



## OPEN ACCESS

## EDITED BY

Joel Ochieng,  
University of Nairobi, Kenya

## REVIEWED BY

Fuyi Xu,  
University of Tennessee Health Science Center  
(UTHSC), United States  
Anne W. T. Muigai,  
National Defence University-Kenya, Kenya

## \*CORRESPONDENCE

Oumer Sheriff,  
✉ soumer74@yahoo.com

RECEIVED 09 December 2023

ACCEPTED 16 April 2024

PUBLISHED 24 May 2024

## CITATION

Sheriff O, Ahbara AM, Haile A, Alemayehu K,  
Han J-L and Mwacharo JM (2024), Whole-  
genome resequencing reveals genomic  
variation and dynamics in Ethiopian  
indigenous goats.

*Front. Genet.* 15:1353026.

doi: 10.3389/fgene.2024.1353026

## COPYRIGHT

© 2024 Sheriff, Ahbara, Haile, Alemayehu, Han  
and Mwacharo. 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.

# Whole-genome resequencing reveals genomic variation and dynamics in Ethiopian indigenous goats

Oumer Sheriff<sup>1,2,3\*</sup>, Abulgasim M. Ahbara<sup>4,5</sup>, Aynalem Haile<sup>6</sup>,  
Kefyalew Alemayehu<sup>2,3,7</sup>, Jian-Lin Han<sup>8,9</sup> and  
Joram M. Mwacharo<sup>5,6</sup>

<sup>1</sup>Department of Animal Science, Assosa University, Assosa, Ethiopia, <sup>2</sup>Department of Animal Production and Technology, Bahir Dar University, Bahir Dar, Ethiopia, <sup>3</sup>Biotechnology Research Institute, Bahir Dar University, Bahir Dar, Ethiopia, <sup>4</sup>Department of Zoology, Faculty of Sciences, Misurata University, Misurata, Libya, <sup>5</sup>Animal and Veterinary Sciences Scotland's Rural College (SRUC) and The Centre for Tropical Livestock Genetics and Health (CTLGH), The Roslin Institute Building, Edinburgh, United Kingdom, <sup>6</sup>Resilient Agricultural Livelihood Systems Program (RALSP), International Center for Agricultural Research in the Dry Areas (ICARDA), Addis Ababa, Ethiopia, <sup>7</sup>Ethiopian Agricultural Transformation Institute, Amhara Agricultural Transformation Center, Bahir Dar, Ethiopia, <sup>8</sup>CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China, <sup>9</sup>Livestock Genetics Program, International Livestock Research Institute, Nairobi, Kenya

Ethiopia has about 52 million indigenous goats with marked phenotypic variability, which is the outcome of natural and artificial selection. Here, we obtained whole-genome sequence data of three Ethiopian indigenous goat populations (Arab, Fellata, and Oromo) from northwestern Ethiopia and analyzed their genome-wide genetic diversity, population structure, and signatures of selection. We included genotype data from four other Ethiopian goat populations (Abergelle, Keffa, Gumuz, and Woyto-Guji) and goats from Asia; Europe; and eastern, southern, western, and northern Africa to investigate the genetic predisposition of the three Ethiopian populations and performed comparative genomic analysis. Genetic diversity analysis showed that Fellata goats exhibited the lowest heterozygosity values ( $H_o = 0.288 \pm 0.005$  and  $H_e = 0.334 \pm 0.0001$ ). The highest values were observed in Arab goats ( $H_o = 0.310 \pm 0.010$  and  $H_e = 0.347 \pm 4.35e-05$ ). A higher inbreeding coefficient ( $F_{ROH} = 0.137 \pm 0.016$ ) was recorded for Fellata goats than the  $0.105 \pm 0.030$  recorded for Arab and the  $0.112 \pm 0.034$  recorded for Oromo goats. This indicates that the Fellata goat population should be prioritized in future conservation activities. The three goat populations showed the majority (~63%) of runs of homozygosity in the shorter (100–150 Kb) length category, illustrating ancient inbreeding and/or small founder effects. Population relationship and structure analysis separated the Ethiopian indigenous goats into two distinct genetic clusters lacking phylogeographic structure. Arab, Fellata, Oromo, Abergelle, and Keffa represented one genetic cluster. Gumuz and Woyto-Guji formed a separate cluster and shared a common genetic background with the Kenyan Boran goat. Genome-wide selection signature analysis identified nine strongest regions spanning 163 genes influencing adaptation to arid and semi-arid environments (*HOXC12*, *HOXC13*, *HOXC4*, *HOXC6*, and *HOXC9*, *MAPK8IP2*), immune response (*IL18*, *TYK2*, *ICAM3*, *ADGRG1*, and *ADGRG3*), and production and reproduction (*RARG* and *DNMT1*). Our results provide insights into a thorough understanding of

genetic architecture underlying selection signatures in Ethiopian indigenous goats in a semi-arid tropical environment and deliver valuable information for goat genetic improvement, conservation strategy, genome-wide association study, and marker-assisted breeding.

#### KEYWORDS

**Africa, *Capra hircus*, genome dynamics, pooled heterozygosity, population differentiation, whole genome**

## Introduction

Goats (*Capra hircus*) are of economic, nutritional, and cultural significance to humankind. Their domestication is believed to have happened in the Near East around 11,000 years ago from a mosaic of wild bezoar populations (*Capra aegagrus*) (Zeder, 2008; Daly et al., 2018). Human migrations and trade dispersed goats to diverse environments, and through adaptation, they integrated successfully into these environments.

The detection of genomic regions under positive selection can give insights into phenotypic evolution driven by different breeding objectives or adaptation to local environments (Andersson and Georges, 2004). Over the last decade, many genome-wide selection signatures have been detected in different goat breeds that are associated with production (Wang et al., 2016), litter size (Lai et al., 2016; Guang-Xin et al., 2019; Tao et al., 2021; Wang et al., 2021), adaptation (Kim et al., 2016; Wang et al., 2016), disease resistance (Lee et al., 2016), cashmere fiber (Li et al., 2017) and multiple traits (Guo et al., 2018). However, there are no such studies on the Ethiopian indigenous goats except Berihulay et al. (2019) who however, analyzed only two Ethiopian goat populations, Abergelle and Begait. In general, the studies have demonstrated how positive selection acting on complex traits has changed the genetic composition of domestic goats.

Ethiopia is home to ~52 million goats (CSA, 2021), a great majority of which are indigenous genotypes. They occur in large flock sizes kept by pastoralists in arid and semi-arid environments, while small flock sizes maintained by agro-pastoralists are widely distributed in the highlands (Abegaz et al., 2014). Using morphometric traits, FARM-Africa (1996) classified Ethiopian indigenous goats into 13 populations. However, this classification may have inadvertently classified genetically similar populations as separate entities.

Among the indigenous goat populations of Ethiopia, Arab, Fellata, and Oromo are known to be well-adapted to Ethiopia's semi-arid region of Benishangul-Gumuz (Getinet et al., 2005) and represent the primary livestock species and breeds raised in the area. The region's altitudinal landscape ranges from 550 m to 2,500 m above sea level (Sinmegn et al., 2014), and its rich ancient and recent human socio-economic, political, and cultural history could have impacted the genome landscape of the indigenous goats. For instance, the two main ethnicities in Benishangul-Gumuz (Berta/Arab and Gumuz) were historically closely associated with neighboring areas of Sudan (where they extend) and, to a lesser extent, with the Ethiopian highlands. Various trade routes dating to the Axumite era (100–940 AD) met in Benishangul-Gumuz, where goods, including gold, livestock, iron, coffee, ivory, and honey, were exchanged. The

Axumite kingdom was a trading empire with its hub in Eritrea and northern Ethiopia, which at times extended across most of present-day Eritrea, northern Ethiopia, Western Yemen, and parts of eastern Sudan (Fattovich, 2019). Furthermore, between 1979 and the mid-1980s, there was a relocation of a large population of inhabitants from the Ethiopian highlands (Amhara) to Benishangul-Gumuz due to drought and famine (Mulatu, 1991; Gebru, 2009). The region is also known for its intense solar radiation, feed and water scarcity, and tsetse infestation (Duguma et al., 2015). We, therefore, hypothesize that these historical, contemporary, and environmental events could have impacted the genetic makeup of indigenous goats from Benishangul-Gumuz, which could have contributed to their genetic differentiation, admixture with other Ethiopian indigenous goats, and adaptation to such an environment.

In this study, using whole-genome sequence data, we investigated i) the genetic diversity and population structure of three Ethiopian indigenous goat populations while mapping their genetic profiles to other goats from Africa, Asia, and Europe and ii) signatures of selection associated with adaptation to arid and semi-arid environments, immune response, and production and reproduction. For the latter, we present an assessment of the genome structure between Arab, Fellata, and Oromo goats descended from a semi-arid tropical environment in Ethiopia and that of Tibetan goats from a cool highland temperate environment in China and exposed to contrasting selection pressures, natural *versus* artificial, between the following population pairs: Arab vs. Tibetan, Fellata vs. Tibetan, and Oromo vs. Tibetan.

## Materials and methods

### Animals and whole-genome sequencing

Thirty (30) unrelated animals representing three northwestern Ethiopia indigenous goat populations (Arab, Fellata, and Oromo) were sampled for this study. Whole blood was collected from each individual by puncturing the jugular vein using EDTA-coated vacutainer tubes while adhering to the guidelines on animal welfare and care of the Ministry of Livestock and Fisheries of the Federal Democratic Republic of Ethiopia. Genomic DNA was extracted from whole blood using DNeasy® Blood and Tissue kit (<https://www.QIAGEN.com>, accessed on 11 September 2019) following the manufacturer's protocol with a few modifications. The integrity of the extracted DNA was checked in a 1% agarose gel. The concentration and purity of the DNA were determined by spectrophotometer readings at 260 nm and 280 nm, respectively (DeNovix Inc., Wilmington, DE, USA). Whole-genome sequencing

at a depth of 10X was done with the NovaSeq 6000™ platform at Tianjin Noozhiyuan Technology Co., Ltd.

Genotype data from 60 animals representing ten goat populations were obtained from the VarGoats project (available from the European Nucleotide Archive (ENA), project number PRJEB37507, accessed on 16 August 2021) in FASTQ format and from which we extracted filtered genotype data. We included Ethiopian (Abergelle, Keffa, Gumuz, and Woyto-Guji) and Kenyan (Boran) goat populations to represent East African goats, Malawian (Thylo) to represent South African goats, Malian (Guera) to represent West African goats, Moroccan (Unknown) to represent North African goats, French (Saanen) to represent European goats, and Chinese (Tibetan) to represent Asian goats. A detailed description of the environmental characteristics of the geographic areas of the study populations is given in [Supplementary Table S1](#).

## Read alignment and variant calling

Sequence read quality was evaluated with FastQC v0.11.5. Illumina adapter reads were trimmed with Trimmomatic v0.36 (Bolger et al., 2014) based on the following criteria: Slidingwindow:4:15, Leading:3, Trailing:3, and Minlen:36. The clean reads were mapped to the ARS1 *C. hircus* reference genome assembly (RefSeq number GCF\_001704415.1) (Bickhart et al., 2017) using the Burrows–Wheeler Aligner (BWA-MEM algorithm v0.7.17) (Li, 2013) with default parameters except for `-t 8 -M` and `-R` to add read groups. Each sequence alignment map (SAM) (Li and Durbin, 2009) was sorted and converted to a binary alignment map (BAM) (Li and Durbin, 2009) using Picard (<https://broadinstitute.github.io/picard/>) “SortSam” v2.22.8. The same software was also used to mark duplicate reads with “MarkDuplicates.” Each BAM file was indexed using “BuildBamIndex,” and read groups were added using “AddOrReplaceReadGroups” of Picard v2.22.8. GATK (Genome Analysis Toolkit; McKenna et al., 2010) BaseRecalibrator and ApplyBQSR v4.1.7.0 were used for base quality score recalibration. The “known-sites” file that is necessary for the BQSR step was computed for each individual. GATK HaplotypeCaller v4.1.7.0 (Poplin et al., 2018) was run to call the variants in the genomic variant call format (GVCF) mode using the GVCF parameter on the pre-processed BAM files. GATK’s CombineGVCFs v4.1.7.0 was then used to aggregate all GVCF files per scaffold. GenotypeGVCFs v4.1.7.0 of GATK was used to perform joint genotyping and output multisample raw variant call format (VCF) (Danecek et al., 2011) per chromosome/scaffold.

