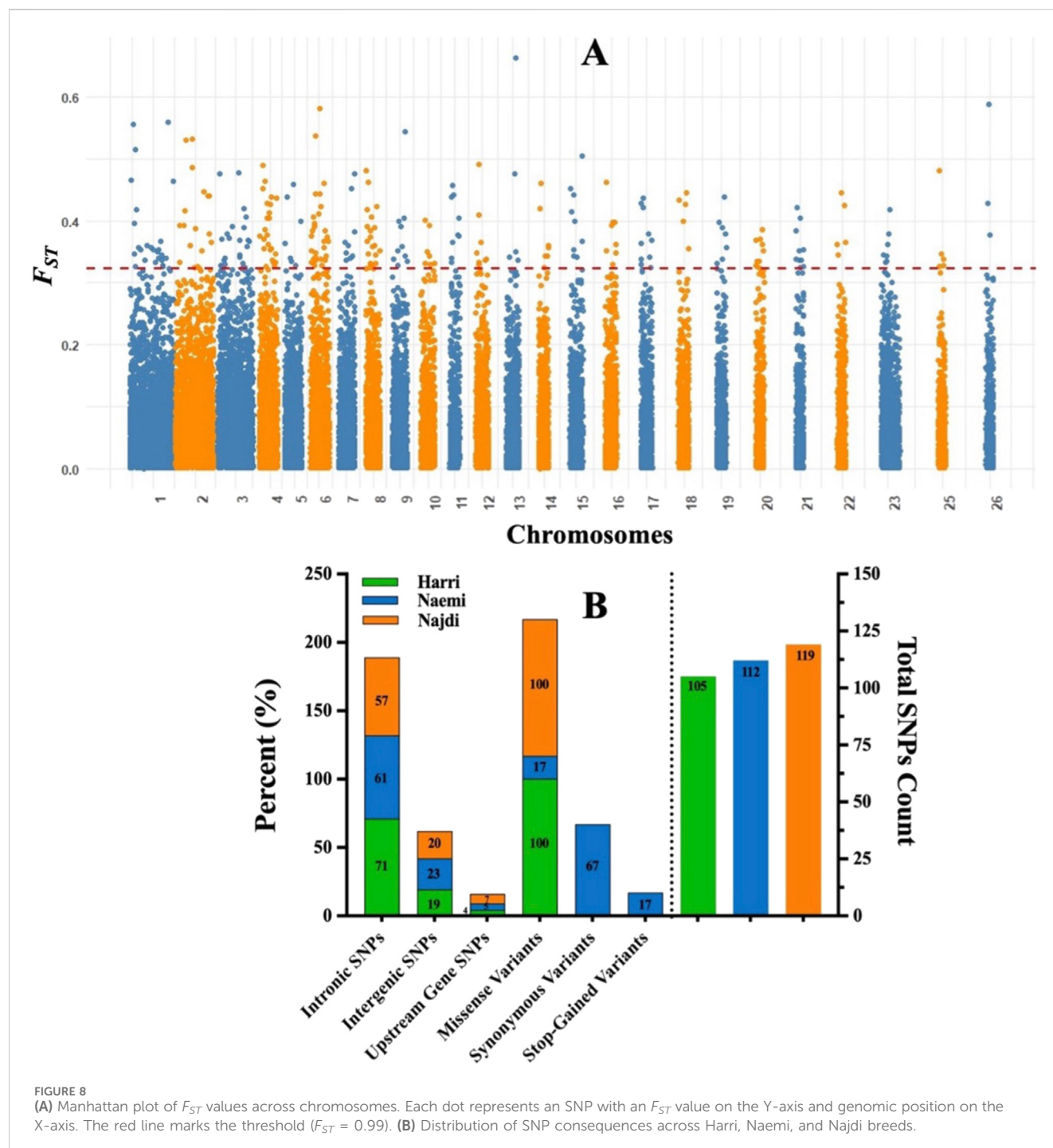
Given the robustness of genomic data over classical pedigree analysis, the values of  $H_o$  and  $H_e$  were determined, aiding in the accurate quantification of inbreeding without the need for pedigree data (Curik et al., 2014; Druet and Gautier, 2017). In this connection, we evaluated the inbreeding coefficient ( $F_{IS}$ ) that has exhibited a negative value (−0.055) for Naemi, while Harri and Najdi showed  $F_{IS}$  of 0.027 and 0.092, respectively. The observed data for Harri and Najdi were  $F_{IS} > 0$ , which clearly indicated that both of them have inbreeding in their population, as also evident by lower values of  $H_o$  (0.342 and 0.319). In contrast, Naemi has  $F_{IS} < 0$  [ $H_o = 0.371$  higher than  $H_e$  (0.352)], affirming that its population has greater



heterozygosity than the expected population, such as with outbreeding or cross-breeding (Ajmone-Marsan et al., 2023).

