

Article

Open Access

# Genomic analysis of indigenous goats in Southwest Asia reveals evidence of ancient adaptive introgression related to desert climate

Hojjat Asadollahpour Nanaei<sup>1, #</sup>, Yudong Cai<sup>1, #</sup>, Akil Alshawi<sup>2</sup>, Jiayue Wen<sup>1</sup>, Tanveer Hussain<sup>3</sup>, Wei-Wei Fu<sup>1</sup>, Nai-Yi Xu<sup>1</sup>, Abdulameer Essa<sup>4</sup>, Johannes A. Lenstra<sup>5</sup>, Xihong Wang<sup>1, \*</sup>, Yu Jiang<sup>1, \*</sup>

<sup>1</sup> Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China

<sup>2</sup> Department of Internal and Preventive Medicine, College of Veterinary Medicine, University of Baghdad, Iraqi Ministry of Higher Education and Scientific Research, Bagdad 10090, Iraq

<sup>3</sup> Department of Molecular Biology, Virtual University of Pakistan, Lahore 54000, Pakistan

<sup>4</sup> Animal Genetics Resources Department, Directorate of Animal Resources, the Ministry of Iraqi Agriculture, Bagdad 10081, Iraq

<sup>5</sup> Faculty of Veterinary Medicine, Utrecht University, 3508 TD Utrecht, The Netherlands

## ABSTRACT

Understanding how evolutionary pressures related to climate change have shaped the current genetic background of domestic animals is a fundamental pursuit of biology. Here, we generated whole-genome sequencing data from native goat populations in Iraq and Pakistan. Combined with previously published data on modern, ancient (Late Neolithic to Medieval periods), and wild *Capra* species worldwide, we explored the genetic population structure, ancestry components, and signatures of natural positive selection in native goat populations in Southwest Asia (SWA). Results revealed that the genetic structure of SWA goats was deeply influenced by gene flow from the eastern Mediterranean during the Chalcolithic period, which may reflect adaptation to gradual warming and aridity in the region. Furthermore, comparative genomic analysis revealed adaptive introgression of the

*KITLG* locus from the Nubian ibex (*C. nubiana*) into African and SWA goats. The frequency of the selected allele at this locus was significantly higher among goat populations located near northeastern Africa. These results provide new insights into the genetic composition and history of goat populations in the SWA region.

**Keywords:** Goat; Adaptation; *KITLG*; Southwest Asia; Nubian ibex

## INTRODUCTION

The Southwest Asian (SWA) Fertile Crescent and adjacent areas are considered to be the origin of the domestication of most livestock and plant species (Zeder, 2008). However, this important geographic area has suffered from prolonged droughts over the past few thousand years. For instance, paleoenvironmental reconstructions suggest that the eastern

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2023 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 09 August 2022; Accepted: 11 October 2022; Online: 13 October 2022

Foundation items: This work was supported by the National Natural Science Foundation of China (32050410304, 32002140, 31822052, 91431572381) and National Thousand Youth Talents Plan to Y.J.

#Authors contributed equally to this work

\*Corresponding authors, E-mail: rebecca\_830410@163.com; yu.jiang@nwfau.edu.cn

part of this region gradually shifted to drier and warmer conditions at the end of the mid-Holocene (Clarke et al., 2016; Sun et al., 2021), coinciding with the early development of agriculture. The most drought-affected regions of the Fertile Crescent are located in Iraq and Iran, which are also part of the hot-arid region of the Afro-Asian desert belt (Ashraf et al., 2021; Kaniewski et al., 2012). As a result, indigenous domesticated animals of the region are well adapted to the local hot-desert environment (Al-Thuwaini et al., 2019; Asadollahpour Nanaei et al., 2022; Saadatabadi et al., 2021).

Gene flow between populations allows the spread of adaptive alleles (Arnold & Kunte, 2017; Razgour et al., 2018). Studies in humans and other species have shown that heterogeneity in ancestry and introgression can offer insights into evolutionary processes (Kingston et al., 2017). Furthermore, studies of ancient introgression in farm animals have indicated that adaptive introgression is an important resource of genetic variation in natural populations and may contribute to adaptations during domestication (Chen et al., 2018). For example, genomic evidence supports the occurrence of introgression in the genus *Bos*, with bovine introgression in Mongolian yaks in East Asia and banteng (*Bos javanicus*) introgression in Chinese Hainan cattle (Decker et al., 2014; Medugorac et al., 2017).