## Filtering process

The multisample raw VCF of each goat population was filtered by VariantRecalibrator v4.1.7.0 of GATK. Two training resources, one with true sites (“known = false, training = true, truth = true, prior = 15.0”) and the other with non-true sites (“known = true, training = false, truth = false, prior = 2.0”) used for recalibration

were dbSNP variants obtained from Ensembl v105. The variant call annotations (for SNPs and InDels) DepthOfCoverage (DP), QualByDepth (QD), RMSMappingQuality (MQ), MappingQualityRankSumTest (MQRankSum), ReadPosRankSumTest (ReadPosRankSum), FisherStrand (FS), and StrandOddsRatio (SOR) were used for VariantRecalibrator. Based on the SNP tranches ([Supplementary Figure S1](#)), no false-positive variants were observed in the 90 tranche. Therefore, the 99 tranche was included to increase the sensitivity of variant discovery. We considered the highest tranche (99.9–100) as a false positive and excluded it. The remaining variants (SNPs and InDels) were then recalibrated at the truth sensitivity filter level (tranche) of 99 using ApplyVQSR v4.1.7.0 of GATK. In general, the entire filtration process resulted in a high confidence set of 35,161,094 biallelic autosomal SNPs and 3,737,445 InDels. The final set of SNPs was annotated using variant effect predictor (VEP) v104.3.

## Genomic diversity analyses

Genomic diversity for each population was analyzed using various metrics, including the proportion of polymorphic SNPs ( $Pn$ ), nucleotide diversity ( $\pi$ ), and genomic expected ( $He$ ) and observed ( $Ho$ ) heterozygosity. Estimates of  $Pn$ —the fraction of total SNPs that displayed both alleles—were calculated as the proportion of SNPs with minor allele frequency (MAF) greater than 0.01. The  $\pi$  analysis (Nei and Li, 1979) is a method that uses SNPs to calculate the average difference between any two nucleotide sequences in a population. The  $\pi$  values were computed based on the sliding window method within 100-kb windows with 50 kb step size along the autosomes using the VCFtools v0.1.15 (Danecek et al., 2011). The  $Ho$  was calculated as the proportion of total heterozygous SNPs to the total number of sites counted in each genome, based on the ARS1 *C. hircus* reference genome. The observed and expected heterozygosity was calculated using the “--het” option of PLINK v1.9 (Purcell et al., 2007) for each genome and then averaged for each population.

The average pairwise genetic distance (D) between individuals within a population was calculated in PLINK v1.9 (Purcell et al., 2007). The average proportion of alleles shared between two individuals was calculated using the “--genome” command line in PLINK v1.9 as:  $D_{ST} = IBS_2 + (0.5 \times IBS_1) / N$ , where  $IBS_1$  and  $IBS_2$  represent the number of loci that share either one or two alleles that are identical by state (IBS) in pairwise comparisons between individuals, respectively, and N is the number of loci tested. The genetic distance between all pairwise combinations of individuals was calculated as  $D = 1 - D_{ST}$ .

The runs of homozygosity (ROHs) were estimated with PLINK v1.9 by invoking the “--homozyg” option. ROHs are uninterrupted stretches of homozygous genotypes common among individuals within a population (McQuillan et al., 2008; Marras et al., 2015). The degree of ROH variation between populations was characterized based on differences in the length and number of ROH fragments between populations. The following PLINK parameters and thresholds (Purcell et al., 2007) were applied to define an ROH region: (i) minimum number of SNPs in ROH or in sliding window = 50, (ii) minimum length of ROH = 100 kb, (iii)

minimum number of missing SNP in the ROH = 1, (iv) minimum allowed density of SNPs within a run = 1 SNP/100 kb, (v) minimum number of heterozygous SNPs in each ROH = 1, and (vi) maximum gap between consecutive homozygous SNPs = 1 Mb. ROHs enable reliable estimation of the level of inbreeding. We estimated the ROH-derived genomic inbreeding coefficient ( $F_{roh}$ ) following McQuillan et al. (2008) as  $F_{roh} = \sum L_{roh}/L_{auto}$ , where  $L_{roh}$  is the total length of ROH of each individual in the genome, and  $L_{auto}$  is the length of the goat autosomal genome (~2400 Mb) (Kumar et al., 2018). To compare the ROH length between populations, four length categories were allocated: 100–150 Kb, >150–250 Kb, >250–400 Kb, and >400 Kb (classified as ROH<sub>100–150Kb</sub>, ROH<sub>150–250Kb</sub>, ROH<sub>250–400Kb</sub>, and ROH<sub>>400Kb</sub>, respectively). For each ROH length category, we summed all ROH per animal and averaged this per population. In order to investigate the potential of our approach, we focused on two length classes: from 100 kb to 400 kb to investigate ancient events and >400 kb to address recent events.

## Population genetic structure analyses

We used three methods to detect the genetic differentiation and structure among the studied goat populations: a principal component analysis (PCA) to visualize patterns in relationships between individuals using the “--pca” command line in PLINK v1.9 (Purcell et al., 2007); Treemix v.1.13 (Pickrell and Pritchard, 2012) to construct a maximum likelihood (ML) tree and visualize population splits and the directionality of gene flows between populations; and ADMIXTURE v.1.3 (Alexander et al., 2009) to perform the ancestry proportion of individual genomes.

The PCA was performed at three levels: global, East African, and Ethiopian. To summarize the relationship among individuals, we plotted the first two eigenvectors. For this analysis, we filtered out SNPs with minor allele frequency (MAF) < 0.01, leaving 27,728,833 SNPs that were used. The graphical display of PCA results (PC1 and PC2) was visualized with GENESIS (Buchmann and Hazelhurst, 2014).

In the TreeMix analysis, the number of migration events ( $m$ ) varied between one (1) (migration between two populations) and thirteen (13) (migration between all populations). The TreeMix plotting script was used to visualize the trees and the proportion and direction of gene flow events between the 13 goat populations with R. The number of gene flow events that best fit the data was identified using the fraction of the variance in the sample covariance matrix explained by the model covariance matrix (Pickrell and Pritchard, 2012). To evaluate the confidence of the trees' edges and nodes, 1000 bootstrap replicates were performed. Before running migration events, it is important to define the position of the root using prior information about known out-groups (Pickrell and Pritchard, 2012). We used the Tibetan goat population as the root (-root) in our analyses. Furthermore, to test for admixture among breeds and to assess the statistical significance of gene flow events, the THREEPOP and FOURPOP functions (implemented in TreeMix) were run to calculate the  $f_3$  and  $f_4$  statistics, respectively (Ahbara et al., 2019). For the  $f_3$  test (A; B, C), the putative admixture of a target population (A) is tested against two source populations (B, C). A

significantly negative value of the resulting Z-score indicates A is admixed. The  $f_4$  test (A, B; C, D) investigates the tree topology of four populations. A significantly positive Z-score suggests gene flow between populations A and C or B and D that surpasses any gene flow between populations A and D or B and C. A significantly negative Z-score, however, indicates gene flow between populations A and D or B and C that surpasses any gene flow between A and C or B and D. Standard errors were estimated using blocks of 500 SNPs.

For the ADMIXTURE analysis, we used 8,344,942 autosomal SNPs that were retained after we pruned the 27,728,833 SNPs used in PCA for linkage disequilibrium (LD). ADMIXTURE was run for  $K = 2$  to  $K = 13$ . A five-fold cross-validation (CV) procedure was applied to determine the optimal number of clusters (K). GENESIS (Buchmann and Hazelhurst, 2014) was used to display the ADMIXTURE OUPUTS.

## Genomic selection signature analysis

We implemented two methods for the selection signature analysis: i) genetic differentiation based on  $F_{ST}$  (Weir and Cockerham, 1984) and ii) within-group pooled heterozygosity ( $H_p$ ) (Rubin et al., 2010). A sliding window approach was used to perform the  $F_{ST}$  analyses with VCFtools (v.0.1.15) using a 100-kb sliding window size ( $\geq 30$  SNPs) and a 25-kb step size according to the method used previously (Lai et al., 2016; Guo et al., 2018) following the equations:  $F_{ST} = 1 - (p_1q_1 + p_2q_2)/2prqr$ ; where  $p_1$ ,  $p_2$  and  $q_1$ , and  $q_2$  are the frequencies of alleles A and a in the first and second group of the test populations, respectively, and  $pr$  and  $qr$  are the frequencies of alleles A and a, respectively, across the tested groups (Zhi et al., 2018). The  $F_{ST}$  values were standardized into Z-scores as follows:  $Z(F_{ST}) = (F_{ST} - \mu(F_{ST}))/\sigma(F_{ST})$ ; where  $\mu$  is the mean, and  $\sigma$  is the standard deviation of  $F_{ST}$  derived from all the windows tested between test groups. Finally, the top 1% of windows showing both extremely high  $F_{ST}$  values (the top 1% of  $Z(F_{ST})$  distributions) and extremely low  $H_p$  scores (the bottom 1% of  $ZH_p$  distributions) (Guo et al., 2018) were proposed to be under selection.

The  $H_p$  values were also estimated using a window size of 100 kb and a sliding step size of 25 kb (Guo et al., 2018) with the formula:  $H_p = (2 \sum nMaj \sum nMin) / (\sum nMaj + \sum nMin)^2$ ; Where  $\sum nMaj$  and  $\sum nMin$  are the sums of major and minor allele frequencies, respectively, for all the SNPs in the 100-kb window. The values for the  $H_p$  calculated for each window size were then Z-transformed using the equation:  $Z(H_p) = (H_p - \mu(H_p))/\sigma(H_p)$ , where  $\mu$  is the mean and  $\sigma$  is the standard deviation of  $H_p$ . Because sex chromosomes and autosomes are subjected to different selective pressures and have different effective population sizes, we calculate the  $ZH_p$  for autosomes of specific breeds only. Putative genomic regions that overlapped between these two approaches were defined as candidate selective regions.

## Candidate gene analysis

Based on genome annotation, a gene was deemed to show evidence of being under selection if it overlapped with an outlier

TABLE 1 Estimates of genetic diversity parameters for each of the 13 populations analyzed in this study.

Population	N	Pn	Ho	He	F <sub>ROH</sub>	π	D
			Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Arab	10	0.928	0.310 ± 0.010	0.347 ± 4.35e-05	0.105 ± 0.030	0.0020 ± 0.01	0.280 ± 0.033
Fellata	10	0.944	0.288 ± 0.005	0.334 ± 0.0001	0.137 ± 0.016	0.0021 ± 0.001	0.281 ± 0.005
Oromo	10	0.930	0.306 ± 0.012	0.345 ± 0.0002	0.112 ± 0.034	0.0020 ± 0.001	0.281 ± 0.016
Abergelle	6	0.933	0.299 ± 0.074	0.340 ± 0.003	0.121 ± 0.210	0.0019 ± 0.001	0.280 ± 0.004
Keffa	6	0.924	0.254 ± 0.115	0.342 ± 0.004	0.261 ± 0.326	0.0017 ± 0.001	0.264 ± 0.001
Gumuz	6	0.932	0.334 ± 0.025	0.345 ± 8.80e-05	0.032 ± 0.072	0.0022 ± 0.001	0.269 ± 0.005
Woyto-Guji	6	0.939	0.335 ± 0.009	0.339 ± 9.06e-05	0.016 ± 0.027	0.0023 ± 0.001	0.296 ± 0.017
Boran	6	0.941	0.330 ± 0.011	0.337 ± 0.0002	0.020 ± 0.032	0.0022 ± 0.001	0.265 ± 0.002
Unknown	6	0.960	0.304 ± 0.023	0.328 ± 5.15e-05	0.072 ± 0.071	0.0022 ± 0.001	0.272 ± 0.007
Thyolo	6	0.905	0.332 ± 0.048	0.357 ± 7.30e-05	0.068 ± 0.135	0.0019 ± 0.001	0.274 ± 0.009
Guera	6	0.934	0.322 ± 0.029	0.346 ± 0.0007	0.069 ± 0.084	0.0020 ± 0.001	0.278 ± 0.016
Saanen	6	0.915	0.331 ± 0.038	0.369 ± 0.0011	0.174 ± 0.101	0.0018 ± 0.001	0.306 ± 0.012
Tibetan	6	0.932	0.330 ± 0.026	0.334 ± 0.0006	0.117 ± 0.077	0.0021 ± 0.001	0.275 ± 0.025

Pn, polymorphic SNPs; Ho, observed heterozygosity; He, expected heterozygosity; F<sub>ROH</sub>, ROH-based inbreeding; π, nucleotide diversity; D, pairwise genetic distance.

genomic window based on both *ZF<sub>ST</sub>* and *ZHp*. Hence, we annotated the candidate genomic regions using the *Ensembl BioMart* tool (<https://www.ensembl.org/biomart>, accessed on 16 February 2024) based on the Caprine reference genome (*ARSI*). Using *C. hircus* as a background species, we performed functional enrichment analysis of all the annotated genes using Database for Annotation, Visualization, and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/content.jsp?file=release.html>, accessed on 16 February 2024) v6.8 (Huang et al., 2009). Categories with the threshold of adjusted *p*-value <0.05 after the Bonferroni correction were defined as significantly enriched terms and pathways. To infer gene functions, we consulted the NCBI database (<https://www.ncbi.nlm.nih.gov>) and reviewed the literature.