The limitations of  $F_{IS}$  to determine inbreeding in more ancient populations have further prompted us to investigate the genomic coverage using the ROH, which was selectively used to analyze inbreeding when SNPs and WGS data are available. Using frequency and length of ROH, understanding population history, including bottlenecks, and identification of selection signatures were advantageous (Bruford et al., 2015; Ceballos et al., 2018; Meyermans et al., 2020). The total length of ROH in Najdi and Harri was quite long, which clearly indicated recent inbreeding and low genetic diversity in these breeds, possibly due to deleterious homozygous variants (Pryce et al., 2014; Peripolli et al., 2017). In contrast, the total length of ROH in Naemi was low, signifying a historical reduction in their population size, with little recent inbreeding and greater genetic diversity (Peripolli et al., 2017).

Najdi, Naemi, and Harri ROHs revealed specific variations in their inbreeding pattern. Among the three breeds, Najdi and Harri showed the highest inbreeding coefficients of  $F_{ROH} = 0.053$  and  $0.023$ , respectively, indicating historical isolation or intensive selection. Similar to these patterns, high  $F_{ROH}$  values (0.06 and 0.14) were also reported for Czech sheep breeds (Sumava and Wallachian) (Machová et al., 2023). Moreover, the  $F_{ROH}$  values of Najdi and Harri reflected a comparable level of inbreeding as also found in Sopravissana (0.052), Ovino delle Langhe (0.052), and Valle del Belice (0.067) indigenous sheep breeds of Italy (Persichilli et al., 2021). To characterize the divergence, PCA-based population structure analysis was performed. PC1 vs. PC2 resulted in the unique clustering of all three breeds. PC1 vs. PC3 showed genetic isolation of Najdi, while Harri and Naemi overlapped with each other, indicating a shared genetic background. To verify the ancestral, pure, and hybrid



**FIGURE 8**  
**(A)** Manhattan plot of  $F_{ST}$  values across chromosomes. Each dot represents an SNP with an  $F_{ST}$  value on the Y-axis and genomic position on the X-axis. The red line marks the threshold ( $F_{ST} = 0.99$ ). **(B)** Distribution of SNP consequences across Harri, Naemi, and Najdi breeds.

populations, we further performed the admixture analysis (Álvarez et al., 2004; Al-Atiyat et al., 2018). It is clearly evident that nearly all individuals of Najdi were characterized as a separate gene pool, forming a consistent, homogenous cluster, affirming their genetic distinctiveness. In contrast, Harri and Naemi showed varying levels of mixing, which is in close agreement with previous reports based on SNP analysis of several regional sheep populations, where gene flow and isolation have led to complex genetic substructures (Ceccobelli et al., 2023; Machová et al., 2023). This pattern is further supported by IBD analysis, where Najdi showed the highest intra-breed haplotype sharing and minimal overlap with

Naemi, confirming its genetic isolation. These trends align with previous findings using IBD to detect breed-specific structure in regional sheep breeds (Revelo et al., 2022). In contrast, admixture analysis using 19 microsatellite loci reported close similarity between Najdi and the gene pool of Naemi (Al-Atiyat et al., 2018).

The genetic differentiation between breeds was further determined using  $F_{ST}$ , which also separated Najdi from Naemi by a greater value. Harri and Naemi were found closer to each other with low  $F_{ST}$  values (Supplementary Table S2). The analysis of  $F_{ST}$  using identity-by-state relatedness, as evident in the heatmap, further differentiated individual sheep belonging to different

breeds. Subsequently, the distribution of SNPs across genomic regions identified 336 highly distinct SNPs, predominantly consisting of intronic and intergenic variants at different loci, signifying breed-specific adaptations to environmental or production-related stresses. Recently, the thermo-physiological and molecular adaptations in Saudi sheep breeds emphasized the advantage of SNP markers as selection signatures to spot heat tolerance traits (Samara et al., 2024).