As one of the earliest domesticated animals (Alberto et al., 2018; Daly et al., 2018), archeozoological dating suggests that goats (*Capra hircus*) were domesticated as early as 10 000 years before present (BP) in the area spanning southeastern Anatolia, western Iran, and northern Iraq (Zagros Mountains). After domestication of their wild ancestors (bezoars), goats followed the spread of agriculture, reaching the Mediterranean 8 000 BP and north Africa 7 000 BP (Daly et al., 2018; Zeder & Hesse, 2000). Currently, the global goat population has grown to more than one billion, and goat farming is a common animal production sector in many countries, especially in the Middle East and Mediterranean (Gooki et al., 2018, 2019; Moghbeli et al., 2013).

Recent studies have provided a snapshot of domestic goat genetic diversity and geographic distribution patterns of mitochondrial DNA (mtDNA) and nuclear DNA variants. Although mtDNA sequences suggest weak geographic differentiation in present-day goat populations (Luikart et al., 2001), genome-wide single nucleotide polymorphism (SNP) studies suggest otherwise (Bertolini et al., 2018; Colli et al., 2018). A recent paleogenomic study of ancient genomes from SWA revealed three distinct Neolithic goat groups from different parts of the Fertile Crescent that contributed differentially to present-day populations in Europe, Africa, and Asia (Daly et al., 2018). Furthermore, a minority of early domesticated goats from the Zagros Mountains show a relatively strong genetic affinity to the bezoar (Daly et al., 2021). This suggests that domestication and post-domestication splits do not preclude subsequent contact between wild and domestic goats or between goat groups from different geographic regions.

In this study, we performed a large-scale genomic analysis of goat populations, including 50 ancient, 211 domestic, and 72 wild *Capra* genomes, to reconstruct the historical migration

and introgression processes that shaped the current genetic architecture of domestic goats in SWA and to identify the genetic basis underlying variation in desert climate adaptation in the region.

## MATERIALS AND METHODS

### Sample collection, DNA extraction, and sequencing

Whole-blood samples (15 mL) were collected from indigenous goats in Iraq ( $n=23$ ) and Pakistan ( $n=25$ ). DNA was isolated from whole blood using standard phenol-chloroform extraction (Köchli et al., 2005). Paired-end sequence data for all individuals were generated using the Illumina HiSeq 2500 platform (USA). In addition, previously published genome sequences from ancient ( $n=50$ ), modern ( $n=163$ ), and wild goats ( $n=72$ ) were obtained from the Sequence Read Archive ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Details of samples used in this study are described in Supplementary Tables S1–S3. High-quality reads from the present study and published data were aligned against the most recent goat reference genome (ARS1, GCF\_001704415.1) using the Burrows-Wheeler Aligner v0.7.17 (BWA) algorithm (Li & Durbin, 2010). Variant calling of sequence data was performed using GATK v4.2.4 (McKenna et al., 2010), and genotype likelihoods were calculated using the ANGSD program (Korneliussen et al., 2014). Finally, to minimize false-positive results, transition sites were discarded in ancient samples.

### Genetic diversity, population structure, and phylogenetic relationship

We investigated different genomic diversity parameters, including nucleotide diversity, inbreeding coefficient ( $F$ ), and linkage disequilibrium (LD) decay using PLINK v1 (Purcell et al., 2007), with default settings. A maximum-likelihood (ML) tree was constructed for the whole-genome SNP set of all modern domestic goats. The Interactive Tree of Life (ITOL) tool (<https://itol.embl.de/>) was used to display the topological structure. ADMIXTURE v1.3.0 was used to characterize the structure of modern goat populations (Alexander et al., 2009). We increased the number of ancestral clusters ( $k$ ) from  $k=2$  to  $k=6$  and ran the analysis 100 times for each  $k$ . Principal component analysis (PCA) was conducted using SmartPCA implemented in the EIGENSOFT package, with the “lsqproject” and “autoshrink” options (Patterson et al., 2006). To identify the population structure among all goat populations, we performed a haplotype-based approach among all individuals using ChromoPainter and fineSTRUCTURE to explore patterns of haplotype sharing (Lawson et al., 2012).