## Results

### Genome sequence mapping and SNP calling

The SNP and InDel summary statistics of the sequence parameters for each population are shown in [Supplementary Tables S2, S3](#), respectively. The highest number of SNPs (18,266,925) and InDels (2,033,758) were detected in the Unknown (UNK) population from Morocco, while the lowest number of SNPs (14,389,837) and InDels (1,739,360) were identified in Saanen (SAN; Europe) breed and Gumuz (GUM; Ethiopia) breed, respectively. The average number of SNPs per sample ranged from 6,591,579 (Keffa (KEF); Ethiopian) to 8,866,137 (Woyto-Guji (WGU); Ethiopian). The average number of InDels ranged from 660,074 (Keffa; KEF) to 976,619 (Abergelle (ABR); Ethiopian). The heterozygosity to homozygosity (het/hom) ratio ranged from 1.03 (Keffa (KEF)) to 1.62 (Unknown (UNK)). A comparison with the *C. hircus* dbSNP database revealed that 29%–

35% of the SNPs and 78%–82% of the InDels were not present in the database.

### Genomic diversity

Genomic diversity metrics (mean ± standard deviation (SD); [Table 1](#)) for each population were assessed by estimating the proportion of polymorphic SNPs (Pn), observed (Ho) and expected (He) heterozygosity, ROH-based inbreeding (*F<sub>roh</sub>*); nucleotide diversity (π), and average pairwise genetic distance between individuals within populations/breeds (D). The Pn ranged between 90.5% and 96%. The highest Pn was in the unknown breed (Morocco) and the lowest in Thylo (Malawi). Keffa goats had the lowest Ho (0.254 ± 0.115), π (0.0017 ± 0.001) and D<sub>ST</sub> (0.264 ± 0.001) but the highest *F<sub>roh</sub>* (0.261 ± 0.326). On the other hand, Woyto-Guji had the highest Ho (0.335 ± 0.009) and π (0.0023 ± 0.001) but the lowest *F<sub>roh</sub>* (0.016 ± 0.027). The value of He was always greater than that of Ho in each breed.

### Runs of homozygosity (ROHs)

The frequency of ROHs was calculated from homozygous sequences across each genome for the four genome length categories (ROH<sub>100–150Kb</sub>, ROH<sub>150–250Kb</sub>, ROH<sub>250–400Kb</sub>, and ROH<sub>>400Kb</sub>) analyzed here ([Figure 1](#)). A total of 57,910 ROH segments were identified across the 78 individuals. Thylo had the highest number (10,979) of ROH segments, while Saanen had the lowest number (909) of ROH segments ([Table 2](#)). The ROHs comprised mostly (58%) the shorter (ROH<sub>100–150Kb</sub>) length category, while a very small proportion (~4%) of the ROH length category was detected in the longer ROH segment (ROH<sub>>400Kb</sub>).

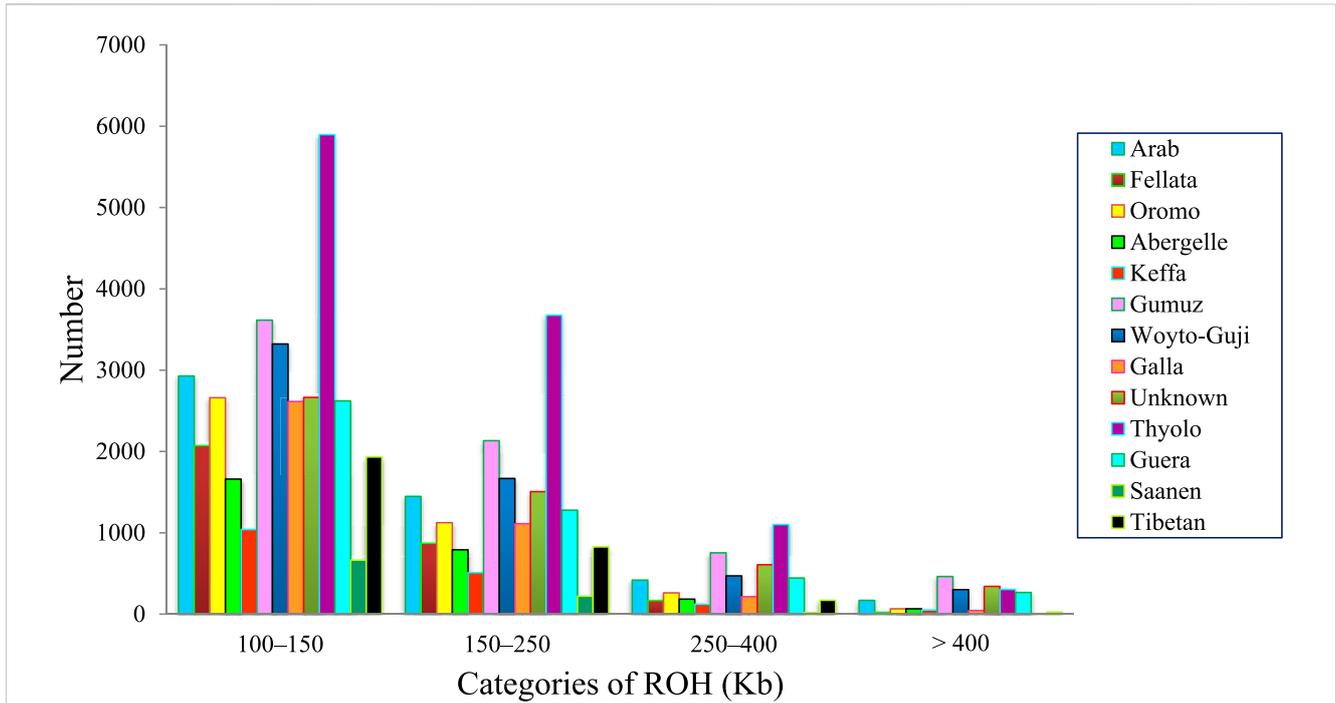


FIGURE 1 Number of ROHs for the four genome length categories (ROH<sub>100–150Kb</sub>, ROH<sub>150–250Kb</sub>, ROH<sub>250–400Kb</sub>, and ROH<sub>>400Kb</sub>) of the study goat populations.

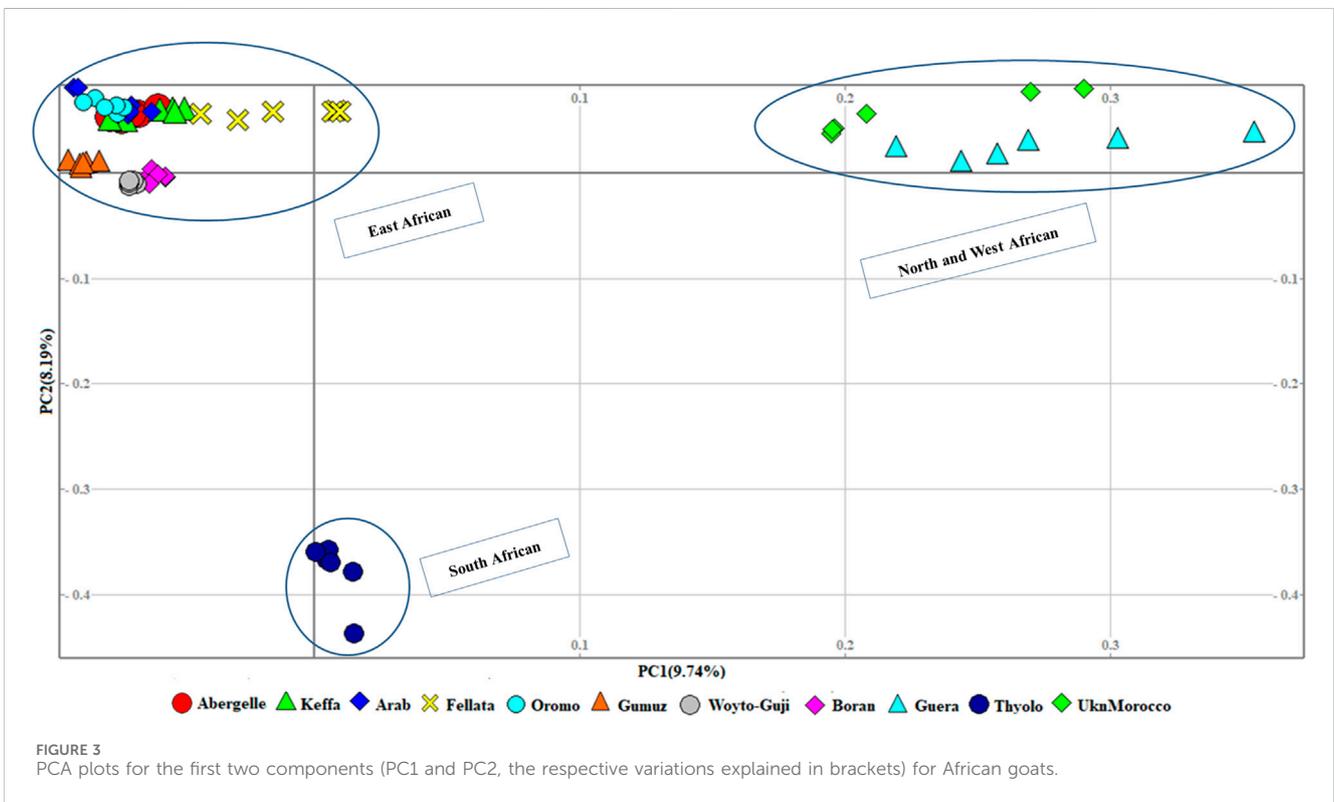
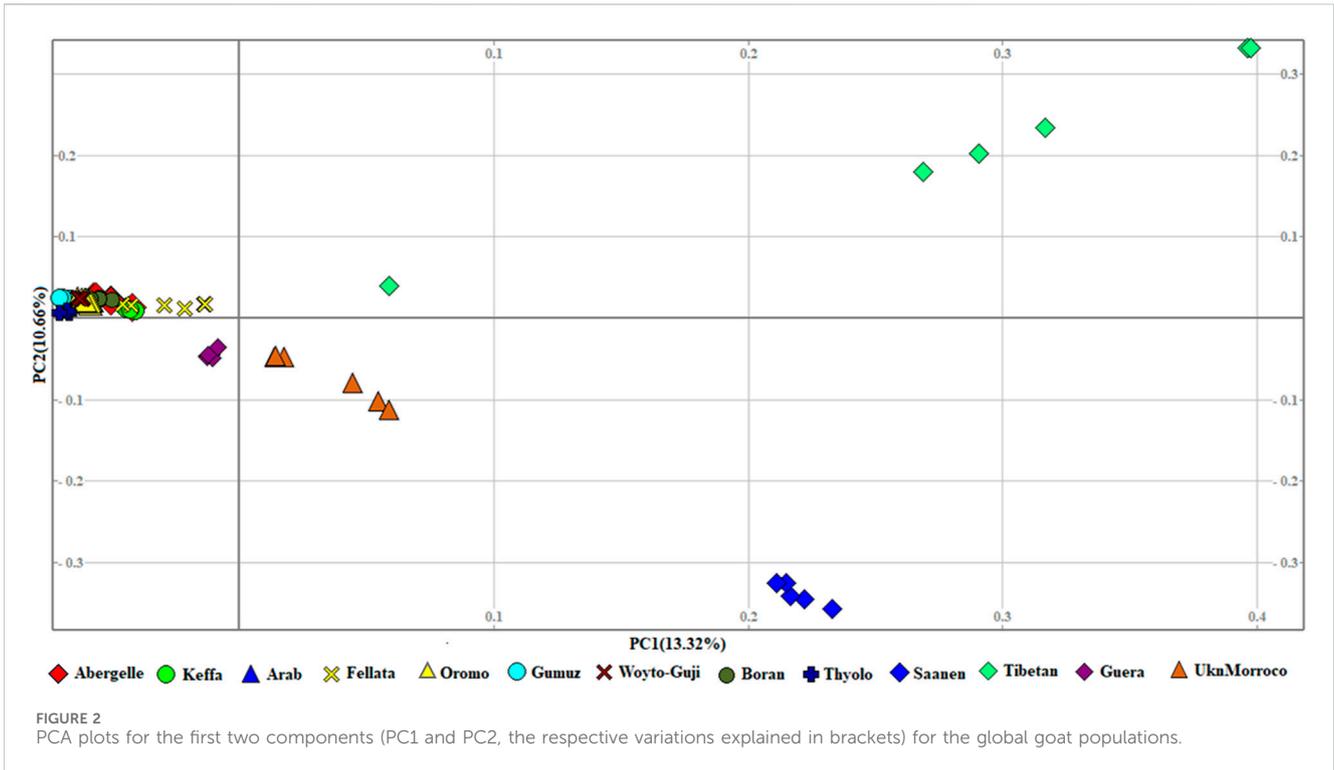
TABLE 2 Statistics of ROH observed in the study goat populations under different length classes.