Although SNP50 genotyping offers high coverage, whole-genome sequencing could reveal rare structural variants. Future studies should integrate phenotype-genotype correlations and prioritize the development of a custom SNP chip for local breeds to refine breeding strategies.

## Conclusion

This study provides high-resolution genomic insights into Najdi, Naemi, and Harri sheep breeds of Saudi Arabia using 50K SNP genotyping. Assessment of key genomic parameters (MAF,  $F_{ROH}$ ,  $H_e$ ,  $F_{IS}$ ,  $F_{ST}$ , PCA, and admixture) demonstrated high inbreeding and genetic isolation in Najdi, while Naemi demonstrated the greatest genetic diversity. These findings underscore the urgent need to conserve Najdi sheep in order to preserve their unique genetic heritage.

Future research should integrate whole-genome sequencing, phenotypic assessments, and functional annotation and develop a custom SNP chip to support sustainable breeding and management of Saudi Arabia's indigenous sheep breeds.

## Data availability statement

The data presented in the study are deposited in the EVA repository, accession number PRJEB97998 available at <https://www.ebi.ac.uk/eva/?eva-study=PRJEB97998>.

## Ethics statement

The animal study was approved by King Saud University (KSU-SE-24-58). The study was conducted in accordance with the local legislation and institutional requirements.

## Author contributions

AA: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, and Writing – original draft. MI: Methodology, Validation, and Writing – review and editing. FA:

Data curation, Resources, and Writing – review and editing. QS: Conceptualization, Supervision, and Writing – review and editing. MA: Conceptualization, Supervision, and Writing – review and editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors would like to thank Deanship of scientific research in King Saud University for funding and supporting this research through the initiative of DSR Graduate Students Research Support (GSR).

## Acknowledgments

The authors would like to extend their sincere appreciation to King Saud University, Riyadh, Saudi Arabia, for its continuous support.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1646127/full#supplementary-material>

## References

- Abied, A., Bagadi, A., Bordbar, F., Pu, Y., Augustino, S. M. A., Xue, X., et al. (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
- Adam, A. A. G., Hamza, N. B., Salim, M. A. W., and Khalil, K. S. (2015). Estimation of genetic diversity between three Saudi sheep breeds using DNA markers. *Afr. J. Biotechnol.* 14 (41), 2876–2882. doi:10.5897/AJB2015.14602