### Exploration of selective sweep regions

For selective sweep analysis, the  $F_{ST}$ ,  $\theta\pi$ , and Tajima's  $D$  statistics were determined to detect regions under selection. We calculated genome-wide weighted  $F_{ST}$  (Weir & Cockerham, 1984) as it is a more precise measure of average genetic distance between groups with unequal samples (Spolaore & Wacziarg, 2009). We then estimated nucleotide diversity  $\theta\pi$  using VCFtools v0.1.13 (--window-pi 50000 --window-pi-step 25000) (Danecek et al., 2011). For the entire

genome, sliding window analysis was performed with a window size of 50 kb and a step size of 25 kb. Average  $F_{ST}$  and  $\log_2(\theta\pi_{SWA}/\theta\pi_{Northeast\ Asia})$  values of SNPs in each window were calculated.

### Gene flow and demographic history analyses

To assess introgression patterns among all groups, we performed outgroup- $f_3$ , D, F-branch, and F4-ratio analyses using ADMIXTOOLS v7.0.2 (Patterson et al., 2012) and Dsuite software (Malinsky et al., 2021). We highlighted introgression candidates as trios with significant D values (Bonferroni-corrected  $P < 0.001$ ) and  $f_4$ -ratios ( $> 0.05$ ) between populations. Local ancestry of the genome haplotypes was further inferred using Loter (Dias-Alves et al., 2018). We used SMC++ (v1.10.0) to infer patterns of effective population size and population separations over historical time (Terhorst et al., 2017). TreeMix v1.13 (Pickrell & Pritchard, 2012) was used to model gene flow among the modern domestic goat groups. We chose argali as the outgroup to form the root of the tree rather than domestic goat breeds to avoid bias.

## RESULTS

### Datasets, genetic diversity, and population structure

We generated whole genomes from native goats in Iraq ( $n=23$ ) and Pakistan ( $n=25$ ) and combined our previously published data from modern Iranian and worldwide goats ( $n=163$ ; Supplementary Table S1). In addition, we obtained 50 complete ancient genomes of domestic goats from Anatolia, the Balkans, eastern Mediterranean (Jordan and Israel, traditionally denoted as Levant), Iran, and Turkmenistan from an earlier study (Daly et al., 2018; Supplementary Tables S1, S2), resulting in a total dataset of 261 domestic goats. The ancient samples included Neolithic (~8500–5000 BC), Chalcolithic (~5100–3700 BC), Bronze Age (~3300–990 BC), Iron Age (~700–330 BC), and Medieval (~1500–901 AD) individuals. We further merged these data with 72 wild caprid species, including different ibex-like species, bezoars, and markhor goats to provide a comprehensive assessment (Supplementary Table S3). The 211 present-day domestic samples were assigned to eight geographic subgroups, namely Europe (EUR), West Africa (WAFR), East Africa (EAFR), Iran (IRN), Iraq (IRQ), Pakistan (PKS), Bangladesh (BNG), and China (CHN) (Figure 1A).

Average nucleotide diversity ( $\theta\pi$ ) for all individuals was estimated at  $1.79e-3$  and both Iranian ( $1.89e-3$ ) and Iraqi ( $1.88e-3$ ) groups showed slightly higher diversity than other goat groups (Supplementary Figure S1A). The LD was clearly higher for PKS goats than for the other populations (Supplementary Figure S1B). Individual inbreeding coefficients (F) were relatively high for European (~0.22) goats but lower for both Iranian (~0.09) and Iraqi (~0.10) goats (Supplementary Figure S2).

A rooted ML phylogenetic tree based on whole-genome sequencing data (Supplementary Figure S3) showed geographic clusters EUR-WAFR-EAFR; IRQ-IRN; PKS-BNG and CHN. Furthermore, modern European goats were grouped with Neolithic populations from the western part of the Fertile Crescent, whereas Neolithic eastern goats were