Population	Total number of ROH	ROH number in different ROH length categories			
		ROH <sub>100–150Kb</sub>	ROH <sub>150–250Kb</sub>	ROH <sub>250–400Kb</sub>	ROH <sub>&gt;400Kb</sub>
Arab	4964	2929	1448	418	169
Fellata	3133	2071	871	168	23
Oromo	4113	2661	1124	262	66
Abergelle	2701	1660	791	184	66
Keffa	1721	1040	510	121	50
Gumuz	6964	3615	2133	754	462
Woyto-Guji	5761	3322	1668	470	301
Boran	3984	2614	1112	214	44
Unknown	5119	2665	1506	607	341
Thyolo	10979	5897	3676	1102	304
Guera	4608	2621	1278	443	266
Saanen	909	664	220	20	5
Tibetan	2954	1932	827	170	25
Overall	57910	33691 (58.18%)	17164 (29.64%)	4933 (8.52%)	2122 (3.66%)

### Population structure

The genetic relationship between individuals was analyzed with PCA, which we performed at three levels: global, East African, and Ethiopian. At the global level, the PCA separated the 13 populations

into four groups that corresponded to their geographic origin, viz, 1) Asian, 2) European, 3) South and East African, and 4) North and West African. In this PCA, PC1 and PC2 explained 13.32% and 10.66%, respectively, of the variation in the entire genetic data (Figure 2).



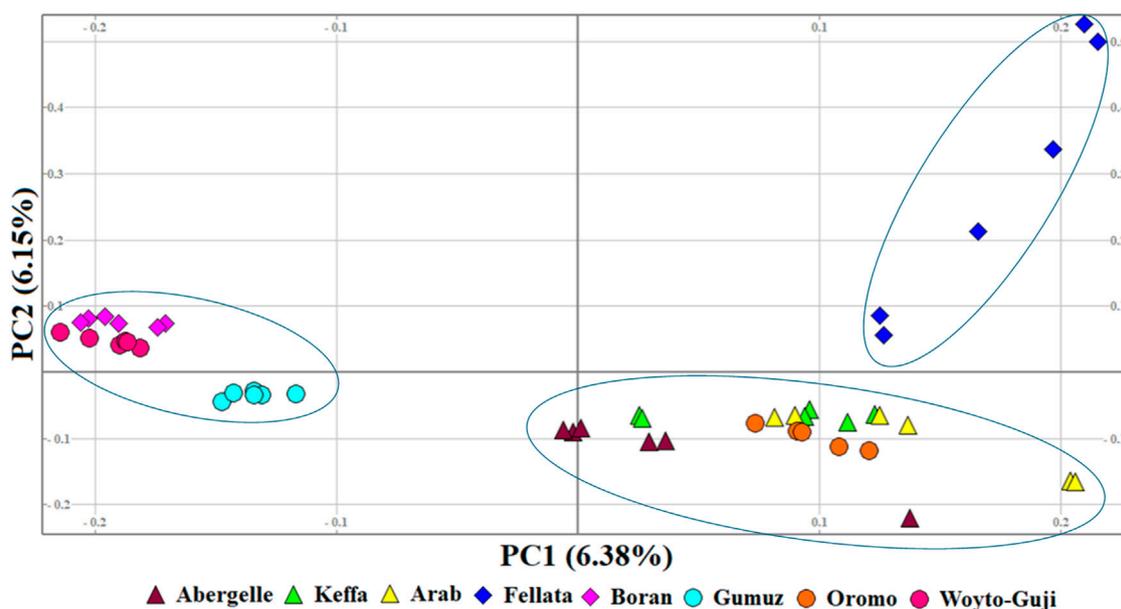


FIGURE 4  
PCA plots for the first two components (PC1 and PC2, the respective variations explained in brackets) for the East African goats.

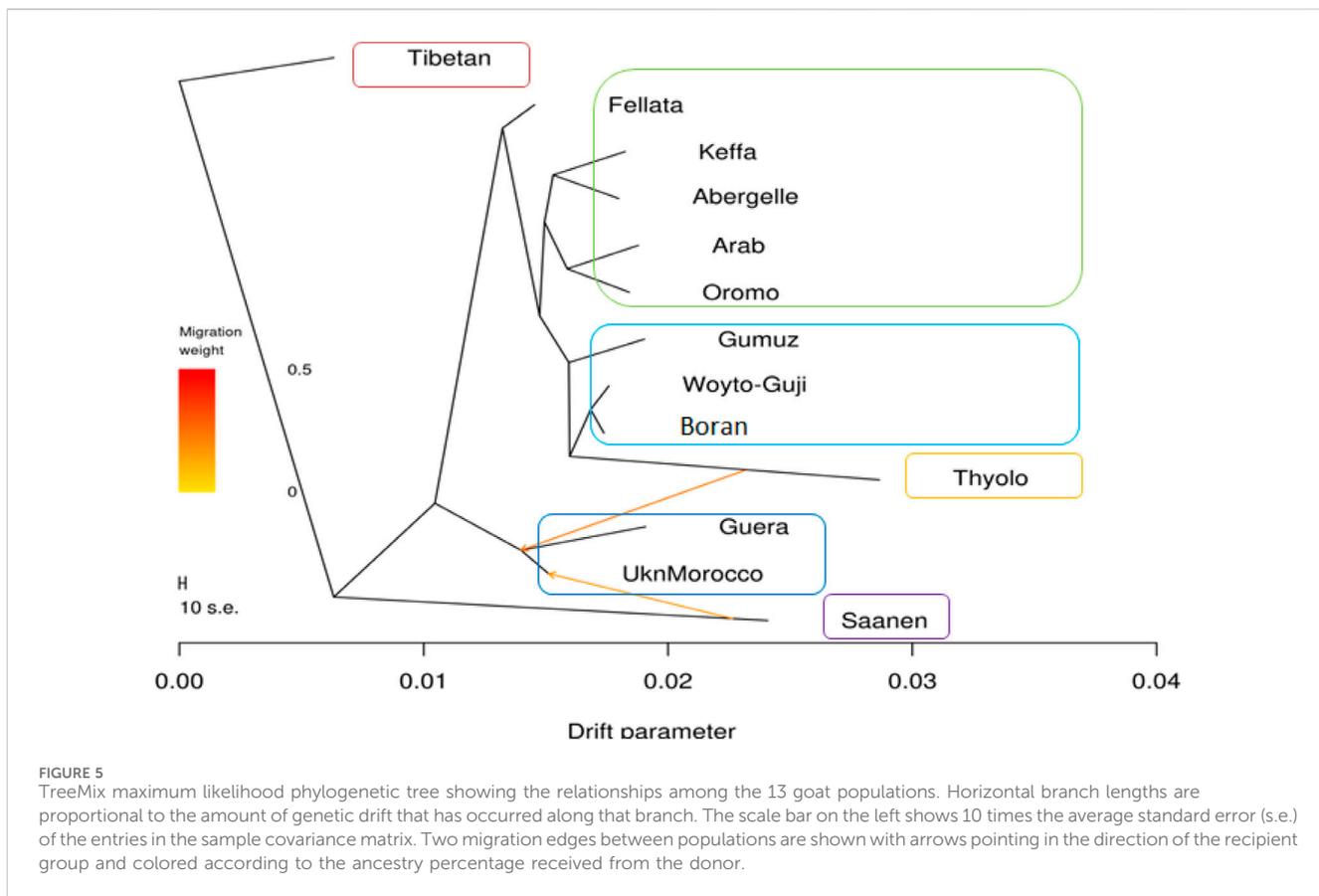
At the second level, which aimed to obtain a clearer picture of the genetic variation among 11 African populations (Figure 3), the PCA revealed three genetic clusters, viz, 1) East African, 2) North and West African, and 3) South African. In contrast to the global PCA, there is the separation of East African goats from the South African. In this dataset, PC1 and PC2 accounted for 17.93% of the total variation.

To illustrate the genetic variation among East African goats, we performed the PCA with only the eight East African goat populations (Figure 4). It revealed three genetic clusters: Cluster 1 (Boran, Gumuz, and Woyto-Guji), Cluster 2 (Abergelle, Arab, Keffa, and Oromo), and Cluster 3 (Fellata). In this dataset, PC1 and PC2 explained 12.53% of the total variation.

The phylogenetic tree generated with TreeMix (Figure 5) revealed four genetic groups in the 13 goat populations. The groups are similar to the genotype groups seen in the PCA plot (Figure 2). As expected, it clearly distinguished Saanen and Tibetan goats. However, it clustered the South and East African goats in one group, separate from the North and West African goats, which were in another group. It also revealed two gene flow events, one from Saanen to Unknown and the other from Thyolo to Guera. Some of the populations, including Saanen, Thyolo, and Tibetan, are shown to have long branches corresponding to the amount of genetic drift.

The  $f_3$  test failed to reveal any gene flow events between the populations analyzed (Supplementary Table S4). A complex pattern of gene flow between the study populations, which cannot be explained by a three-way model, may explain this result. The  $f_4$  test highlighted possibilities of gene flow among various populations. The highest Z-scores ( $>|54|$ ) were between Boran and Abergelle and between Unknown and Saanen (Supplementary Table S5). The lowest Z-scores ( $>|54|$ ) were between Boran and Woyto-Guji and between Saanen and Abergelle.

The graphical summary of the results of ADMIXTURE for  $2 \leq K \leq 13$  is presented in Figure 6. The cross-validation (CV) error was lowest at  $K = 6$  (Supplementary Figure S2), suggesting the presence of six optimal genetic backgrounds in the 13 study goat populations. At this  $K$ -value, Saanen and Tibetan each show separate genetic backgrounds, suggesting they are genetically distinct. This separation is also revealed by the PCA (Figure 2) and Treemix (Figure 5). The Guera and Unknown breeds share the third genetic background, but the Unknown breed shows a small proportion of the Saanen background in its genome, which corresponds with the gene flow results from Saanen to Unknown in the TreeMix (Figure 5). Thyolo has the fourth genetic background, while Gumuz, Woyto-Guji, and Boran share the fifth genetic background. The sixth genetic background occurs in Abergelle, Keffa, Fellata, Arab, and Oromo. Some results are noteworthy. At  $K = 5$ , Thyolo, Guerra, and Unknown share a common genetic background, but Thyolo is separate with a different background at  $K = 6-13$ . The TreeMix also separates Thyolo and shows that it is closer to the populations from Ethiopia and Kenya (Figure 5). Similarly, the PCA completely separates Thyolo, which corresponds with the ADMIXTURE results for  $K = 6-13$  (Figure 3). At  $K = 6$ , the background that defines Gumuz, Woyto-Guji, and Boran is observed in Abergelle, Arab, and Keffa, while Fellata shows the presence of the Tibetan background and the background defining Guera and Unknown in its genome. At  $K = 7$ , Gumuz diverges from Boran and Woyto-Guji, but its genetic background is now observed in Abergelle, Arab, Keffa, and Oromo. At this  $K$ -value, Fellata shares a background with Abergelle and Oromo goats but is separate from Keffa and Arab (Figure 6). This separation is also revealed by both PCA (Figure 4) and TreeMix (Figure 5). Abergelle, on the other hand, clusters with Arab, Keffa, and Oromo in both PCA and TreeMix, but ADMIXTURE shows Abergelle with a



different genetic background at  $K = 7-13$ , which it shares with Fellata. Arab, which clusters with Abergelle, Keffa, and Oromo in PCA (Figure 4) and TreeMix (Figure 5), shares some genetic background with these populations at  $K = 7-13$  in ADMIXTURE. Gumuz clusters with Woyto-Guji and Boran in PCA (Figures 3, 4) but separates from Woyto-Guji and Boran in TreeMix (Figure 5). This separation is also revealed by ADMIXTURE at  $K = 5, 6$ , and  $8$  but not at  $K = 7$  and  $K = 9-13$ .

At  $K = 6$ , six genetic clusters were observed. These are designated as Cluster 1, which comprised Arab, Fellata, Oromo, Abergelle, and Keffa; Cluster 2, which included Gumuz, Woyto-Guji, and Boran; Cluster 3 consisted of Thyolo; Cluster 4 embraced Unknown and Guera; and Tibetan and Saanen comprised their own individual clusters.