- Ajmone-Marsan, P., Boettcher, P. J., Colli, L., Ginja, C., Kantanen, J., and Lenstra, J. A. (2023). *Genomic characterization of animal genetic resources – practical guide*. Rome: FAO Animal Production and Health Guidelines No. Vol. 32.
- Al-Atiyat, R. M., Aljumaah, R. S., Alshaikh, M. A., and Abudabos, A. M. (2018). Microsatellite-based genetic structure and diversity of local arabian sheep breeds. *Front. Genet.* 9, 408–2018. doi:10.3389/fgene.2018.00408
- Alamer, M., and Al-hozab, A. (2004). Effect of water deprivation and season on feed intake, body weight and thermoregulation in awassi and najdi sheep breeds in Saudi Arabia. *J. Arid Environ.* 59 (1), 71–84. doi:10.1016/j.jaridenv.2004.01.003
- Aldosari, F. O. (2018). Gender participation in sheep and goat farming in Najran, southern Saudi Arabia. *Saudi J. Biol. Sci.* 25 (1), 144–148. doi:10.1016/j.sjbs.2017.03.013
- Alhafi-Alotaibi, B., Muddassir, M., Alharbi, M., S-Aljohani, E., and Ahmad-Shah, A. (2025). Producers' perspectives of modern sheep breeds in Saudi Arabia. *Agrociencia*, 1–16. doi:10.47163/agrociencia.v59i2.3316
- Álvarez, I., Royo, L. J., Fernández, I., Gutiérrez, J. P., Gómez, E., and Goyache, F. (2004). Genetic relationships and admixture among sheep breeds from northern Spain assessed using microsatellites. *J. Animal Sci.* 82 (8), 2246–2252. doi:10.2527/2004.8282246x
- Browning, B. L., and Browning, S. R. (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* 194 (2), 459–471. doi:10.1534/genetics.113.150029
- Bruford, M. W., Ginja, C., Hoffmann, I., Joost, S., Orozco-terWengel, P., Alberto, F. J., et al. (2015). Prospects and challenges for the conservation of farm animal genomic resources, 2015–2025. *Front. Genet.* 6, 314–2015. doi:10.3389/fgene.2015.00314
- Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., and Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* 19 (4), 220–234. doi:10.1038/nrg.2017.109
- Ceccobelli, S., Landi, V., Senczuk, G., Mastrangelo, S., Sardina, M. T., Ben-Jemaa, S., et al. (2023). A comprehensive analysis of the genetic diversity and environmental adaptability in worldwide merino and Merino-derived sheep breeds. *Genet. Sel. Evol.* 55 (1), 24. doi:10.1186/s12711-023-00797-z
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4 (1), s13742-13015-10047-13748. doi:10.1186/s13742-015-0047-8
- Curik, I., Ferenčaković, M., and Sölkner, J. (2014). Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livest. Sci.* 166, 26–34. doi:10.1016/j.livsci.2014.05.034
- Deniskova, T., Dotsev, A., Lushihina, E., Shakhin, A., Kunz, E., Medugorac, I., et al. (2019). Population structure and genetic diversity of sheep breeds in the Kyrgyzstan. *Front. Genet.* 10, 1311–2019. doi:10.3389/fgene.2019.01311
- Druet, T., and Gautier, M. (2017). A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Mol. Ecol.* 26 (20), 5820–5841. doi:10.1111/mec.14324
- Edea, Z., Dessie, T., Dadi, H., Do, K.-T., and Kim, K.-S. (2017). Genetic diversity and population structure of Ethiopian sheep populations revealed by high-density SNP markers. *Front. Genet.* 8, 218–2017. doi:10.3389/fgene.2017.00218
- El-Zarei, M., A. A. L. S., S. A., A. A., and Mousa, E. (2023). Phenotypic characterization of Indigenous sheep breeds in Saudi Arabia. *J. Agric. Sci.* 15, 16. doi:10.5539/jas.v15n8p16
- FAO (2022). *Food and agriculture organization of united nations and ministry of environment, water and agriculture. Strategies for strengthening small ruminant breeding systems for a sustainable genetic improvement program in KSA*. Riyadh, Saudi Arabia: FAO. Available online at: [https://www.mewa.gov.sa/reefea/attachments/الاشية%20مربي%20صغار%20طاع%20ل/2022-Strategies%20small%20ruminant%20breeding%20systems\\_KSA\\_TC.pdf](https://www.mewa.gov.sa/reefea/attachments/الاشية%20مربي%20صغار%20طاع%20ل/2022-Strategies%20small%20ruminant%20breeding%20systems_KSA_TC.pdf) (Accessed May 21, 2025).
- FAO (2023a). *Food and Agriculture Organization of the United Nations, Commission On Genetic Resources For Food And Agriculture, Intergovernmental Technical Working Group On Animal Genetic Resources For Food And Agriculture, Twelfth Session, Rome, 18-20 January 2023, Status And Trends Of Animal Genetic Resources-2022*. Riyadh, Saudi Arabia: FAO. Available online at: <https://openknowledge.fao.org/server/api/core/bitstreams/c8121641-5385-404f-bf98-6a584bec2e7b/content> (Accessed May 21, 2025).
- FAO (2023b). *Food and agriculture organization of the united nations, FAO technical cooperation programme in the Kingdom of Saudi Arabia*. Riyadh, Saudi Arabia: FAO. doi:10.4060/cc8564en