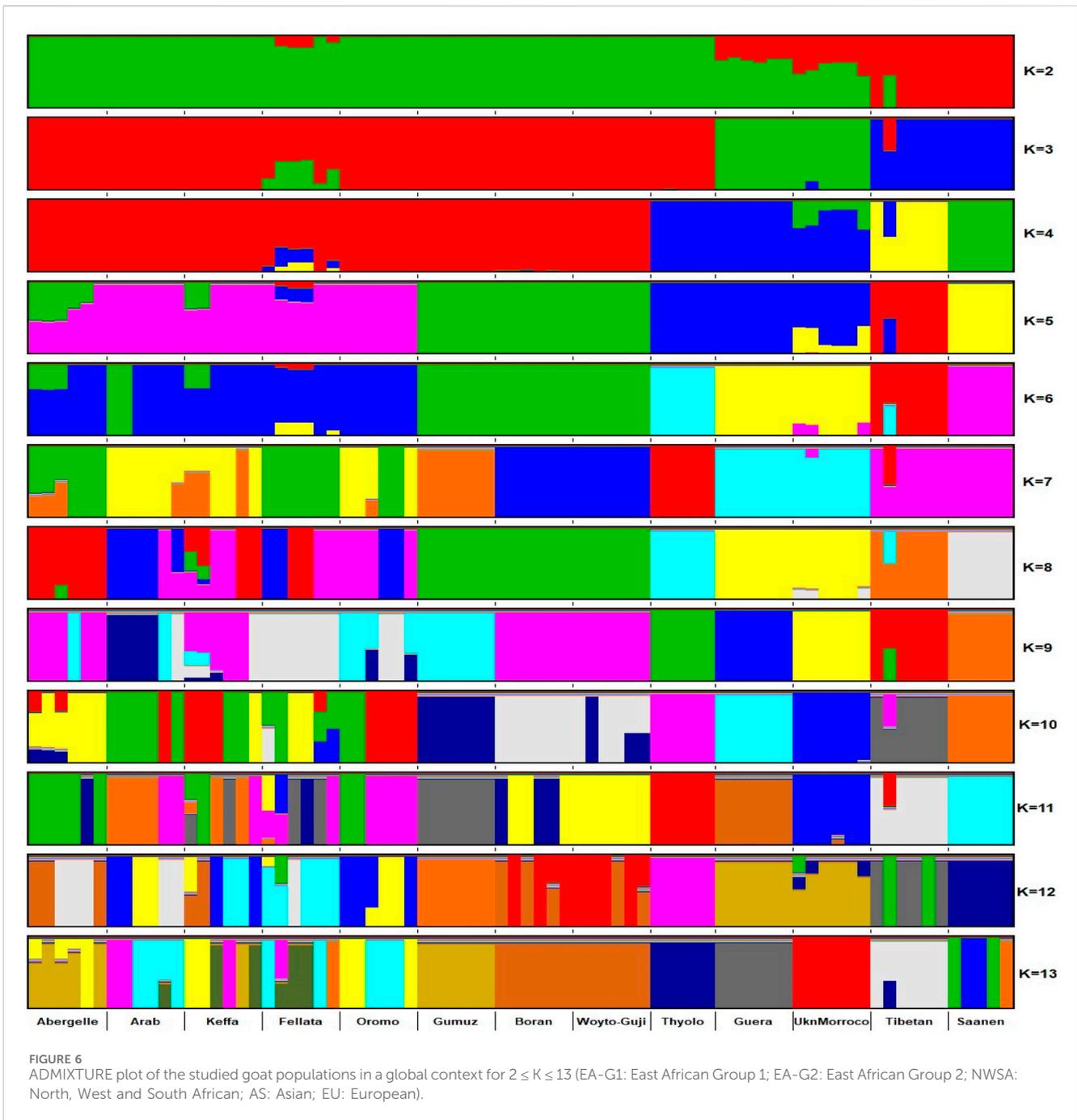
(as described earlier) in each population,  $\sim 987$  windows were assessed per population. Accordingly, genomic regions with high  $ZF_{ST}$  ( $ZF_{ST} \geq 3.20$  (range = 3.20–8.22), 3.29 (range = 3.29–7.96), and 3.24 (range = 3.24–7.34); corresponding  $F_{ST} \geq 0.40, 0.35, 0.40$ ) in the Arab vs. Tibetan, Fellata vs. Tibetan and Oromo vs. Tibetan, respectively, were defined as selection signatures (Figure 7). Based on these criteria, 987 genomic regions were detected for each of the Arab, Fellata, and Oromo populations (Supplementary Tables S6–S8). Among the three populations, there was variation in the distribution of the regions across the genome. In the Arab and Oromo populations, chromosome 2 showed a strong selection signal, whereas in the Fellata population, chromosomes 8 and 18 showed stronger sweeps than the other chromosomes (Figure 7).

### Selection signatures— $F_{ST}$ approach

A total of 98,660 chromosome windows were assessed for each pairwise comparison (Arab vs. Tibetan, Fellata vs. Tibetan, and Oromo vs. Tibetan). The genome-wide distributions of the standardized  $F_{ST}$  values across the genomes for each goat population are shown in Supplementary Figure S3. The average  $F_{ST}$  values were 0.14, 0.11, and 0.13 in Arab, Fellata, and Oromo goats, respectively, suggesting moderate differentiation among the three populations, and this provides insights into their common genetic backgrounds and possible gene flow. Using the top 1% outlier windows as a cutoff threshold

### Selection signatures— $H_p$ approach

We also calculated the Z-transformed pooled heterozygosity ( $ZH_p$ ) in 100-kb sliding windows and a 25-kb step size to detect selection sweeps in the three Ethiopian indigenous goat populations. Accordingly, 98,573 windows were detected per population. The overall average  $H_p$  values across all the windows were 0.19, 0.19, and 0.18 in the Arab, Fellata, and Oromo goats, respectively (Supplementary Figure S4). Similarly, only outliers falling within the bottom 1% ( $\sim 986$  windows) with low  $ZH_p$  values ( $ZH_p \leq -2.73, -2.59, \text{ and } -2.72$ ; corresponding  $H_p \leq 0.08, 0.09, \text{ and } 0.08$  in the Arab, Fellata, and Oromo goats, respectively, were considered signatures of selection.

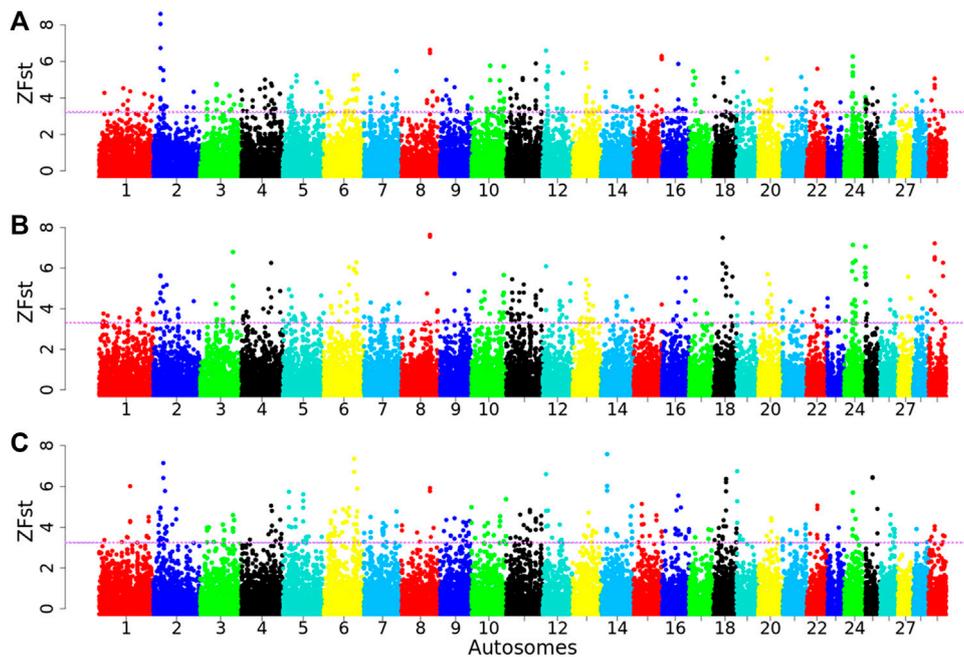


Accordingly, ~983 genomic regions were detected in each population (Supplementary Tables S9–S11). Figure 8 displays the Manhattan plot of the  $ZH_p$  values for several comparisons in each population. In both the Arab and the Fellata goats, chromosome 15 was shown to have the highest signs for selection signatures, while chromosome 14 was detected in the Oromo population.

### Overlapping selection signature regions, genes identified, and enrichment analysis

Based on the overlap of the top 1%  $ZF_{ST}$  and the bottom 1%  $ZH_p$ , 250, 174, and 285 putative selection signature regions were

detected for Arab, Fellata, and Oromo goats, respectively (Supplementary Tables S12–S14). Supplementary Tables S15–S17 show a set of functional genes that are novel and reported after annotation of those overlapping genomic regions for Arab, Fellata, and Oromo goat populations, respectively. A total of 206 genes within 65 putative selection signature regions were detected to be associated with Arab goats (Supplementary Table S15). Similarly, 107 genes within 42 candidate genomic regions were identified in Fellata goats (Supplementary Table S16), and 195 genes within 63 genomic regions were identified in Oromo goats (Supplementary Table S17). In total, 508 genes were detected within 170 regions for the three goat populations (Supplementary Tables S15–S17). Among the 170 candidate regions, nine regions



**FIGURE 7** Manhattan plots for selection sweep analysis (A) between Arab vs. Tibetan, (B) between Fellata vs. Tibetan, and (C) between Oromo vs. Tibetan goat populations performed using the standardized population differentiation (ZFST) approach. The horizontal line represents the arbitrary threshold for ZFst.

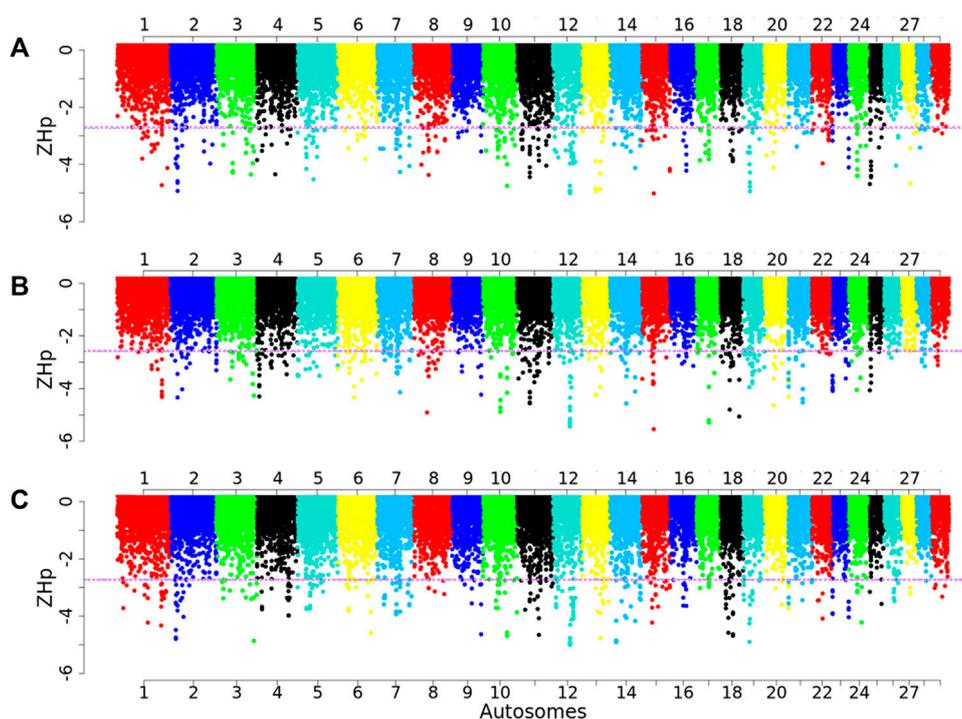
were the strongest, including CHI5, CHI7, and CHI19 in Arab goats (Supplementary Table S15); CHI5, CHI7, and CHI18 in Fellata goats (Supplementary Table S16); and CHI5, CHI15, and CHI19 in Oromo goats (Supplementary Table S17). These nine regions spanned a total of 163 candidate genes (Supplementary Tables S15–S17). Enrichment analysis was performed with the 163 genes using DAVID on the goat gene set (*C. hircus*) with default settings. Based on the annotation, of the 163 genes, 125 were significantly ( $p \leq 0.05$ , Bonferroni correction) enriched, and we found nine biological process (BP) terms, five cellular component (CC) terms, six molecular function (MF) terms, and one KEGG pathway (Supplementary Tables S18–S21).

## Discussion

In the present study, we investigated the genome architecture of three northwestern populations of Ethiopian indigenous goats and benchmarked them against other goat populations from eastern, western, northern, and southern Africa, as well as two exotic breeds, Saanen and Tibetan, from Europe and Asia, respectively. The overall average  $H_o$  and  $H_e$  exceeded 0.322, and the within-population genetic variation (D) was above 0.278, suggesting the study populations are highly genetically diverse. The values for  $H_o$  and  $H_e$  observed in this study are within the range of global goat diversity (Colli et al., 2018) and close to those reported in Sudanese (Rahmatalla et al., 2017) and Pakistani (Kumar et al., 2018) goats and in the Egyptian Barki goat breed (Kim et al., 2016). The higher variability observed within the Ethiopian goat

populations could be attributed to uncontrolled mating that occurs as a result of the common practice of flocks utilizing communal grazing and watering points. The absence of artificial selection, high levels of admixture in these populations, and the introduction and crossbreeding of exotic goats into Ethiopia can also be possible explanations for the increased heterozygosity. Artificial selection within a population may benefit from high variation within populations, especially in areas where community-based goat breeding programs (CBBP) are practiced. CBBP, in which artificial selection occurs within a population, provides a good framework for the implementation of genomic selection in smallholder production systems (Mrode et al., 2018; Rezik et al., 2021).

Runs of homozygosity (ROH) are two contiguous identical by descent (IBD) genomic segments (Gibson et al., 2006) that arise from an increased level of relatedness between individuals within a population or through positive selection (Kijas, 2013). Estimates of ROHs can be used to assist with the interpretation of genomic inbreeding and give insights into population history (Purfield et al., 2012; Zavarez et al., 2015). The latter is particularly important for African indigenous livestock, which are characterized by a lack of written pedigree data (Kosgey et al., 2006). According to Purfield et al. (2012), short ROHs are most likely correlated to ancient inbreeding or potential ancient bottlenecks, whereas long ROHs are more likely associated with relatively recent inbreeding. In the present study, the Ethiopian indigenous goat populations showed their majority of ROHs in the short (100–150 Kb) length category, which is in agreement with the results obtained for other goats (Kim et al., 2016; Brito et al., 2017; Onzima et al., 2018; Islam et al., 2019).



**FIGURE 8** Manhattan plots performed using the standardized pool heterozygosity (*ZHp*) approach for each 100-kb sliding window with a 25-kb step size across all autosomes in the (A) Arab, (B) Fellata, and (C) Oromo goat populations. The horizontal line represents the arbitrary threshold for *ZHp*.

The accumulation of ROHs in the short-length category indicates that the study goats could have been initially established by small founding populations but were not particularly affected by recent inbreeding.

A combination of PCA, ADMIXTURE, and TreeMix tools provided an insight into the genetic structure of the three indigenous goat populations from northwestern Ethiopia and referenced other Ethiopian indigenous goat populations as well as goat populations from Africa and one each from Europe and Asia. The ADMIXTURE tool revealed two distinct genetic clusters. At  $K = 5$ , Gumuz and Woyto-Guji were grouped together, while Abergelle, Arab, Fellata, Keffa, and Oromo were in a different cluster and showed some degree of admixture. This result was supported by the TreeMix result. Our finding on these two genetic groups mirrors the previous findings from mitochondrial DNA analyses of 13 Ethiopian goat populations, which identified two haplogroups (A and G) (Tarekegn et al., 2018), suggesting the presence of two deep ancient ancestries in Ethiopia. However, based on the current dataset, it is difficult to infer whether the two groups introduced/arrived in Ethiopia together or independently. The Gumuz and Woyto-Guji populations are geographically isolated (e.g., Gumuz is located in northwestern while Woyto-Guji is located in southern Ethiopia), but they clustered together and were separate from other Ethiopian goat populations. This could be due to their unique genetic compositions and similarity of production environments. We found that the heterozygosity values for Gumuz and Woyto-Guji were very similar ( $H_o = 0.334 \pm$

$0.025$  and  $0.335 \pm 0.009$  for Gumuz and Woyto-Guji, respectively). This was further confirmed by comparable inbreeding coefficient and number of ROHs. The goats also inhabit similar agro-ecology (semi-arid), and both of them are kept by agro-pastoralists. The clustering of Arab, Fellata, and Oromo goats in one group is attributed to geographical proximity. Home tracts of the three goat populations tend to overlap, which may facilitate ease of flock exchange between farmers and favor gene flow among the populations. In general, the study goat populations are characterized by a high level of admixture and a lack of phylogeographic structure, and this agrees with Tarekegn et al. (2018), Luikart et al. (2001), and Naderi et al. (2007; 2008), who reported similar findings in different goat breeds.

We did find a total of 170 overlapping genomic regions in the three populations, spanning 508 candidate genes, by combining the two approaches. Comparable results have been reported for livestock species from similar environments (Kim et al., 2016; Mwacharo et al., 2017; Ayalew et al., 2023). Nine of 170 regions were identified as the strongest signals and covered 163 functional genes. We speculate that these 163 genes represent past and/or on-going selection in the studied goat populations. Many of the 163 candidate genes were associated with diverse physiological, molecular, and cellular processes and pathways (Supplementary Tables S18–S21). This shows that adaptation to semi-arid tropical environmental stressors such as heat, solar radiation, physical exhaustion, feed and water scarcity, parasites, and others is complex and may involve many genomic regions and genes with pleiotropic activities. It also demonstrates the intricacy of adaptation,

involving numerous biological processes and quantitative trait loci, each of which has a small but cumulative effect on the overall phenotypic expression. Among the significant GO terms, the top three that are related to tropical environment adaptations include anterior/posterior pattern specification (GO:0009952), positive regulation of natural killer cell proliferation (GO:0032819), and positive regulation of natural killer (NK) T cell proliferation (GO:0051142) (Supplementary Table S18). Based on their biological functions and information from published studies, several genes that have been reported before and are possibly responsible for the important traits in goats and other domestic livestock were presented.

## Thermo-tolerance genes

Numerous inherent genetic endowments in the Ethiopian indigenous goat populations could be harnessed for better adaptation in semi-arid tropical environments. We identified some of these genetic biomarkers, notably the *HOXC*-cluster (homeobox genes) (*HOXC12*, *HOXC13*, *HOXC4*, *HOXC6*, and *HOXC9*) and *MAPK8IP2* (Supplementary Table S18), that are associated with heat stress. The *HOXC*-cluster was present in Arab and Oromo goat populations and was also pinpointed by GO analysis (GO:0009952), revealing the term anterior/posterior pattern specification. Onzima et al. (2018) reported *HOXC12* and *HOXC13* genes for their involvement in anterior/posterior pattern specification in the Sebei goat breed of Uganda. The genes can also regulate essential traits like keratin and hair follicle differentiation in goats (Wu et al., 2009), sheep (Sander and Powell, 2004), and cattle (Taye et al., 2017; Ayalew et al., 2023). More specifically, the *HOXC13* gene has been reported to influence skin thickness and number of hair follicles in animals. Skin, being the intermediate between the animal body and the surrounding environment, influences thermoregulation positively (Alfonzo et al., 2016) and plays a major role in aiding the adaptation of animals to heat stress. Cattle with relatively thicker skin, such as the thermo-tolerant *Bos indicus*, exhibit better thermoregulation than cattle with thinner skin, like the heat-sensitive *Bos taurus* (Alfonzo et al., 2016). The present study identifies *HOXC13* as one of the biomarkers for thermo-tolerance, which could further aid in long-term breeding goals towards developing agro-ecological zone-specific goat breeds in the study area. The other gene, *MAPK8IP2*, is identified in the candidate genomic region of the Fellata goat population, and it has been found to be involved in different aspects of thermo-tolerance. This gene has been shown to interact with *MAPK8IP1* (Yasuda et al., 2000). The latter is involved in regulating the reaction of cells to heat stress, which raises the transcription activity of multiple heat stress-responsive genes that regulate various processes, such as cell survival, proliferation, and apoptosis in Holstein cows involved in milk production under heat stress (Sigdel et al., 2019). When cells are subjected to heat stress, they release reactive oxygen species (ROS), which result in cellular necrosis and, ultimately, cell death. Interestingly, gene *MAPK8IP1* is involved in suppressing heat stress-induced ROS production and cellular apoptosis (Li et al., 2018).

## Immune response genes

Indigenous goats are known to be well-adapted to various environmental stressors, including disease. A member of interleukins

(IL18) and TYK2 were present in one of the candidate regions identified in Arab and Oromo goat populations. Interleukins are expressed by leukocytes (Maggio et al., 2006) and are likely involved in the activation of the immune response of goats (Amiri et al., 2023) and chickens (Susta et al., 2015; Chen et al., 2016; Truong et al., 2016). In a study conducted on Malabari goats during heat stress, IL18 was identified as a reliable immunological marker that aids in assessing heat stress-mediated immune response alterations (Rashamol et al., 2019). Similarly, Madhusoodan et al. (2019) reported a significant reduction in hepatic IL18 gene expression during heat stress in Salem black goats (Madhusoodan et al., 2019). This gene is considered an important inflammatory marker for quantifying the impact of heat stress on the immune system in goats (Bagath et al., 2019). Similarly, TYK2 plays a vital role in the innate and acquired responses of humans (Lu et al., 2015), including inflammatory conditions resulting from viral infections and autoimmune diseases. Some other important genes identified in the present study, known to be associated with immune response, are members of *ICAM*, such as *ICAM3* in Arab and Oromo goat populations and subfamily members of *ADGRG*, including *ADGRG1* and *ADGRG3* in Fellata goats. *ICAM3*, the gene constitutively expressed on the surface of leukocytes (Xiao et al., 2013), is important in generating an immune cell response (Montoya et al., 2002) through its facilitation of interactions between T cells and dendritic cells (Svajger et al., 2010). The other gene, *ADGRG1*, plays a role in immune regulation (Lala and Hall, 2022) and is expressed in cytotoxic lymphocytes, including natural killer (NK) cells (Della Chiesa et al., 2010). *ADGRG3* has been shown to exert robust effects on B cell development (Wang et al., 2013). Genetic deletion of *ADGRG3* reduces macrophage migration into white adipose tissue (Shi et al., 2016). The fact that *ADGRG3* is robustly expressed in multiple immune cells is fascinating with regard to the recent revelation that *ADGRG3* can be activated by glucocorticoids (Ping et al., 2021), as glucocorticoids are known to exert powerful effects on the physiology of many different cell types in the immune system.

## Genes associated with production and reproduction

We detected genes like *RARG* and *DNMT1* in Arab and Oromo goats, which are associated with production and reproduction. *RARG* plays an important role in milk production, body size, and kidding in Liaoning cashmere goats (Chen et al., 2022). The expression of this gene has been suggested to be required for the development of limb buds and skeletal growth in dairy cows (Bionaz et al., 2015). Furthermore, *RARG* was identified as the most important master regulator of quantitative trait loci (QTL) for milk production in the F<sub>2</sub> population of German Holstein × Charolais crossed cows (Brand et al., 2016). This gene was also associated with litter size in pigs (Rothschild et al., 1997) and the development of bovine embryos (Mohan et al., 2001). In other species, such as mice, *RARG* was reported to be of functional significance involved in spermatogenesis (Aurore et al., 2012), formation of normal alveoli and alveoli elastic fibers in the lung (McGowan et al., 2000), hematopoietic development (Purton, 2007), and correct formation of the axial skeleton, including anteriorization of the cervical and thoracic vertebrae (Wendling et al., 2001). In contrast, its absence in mice resulted in growth deficiency (Lohnes et al., 1993), bone mass reduction (Green et al., 2015), reduced chondrocyte proliferation, decreased

expression and deposition of proteoglycans (Williams et al., 2009), and male sterility (Lohnes et al., 1993). The other gene, *DNMT1*, consistently appears to be involved in reproduction (Kay et al., 2018; Schulz et al., 2018; Bewick et al., 2019; Amukamara et al., 2020; Washington et al., 2021) and egg production and egg viability in *B. tabaci*. *DNMT1* knockdown affected testis size and structure (Washington et al., 2021).

## Conclusion

We used whole-genome sequence data to investigate the level of genetic diversity, population structure, and signatures of selection in three Ethiopian indigenous goat populations. We observed high within-breed genetic diversity, low genetic differentiation, a high level of admixture, and lack of phylogeographic structure in the studied goat populations. The genomic data also identified several potential candidate genes possibly under selection, including adaptation to a hot environment (homeobox genes and *MAPK8IP2*), immune response (*IL18*, *TYK2*, *ICAM3*, and *ADGRG* subfamilies), and production and reproduction (*RARG* and *DNMT1*). Our study is important in the design of goat genetic improvement programs in view of the current and future predicted effects of climate change. It also provides a foundation for future studies to investigate the genome architectures of different ruminant species coexisting in a similar environment.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI SRA database (accession number PRJNA1063878).

## Ethics statement

The animal studies were approved by the Assosa University Research Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was not obtained from the owners for the participation of their animals in this study because before we collected blood samples from each animal, we informed the farmers about the objectives of the study to make informed decisions. During the blood sample collection, animal welfare and health regulations were observed. No further specific permissions were required from the Ethics Committee of the Bahir Dar or Assosa Universities, Ethiopia, at the time of the sampling.

## Author contributions

OS: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualization, writing—original draft, and writing—review and editing. AA: formal analysis, methodology, resources, software, and writing—review and editing. AH: conceptualization, funding

acquisition, methodology, project administration, resources, supervision, validation, and writing—review and editing. KA: conceptualization, funding acquisition, project administration, resources, supervision, visualization, and writing—review and editing. J-LH: conceptualization, resources, and writing—review and editing. JM: conceptualization, data curation, funding acquisition, methodology, project administration, resources, software, supervision, validation, visualization, and writing—review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was conducted during OS's PhD study, which was supported by the Ethiopian Ministry of Higher Education.