- Frivot, E., and François, O. (2015). LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6 (8), 925–929. doi:10.1111/2041-210X.12382
- Grasso, A. N., Goldberg, V., Navajas, E. A., Iriarte, W., Gimeno, D., Aguilar, I., et al. (2014). Genomic variation and population structure detected by single nucleotide polymorphism arrays in corriedale, merino and creole sheep. *Genet. Mol. Biol.* 37 (2), 389–395. doi:10.1590/s1415-47572014000300011
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* 30 (19), 2811–2812. doi:10.1093/bioinformatics/btu393
- Hill, W. G., and Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genet. Res.* 8 (3), 269–294. doi:10.1017/S0016672300010156
- Kanaka, K. K., Sukhija, N., Goli, R. C., Singh, S., Ganguly, I., Dixit, S. P., et al. (2023). On the concepts and measures of diversity in the genomics era. *Curr. Plant Biol.* 33, 100278. doi:10.1016/j.cpb.2023.100278
- Machová, K., Marina, H., Arranz, J. J., Pelayo, R., Rychtářová, J., Milerski, M., et al. (2023). Genetic diversity of two native sheep breeds by genome-wide analysis of single nucleotide polymorphisms. *animal* 17 (1), 100690. doi:10.1016/j.animal.2022.100690
- Mahmoud, A. H., Abou-Tarboush, F. M., Rady, A., Al-Anazi, K. M., Farah, M. A., and Mohammed, O. B. (2020). Genetic variability of sheep populations of Saudi Arabia using microsatellite markers. *Indian J. Animal Res.* 54 (4), 409–412. doi:10.18805/ijar.B-775
- MEWA (2021). *Ministry of Environment, Water and Agriculture. Sustainable Rural Agricultural Development (REEF). Total numbers of animals according to health card, unpublished data. Ministry of Environment, Water and Agriculture*. Riyadh: MEWA. Available online at: <https://www.mewa.gov.sa/en/Pages/default.aspx> (Accessed May 21, 2025).
- MEWA (2023). *Ministry of environment, Water and agriculture*. Riyadh: MEWA. Available online at: <https://www.mewa.gov.sa/ar/InformationCenter/Researchs/Reports/GeneralReports/م202023%إيالحاص%إي%20%اب%20%ال%20%ك%20%اب%20%إي> (Accessed May 21, 2025).
- Meydan, H., Rashaydeh, F. S., Al-Atiyat, R. M., and Yildiz, B. (2024). Genetic diversity of awassi sheep reared in fertile crescent based on microsatellites: a review. *J. Animal Plant Sci.* 34 (5), 1099–1111. doi:10.36899/japs.2024.5.0792
- Meyermans, R., Gossen, W., Buys, N., and Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics* 21 (1), 94. doi:10.1186/s12864-020-6463-x
- Musthafa, M. M., Aljummah, R., and Alshaikh, M. (2012). Genetic diversity of najdi sheep based on microsatellite analysis. *Afr. J. Biotechnol.* 11 (83), 14868–14876. doi:10.5897/AJB12.1797
- Peripolli, E., Munari, D. P., Silva, M. V. G. B., Lima, A. L. F., Irgang, R., and Baldi, F. (2017). Runs of homozygosity: current knowledge and applications in livestock. *Anim. Genet.* 48 (3), 255–271. doi:10.1111/age.12526
- Persichilli, C., Gabriele, S., Loi, P., Agata, A. D., Riccardo, N., Paolo, A. M., et al. (2021). Genome-wide diversity of pagliarola sheep residual population and its conservation implication. *Italian J. Animal Sci.* 20 (1), 1695–1705. doi:10.1080/1828051X.2021.1970033
- Pryce, J. E., Haile-Mariam, M., Goddard, M. E., and Hayes, B. J. (2014). Identification of genomic regions associated with inbreeding depression in holstein and jersey dairy cattle. *Genet. Sel. Evol.* 46 (1), 71. doi:10.1186/s12711-014-0071-7
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575. doi:10.1086/519795
- R-CORE Team (2021). *R: a language and environment for statistical computing*. Vienna, Austria: R foundation for statistical computing. Available online at: <https://www.R-project.org/>.
- Revelo, H. A., Landi, V., López-Alvarez, D., Palacios, Y. A., Paiva, S. R., McManus, C., et al. (2022). New insight into the genome-wide diversity and admixture of six Colombian sheep populations. *Genes* 13 (8), 1415. doi:10.3390/genes13081415
- Sabir, J. M., Awad, N. S., Alhallous, E. E., and Mutawakil, M. H. (2013). Assessment of genetic diversity among Saudi sheep breeds for characterization and conservation purposes. *Glob. Vet.* 11 (5), 634–639. doi:10.5829/idosi.gv.2013.11.5.8128
- Samara, E. M., Bahadi, M. A., Khan, M. A., Al-Badwi, M. A., Abdoun, K. A., Afzal, M., et al. (2024). Thermo-physiological and molecular profiling of two Indigenous purebred saudi sheep under acute heat stress conditions. *Trop. Animal Sci. J.* 47 (3), 300–311. doi:10.5398/tasj.2024.47.3.300
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution* 38 (6), 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x
- Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88 (1), 76–82. doi:10.1016/j.ajhg.2010.11.011