## Acknowledgments

The authors would like to acknowledge the Biotechnology Research Institute of Bahir Dar University, Ethiopia for the financial support during blood sample collection and processing [Reference number: 1/4449/1.11.10 (BRI-BDU)]. This study was also partially supported by the CGIAR Research Program on Livestock (CRP Livestock) and accordingly, the authors thank all donors and organizations contributing to the CGIAR Trust Fund. The goat owners in Benishangul Gumuz region and all experts and development agents in the study areas are duly acknowledged for their cooperation during data collection.

## 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.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

## 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.2024.1353026/full#supplementary-material>

## References

- Abegaz, S. K., Mwai, O., Grum, G., Haile, A., Rischkowsky, B., Solomon, G., et al. (2014). "Review of goat research and development projects in Ethiopia," in *ILRI project report* (Nairobi, Kenya: International Livestock Research Institute).
- Ahbara, A., Bahbahani, H., Almathen, F., Al Abri, M., Agoub, M. O., Abeba, A., et al. (2019). Genome wide variation, candidate regions and genes associated with fat deposition and tail morphology in Ethiopian indigenous sheep. *Front. Genet.* 9, 699. doi:10.3389/fgene.2018.00699
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109
- Alfonzo, E. P. M., Barbosa Da Silva, M. V. G., Dos Santos Daltro, D., Stumpf, M. T., Dalcin, V. C., Kolling, G., et al. (2016). Relationship between physical attributes and heat stress in dairy cattle from different genetic groups. *Int. J. Biometeorol.* 60, 245–253. doi:10.1007/s00484-015-1021-y
- Amiri, G. Z., Ayatollahi, M. A., Asadollahpour, N. H., and Esmailzadeh, A. (2023). Comparative genomic analysis uncovers candidate genes related with milk production and adaptive traits in goat breeds. *Sci. Rep.* 30 (1), 8722. doi:10.1038/s41598-023-35973-0
- Amukamara, A. U., Washington, J. T., Sanchez, Z., McKinney, E. C., Moore, A. J., Schmitz, R. J., et al. (2020). More than DNA methylation: does pleiotropy drive the complex pattern of evolution of Dnmt1? *Front. Ecol. Evol.* 8, 4. doi:10.3389/fevo.2020.00004
- Andersson, L., and Georges, M. (2004). Domestic animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5, 202–212. doi:10.1038/nrg1294
- Ayalew, W., Wu, X., Tarekegn, G. M., Sisay Tessema, T., Naboulsi, R., Van Damme, R., et al. (2023). Whole-genome resequencing reveals selection signatures of abigar cattle for local adaptation. *Animals* 13 (20), 3269. doi:10.3390/ani13203269
- Bagath, M., Krishnan, G., Devaraj, C., Rashamol, V. P., Pragna, P., Lees, A. M., et al. (2019). The impact of heat stress on the immune system in dairy cattle: a review. *Res. Vet. Sci.* 126, 94–102. doi:10.1016/j.rvsc.2019.08.011
- Berihulay, H., Liu, X., Li, Y., Jiang, L., and GebrekidanMa, B. Y. (2019). Whole genome resequencing reveals selection SignaturAssociated with important traits in Ethiopian indigenous goat populations. *Front. Genet.* 28 (10), 1190. doi:10.3389/fgene.2019.01190
- Bewick, A. J., Sanchez, Z., Mckinney, E. C., Moore, A. J., Moore, P. J., and Schmitz, R. J. (2019). Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics Chromatin* 12, 6. doi:10.1186/s13072-018-0246-5
- Bickhart, D. M., Rosen, B. D., Koren, S., Sayre, B. L., Hastie, A. R., Chan, S., et al. (2017). Single-molecule sequencing and chromatin conformation capture enable *de novo* reference assembly of the domestic goat genome. *Nat. Genet.* 49, 643–650. doi:10.1038/ng.3802
- Bionaz, M., Osorio, J., and Loo, J. J. (2015). TRIENNIAL LACTATION SYMPOSIUM: nutrigenomics in dairy cows: nutrients, transcription factors, and techniques. *J. Animal Sci.* 93 (12), 5531–5553. doi:10.2527/jas.2015-9192
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu.170
- Brand, B., Scheinhardt, M. O., Friedrich, J., Zimmer, D., Reinsch, N., Ponsuksili, S., et al. (2016). Adrenal cortex expression quantitative trait loci in a German Holstein × Charolais cross. *BMC Genet.* 17, 135. doi:10.1186/s12863-016-0442-x
- Brito, L. F., Kijas, J. W., Ventura, R. V., Sargolzaei, M., Porto-Neto, L. R., Cánovas, A., et al. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics* 18, 229. doi:10.1186/s12864-017-3610-0
- Buchmann, R., and Hazelhurst, S. (2014). *Genesis manual*. Johannesburg: University of the Witwatersrand. Available at: <http://www.bioinf.wits.ac.za/software/genesis/Genesis.pdf>.
- Central Statistical Agency (CSA) (2021). Report on livestock and livestock characteristics (private peasant holdings). *Agric. Sample Surv. 2020/21 [2013 e.c.]*, Vol. II Stat. Bull., 587.
- Chen, R., Liu, J., Zhang, Y., Cai, W., Zhang, X., Xu, Y., et al. (2022). Association analysis between reproduction genes *INHA*, *PGR*, *RARG* with lamb and other traits of Liaoning cashmere goats. *Anim. Biotechnol.* 34 (7), 2094–2105. doi:10.1080/10495398.2022.2077212
- Chen, W. T., Huang, W. Y., Chen, T., Salawu, E. O., Wang, D., Lee, Y. Z., et al. (2016). Structure and function of chicken interleukin-1 beta mutants: uncoupling of receptor binding and *in vivo* biological activity. *Sci. Rep.* 6, 27729. doi:10.1038/srep27729
- Colli, L., Milanesi, M., Talenti, A., Bertolini, F., Chen, M., Crisà, A., et al. (2018). Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genet. Sel. Evol.* 50, 58. doi:10.1186/s12711-018-0422-x
- Daly, K. G., Delsler, P. M., Mullin, V. E., Scheu, A., Mattiangeli, V., Teasdale, M. D., et al. (2018). Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Sci. Mag.* 361, 6397. doi:10.1126/science.aas9411
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Della Chiesa, M., Falco, M., Parolini, S., Bellora, F., Petretto, A., Romeo, E., et al. (2010). GPR56 as a novel marker identifying the CD56dull CD16+ NK cell subset both in blood stream and in inflamed peripheral tissues. *Int. Immunol.* 22, 91–100. doi:10.1093/intimm/dxp116
- FARM-Africa (1996). "Goat types of Ethiopia and Eritrea," in *Physical description and management systems* (Nairobi, Kenya: London, UK and ILRI (International Livestock Research Institute), 76. Published jointly by FARM-Africa.
- Fattovich, R. (2019). From community to state: the development of the aksumite polity (northern Ethiopia and Eritrea), c. 400 BC–AD 800. *J. Archaeol. Res.* 27, 249–285. doi:10.1007/s10814-018-9122-x
- Gebru, T. (2009). *The Ethiopian revolution: war in the horn of Africa*. New Haven: Yale University, 149. doi:10.12987/yale/9780300141634.001.0001
- Gely-Pernot, A., Raverdeau, M., Célébi, C., Dennefeld, C., Feret, B., Klopfenstein, M., et al. (2012). Spermatogonia differentiation requires retinoic acid receptor  $\gamma$ . *Endocrinology* 153 (1), 438–449. doi:10.1210/en.2011-1102
- Getinet, A., Hegde, B. P., Bekele, T., Enyew, N., and Workneh, A. (2005). Phenotypic characterization of goat types in northwestern Ethiopia. *Ethiop. J. Animal Prod.* 5, 13–32.
- Gibson, J., Morton, N. E., and Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15, 789–795. doi:10.1093/hmg/ddi493
- Gifford-Gonzalez, D., and Hanotte, O. (2011). Domesticating animals in Africa: implications of genetic and archaeological findings. *J. World Prehist.* 24, 1–23. doi:10.1007/s10963-010-9042-2
- Green, A. C., Poulton, I. J., Vrahnas, C., Häusler, K. D., Walkley, C. R., Wu, J. Y., et al. (2015). RAR $\gamma$  is a negative regulator of osteoclastogenesis. *J. Steroid Biochem. Mol. Biol.* 150, 46–53. doi:10.1016/j.jsbmb.2015.03.005
- Guang-Xin, E., Duan, X.-H., Zhang, J.-H., Huang, Y.-Fu, Zhao, Y.-Ju, Zhao, Z.-Q., et al. (2019). Genome-wide selection signatures analysis of litter size in Dazu black goats using single-nucleotide polymorphism. *Biotech* 9 (9), 336. doi:10.1007/s13205-019-1869-3
- Guo, J., Tao, H., Li, P., Li, L., Zhong, T., Wang, L., et al. (2018). Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Sci. Rep.* 8, 10405. doi:10.1038/s41598-018-28719-w
- Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., and Rege, J. E. (2002). African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336–339. doi:10.1126/science.1069878
- Hassen, H., Lababidi, S., Rischkowsky, B., Baum, M., and Tibbo, M. (2012). Molecular characterization of Ethiopian indigenous goat populations. *Trop. Anim. Health Prod.* 44, 1239–1246. doi:10.1007/s11250-011-0064-2
- Hatziminaoglou, Y., and Boyazoglu, J. (2004). The goat in ancient civilizations: from the fertile crescent to aegean sea. *Small Rum. Res.* 51, 2. doi:10.1016/j.smallrumres.2003.08.006
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211
- Islam, R., Li, Y., Liu, X., Berihulay, H., Abied, A., Gebreselassie, G., et al. (2019). Genome-wide runs of homozygosity, effective population size, and detection of positive selection signatures in six Chinese goat breeds. *Genes* 10, 938. doi:10.3390/genes10110938
- Kay, S., Skowronski, D., and Hunt, B. G. (2018). Developmental DNA methyltransferase expression in the fire ant *solenopsis invicta*. *Insect Sci.* 25, 57–65. doi:10.1111/1744-7917.12413
- Kijas, J. W. (2013). Detecting regions of homozygosity to map the cause of recessively inherited disease. *Methods Mol. Biol.* 1019, 331–345. doi:10.1007/978-1-62703-447-0\_14
- Kim, E.-S., Elbeltagy, A. R., Aboul-Naga, A. M., Rischkowsky, B., Sayre, B., Mwacharo, J. M., et al. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity* 116, 255–264. doi:10.1038/hdy.2015.94
- Kosgey, I. S., Baker, R. L., Udo, H. M. J., and van Arendonk, J. A. M. (2006). Successes and failures of small ruminant breeding programs in the tropics: a review. *Small Rum. Res.* 61, 1. doi:10.1016/j.smallrumres.2005.01.003
- Kumar, C., Song, S., Dewani, P., Kumar, M., Parkash, O., Ma, Y., et al. (2018). Population structure, genetic diversity and selection signatures within seven indigenous Pakistani goat populations. *Anim. Genet.* 49, 592–604. doi:10.1111/age.12722

- Lai, F. N., Zhai H.L., H. L., Cheng, M., Ma, J. Y., Cheng, S. F., Ge, W., et al. (2016). Whole-genome scanning for the litter size trait associated genes and SNPs under selection in dairy goat (*Capra hircus*). *Sci. Rep.* 6, 38096. doi:10.1038/srep38096
- Lala, and Hall, R. A. (2022). Adhesion G protein-coupled receptors: structure, signaling, physiology, and pathophysiology. *Physiol. Rev.* 102, 1587–1624. doi:10.1152/physrev.00027.2021
- Lee, W., Ahn, S., Taye, M., Sung, S., Lee, H. J., Cho, S., et al. (2016). Detecting positive selection of Korean native goat populations using next-generation sequencing. *Mol. Cells* 39 (12), 862–868. doi:10.14348/molcells.2016.0219
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv:1303.3997. doi:10.48550/arXiv.1303.3997
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324
- Li, H. B., Liu, Y. A., Gu, Z. T., Li, L., Liu, Y. S., Wang, L., et al. (2018). p38 MAPK-MK2 pathway regulates the heat-stress-induced accumulation of reactive oxygen species that mediates apoptotic cell death in glial cells. *Oncol. Lett.* 15, 775–782. doi:10.3892/ol.2017.7360
- Li, X., Su, R., Wan, W., Zhang, W., Jiang, H., Qiao, X., et al. (2017). Identification of selection signals by large-scale whole-genome resequencing of cashmere goats. *Sci. Rep.* 7, 15142. doi:10.1038/s41598-017-15516-0
- Lohnes, D., Kastner, P., Dierich, A., Mark, M., LeMeur, M., and Chambon, P. (1993). Function of retinoic acid receptor  $\gamma$  in the mouse. *Cell* 73, 643–658. doi:10.1016/0092-8674(93)90246-M
- Lu, F., Zhao, M. D., Chen, S. Y., Jia, X. B., and Lai, S. J. (2015). Investigation of genetic susceptibility to nonspecific digestive disorder between TYK2, JAK1, and STAT3 genes in rabbits. *Livest. Sci.* 181, 137–142. doi:10.1016/j.livsci.2015.08.014
- Madhusoodan, A. P., Sejian, V., Afsal, A., Bagath, M., Krishnan, G., Savitha, S. T., et al. (2019). Differential expression patterns of candidate genes pertaining to productive and immune functions in hepatic tissue of heat-stressed Salem Black goats. *Biol. Rhythm Res.* 2, 809–820. doi:10.1080/09291016.2019.1607213
- Maggio, M., Guralnik, J. M., Longo, D. L., and Ferrucci, L. (2006). Interleukin-6 in aging and chronic disease: a magnificent pathway. *J. Gerontol. Biol. Sci.* 61, 575–584. doi:10.1093/gerona/61.6.575
- Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., et al. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Anim. Genet.* 46, 110–121. doi:10.1111/age.12259
- McGowan, S., Jackson, S. K., Jenkins-Moore, M., Dai, H.-H., Chambon, P., and Snyder, J. M. (2000). Mice bearing deletions of retinoic acid receptors demonstrate reduced lung elastin and alveolar numbers. *Am. J. Respir. Cell Mol. Biol.* 23, 162–167. doi:10.1165/ajrcmb.23.2.3904
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi:10.1101/gr.107524.110
- McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., et al. (2008). Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83, 359–372. doi:10.1016/j.ajhg.2008.08.007
- Mohan, M., Malayer, J. R., Geisert, R. D., and Morgan, J. L. (2001). Expression of retinol-binding protein messenger RNA and retinoic acid receptors in preattachment bovine embryos. *Mol. Reprod. Dev.* 60, 289–296. doi:10.1002/mrd.1090
- Montoya, M., Sancho, D., Bonello, G., Collette, Y., Langlet, C., He, H. T., et al. (2002). Role of ICAM-3 in the initial interaction of T lymphocytes and APCs. *Nat. Immunol.* 3, 159–168. doi:10.1038/ni753
- Mrode, R., Tarekegn, G. M., Mwacharo, J. M., and Djikeng, A. (2018). Invited review: genomic selection for small ruminants in developed countries: how applicable for the rest of the world? *Animal* 12, 1333–1340. doi:10.1017/S1751731117003688
- Mulatu, W. (1991). “Resettlement and villagization”. A country study: Ethiopia,” in *Library of congress federal research division*. Editors, Thomas P. Of cansky and LaVerle Berry.
- Mwacharo, J. M., Kim, E. S., Elbeltagy, A. R., Aboul-Naga, A. M., Rischkowsky, B. A., and Rothschild, M. F. (2017). Genomic footprints of dryland stress adaptation in Egyptian fat-tail sheep and their divergence from East African and western Asia cohorts. *Scientific Rep.* 7, 17647. doi:10.1038/s41598-017-17775-3
- Naderi, S., Rezaei, H. R., Pompanon, F., Blum, M. G., Negrini, R., Naghash, H. R., et al. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17659–17664. doi:10.1073/pnas.0804782105
- Naderi, S., Rezaei, H. R., Taberlet, P., Zundel, S., Rafat, S. A., Naghash, H. R., et al. (2007). Large-scale mitochondrial DNA analysis of the domestic goat reveals six maternal lineages with high haplotype diversity. *PLoS ONE* 2, 10. doi:10.1371/journal.pone.0001012
- Nei, M., and Li, W. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* 76, 5269–5273. doi:10.1073/pnas.76.10.5269
- Onzima, R. B., Upadhyay, M. R., Doekes, H. P., Brito, L. F., Bosse, M., Kanis, E., et al. (2018). Genome-wide characterization of selection signatures and runs of homozygosity in Ugandan goat breeds. *Front. Genet.* 9, 318. doi:10.3389/fgene.2018.00318
- Pickrell, J. K., and Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8, e1002967. doi:10.1371/journal.pgen.1002967
- Ping, Y. Q., Mao, C., Xiao, P., Zhao, R. J., Jiang, Y., Yang, Z., et al. (2021). Structures of the glucocorticoid-bound adhesion receptor GPR97-Go complex. *Nature* 589, 620–626. doi:10.1038/s41586-020-03083-w
- Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Aara, G. A., et al. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. *Prepr. A. T. bioRxiv*. doi:10.1101/201178
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795
- Purfield, D. C., Berry, D. P., McParland, S., and Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genet.* 13, 70. doi:10.1186/1471-2156-13-70
- Purton, L. E. (2007). Roles of retinoids and retinoic Acid receptors in the regulation of hematopoietic stem cell self-renewal and differentiation. *PPAR Res.* 87934, 87934. doi:10.1155/2007/87934
- Rahmatalla, S. A., Arends, D., Reissmann, M., Ahmed, A. S., Wimmers, K., Reyer, H., et al. (2017). Whole genome population genetics analysis of Sudanese goats identifies regions harboring genes associated with major traits. *BMC Genet.* 18, 92. doi:10.1186/s12863-017-0553-z
- Rashamol, V. P., Sejian, V., Bagath, M., Krishnan, G., Beena, V., and Bhatta, R. (2019). Effect of heat stress on the quantitative expression patterns of different cytokine genes in Malabari goats. *Int. J. Biometeorol.* 63, 1005–1013. doi:10.1007/s00484-019-01713-1
- Rekik, E., Ahbara, A. M., Abate, Z., Goshme, S., Getachew, T., Haile, A., et al. (2021). Genomic analysis of 10 years of artificial selection in community-based breeding programs in two Ethiopian indigenous sheep breeds. *Anim. Genet.* 53, 447–451. doi:10.1111/age.13190
- Rothschild, M. F., Messer, L. A., and Vincent, A. (1997). Molecular approaches to improved pig fertility. *J. Reproduction Fertil. Suppl.* 52, 227–236. doi:10.1530/biosci.15.0016
- Rubin, C. J., Zody, M., Eriksson, J., Meadows, J. R. S., Sherwood, E., Webster, M. T., et al. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464, 587–591. doi:10.1038/nature08832
- Sander, G. R., and Powell, B. C. (2004). Structure and expression of the ovine Hoxc-13 gene. *Gene* 327 (1), 107–116. doi:10.1016/j.gene.2003.11.006
- Schulz, N. K. E., Wagner, C. I., Ebeling, J., Raddatz, G., Diddens-de Buhr, M. F., Lyko, F., et al. (2018). Dnmt1 has an essential function despite the absence of CpG DNA methylation in the red flour beetle *Tribolium castaneum*. *Sci. Rep.* 8, 16462–16510. doi:10.1038/s41598-018-34701-3
- Shi, J., Zhang, X., Wang, S., Wang, J., Du, B., Wang, Z., et al. (2016). Gpr97 is dispensable for metabolic syndrome but is involved in macrophage inflammation in high-fat diet-induced obesity in mice. *Sci. Rep.* 6, 24649. doi:10.1038/srep24649
- Sigdel, A., Abdollahi-Arpanahi, R., Aguilar, I., and Peñagaricano, F. (2019). Whole genome mapping reveals novel genes and pathways involved in milk production under heat stress in US Holstein cows. *Front. Genet.* 10, 928. doi:10.3389/fgene.2019.00928
- Sinmegn, M. T., Asres, A. G., and Shimeka, T. A. (2014). Determinants of childhood diarrhea among under five children in Benishangul Gumuz regional state, North west Ethiopia. *BMC Pediatr.* 14, 102. doi:10.1186/1471-2431-14-102
- Susta, L., Diel, D. G., Courtney, S., Cardenas-Garcia, S., Sundick, R. S., Miller, P. J., et al. (2015). Expression of chicken interleukin-2 by a highly virulent strain of Newcastle disease virus leads to decreased systemic viral load but does not significantly affect mortality in chickens. *Virol. J.* 12, 222–217. doi:10.1186/s12985-015-0353-x
- Svajer, U., Anderlüh, M., Jeras, M., and Obermajer, N. (2010). C-type lectin DC-SIGN: an adhesion, signalling and antigen-uptake molecule that guides dendritic cells in immunity. *Cell. Signal.* 22 (10), 1397–1405. doi:10.1016/j.cellsig.2010.03.018
- Tao, L., He, X., Jiang, Y., Liu, Y., Ouyang, Y., Shen, Y., et al. (2021). Genome-wide analyses reveal genetic convergence of prolificacy between goats and sheep. *Genes* 12 (4), 480. doi:10.3390/genes12040480
- Tarekegn, G. M., Khayatadeh, N., Liu, B., Osama, S., Haile, A., Rischkowsky, B., et al. (2020). Ethiopian indigenous goats offer insights into past and recent demographic dynamics and local adaptation in sub-Saharan African goats. *Evol. Appl.* 14, 1716–1731. doi:10.1111/eva.13118
- Tarekegn, G. M., Tesfaye, K., Mwai, O., Djikeng, A., Dessie, T., Birungi, J., et al. (2018). Mitochondrial DNA variation reveals maternal origins and demographic dynamics of Ethiopian indigenous goats. *Ecol. Evol.* 8, 1543–1553. doi:10.1002/ece3.3710
- Taye, M., Lee, W., Caetano-Anolles, K., Dessie, T., Hanotte, O., Mwai, O. A., et al. (2017). Whole genome detection of signature of positive selection in African cattle

- reveals selection for thermotolerance. *Animal Sci. Journal* 88, 1889–1901. doi:10.1111/asj.12851
- Truong, A. D., Park, B., Ban, J., and Hong, Y. H. (2016). The novel chicken interleukin 26 protein is overexpressed in T cells and induces proinflammatory cytokines. *Vet. Res.* 47, 65. doi:10.1186/s13567-016-0342-0
- Wang, J. J., Zhang, L. L., Zhang, H. X., Shen, C. L., Lu, S. Y., Kuang, Y., et al. (2013). Gpr97 is essential for the follicular versus marginal zone B-lymphocyte fate decision. *Cell Death Dis.* 4, e853. doi:10.1038/cddis.2013.346
- Wang, K., Liu, X., Qi, T., Hui, Y., Yan, H., Qu, L., et al. (2021). Whole-genome sequencing to identify candidate genes for litter size and to uncover the variant function in goats (*Capra hircus*). *Genomics* 113 (1), 142–150. doi:10.1016/j.ygeno.2020.11.024
- Wang, X., Liu, J., Zhou, G., Guo, J., Yan, H., Li, Y., et al. (2016). Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. *Sci. Rep.* 6, 38932. doi:10.1038/srep38932
- Washington, J. T., Cavender, K. R., Amukamara, A. U., McKinney, E. C., Schmitz, R. J., and Moore, P. J. (2021). The essential role of Dnmt1 in gametogenesis in the large milkweed bug *Oncopeltus fasciatus*. *Elife* 10, e62202. doi:10.7554/eLife.62202
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x
- Wendling, O., Ghyselinck, N. B., Chambon, P., and Mark, M. (2001). Roles of retinoic acid receptors in early embryonic morphogenesis and hindbrain patterning. *Development* 128, 2031–2038. doi:10.1242/dev.128.11.2031
- Williams, J. A., Kondo, N., Okabe, T., Takeshita, N., Pilchak, D. M., Koyama, E., et al. (2009). Retinoic acid receptors are required for skeletal growth, matrix homeostasis and growth plate function in postnatal mouse. *Dev. Biol.* 328, 315–327. doi:10.1016/j.ydbio.2009.01.031
- Wu, J.-H., Zhang, W.-G., Li, J.-Q., Yin, J., and Zhang, Y.-J. (2009). Hoxc13 expression pattern in cashmere goat skin during hair follicle development. *Agric. Sci. China* 8, 491–496. doi:10.1016/S1671-2927(08)60237-0
- Xiao, X., Mruk, D. D., and Cheng, C. Y. (2013). Intercellular adhesion molecules (ICAMs) and spermatogenesis. *Hum. Reprod. Update* 19 (2), 167–186. doi:10.1093/humupd/dms049
- Yasuda, J., Whitmarsh, A. J., Cavanagh, J., Sharma, M., and Davis, R. J. (2000). The JIP group of mitogen-activated protein kinase scaffold proteins. *Mol. Cell Biol.* 19 (10), 7245–7254. doi:10.1128/mcb.19.10.7245
- Zavarez, L. B., Utsunomiya, Y. T., Carmo, A. S., Neves, H. R. H., Carvalheiro, R., Ferencakovic, M., et al. (2015). Assessment of autozygosity in Nellore cows (*Bos indicus*) through high-density SNP genotypes. *Front. Genet.* 6, 5. doi:10.3389/fgene.2015.00005
- Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. *Proc. Natl. Acad. Sci. U. S. A.* 105, 11597–11604. doi:10.1073/pnas.0801317105
- Zhi, D., Da, L., Liu, M., Cheng, C., Wang, X., Wang, Y., et al. (2018). Whole genome sequencing of hulunbuir short-tailed sheep for identifying candidate genes related to the short-tail phenotype. *Gene Genom. Genet.* 3, 377–383. doi:10.1534/g3.117.300307