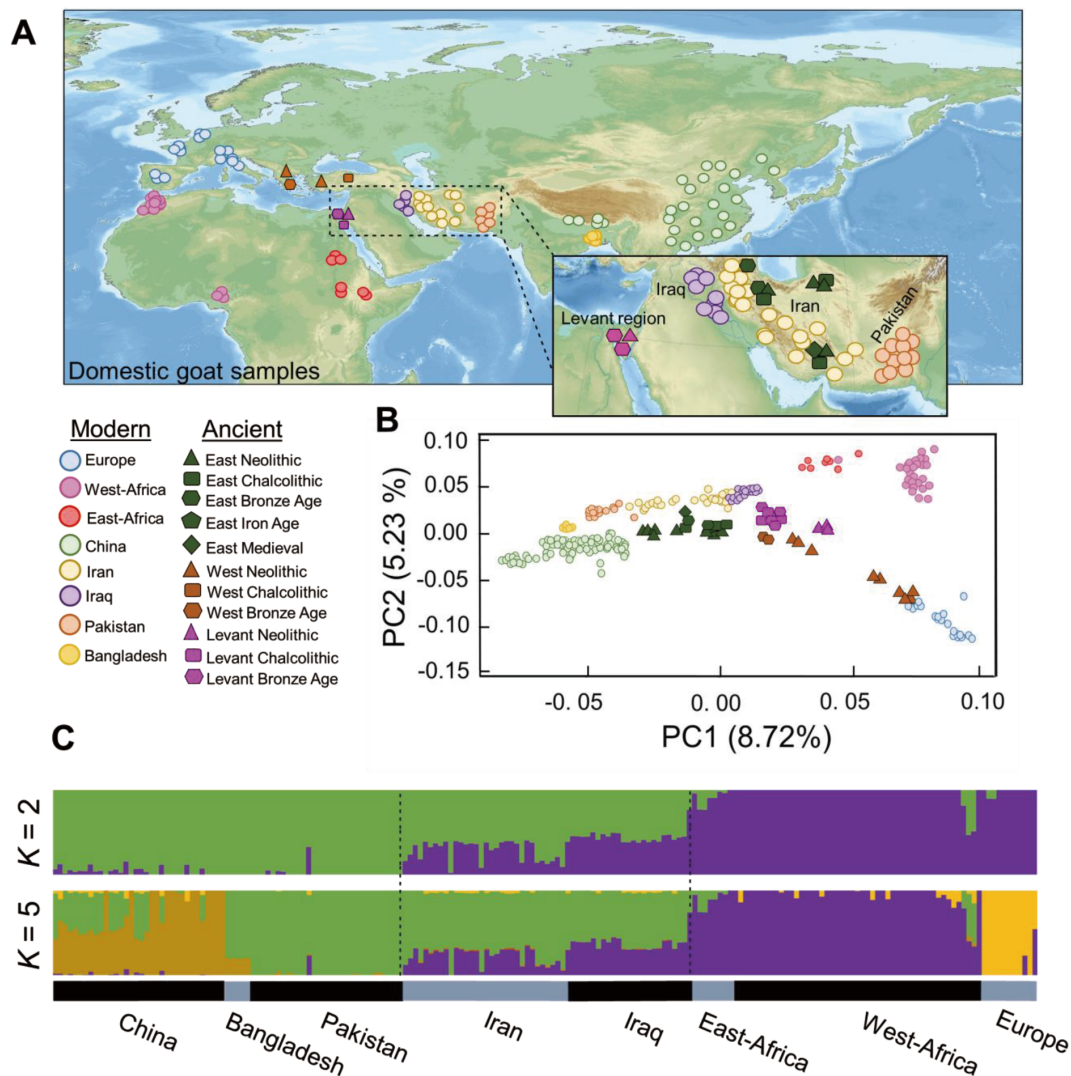
clustered between modern samples from China and South Asia. Based on PCA of modern domestic goats, PC1 showed a genetic cline from Chinese samples at the eastern end to European goats at the western end (Figure 1B), in agreement with the admixture pattern (Figure 1C; Supplementary Figure S4). In contrast, PC2 corresponded to the variation between Africa and northern EUR. As shown by Daly et al. (2018), ancient samples projected on the same plot were positioned near contemporary individuals from the same region. The ChromoPainter and fineSTRUCTURE results are in accordance with Figure 1, showing a low level of haplotype sharing in SWA (IRN and IRQ) samples with most other Asian goat populations (Supplementary Figure S5).

The mitochondrial phylogenetic tree showed that all six previously defined maternal haplogroups, i.e., A, B, C, D, F, and G, were present in our dataset (Supplementary Figure S6A, B). Haplogroup A was dominant in all domestic groups. The PKS samples were the most diverse, with four haplogroups, whereas the EUR, IRQ, and African individuals harbored only two haplogroups. Analysis of Y-chromosomal haplotype diversity showed that the haplogroup Y2A was dominant in both African and Iranian goat individuals, with a high frequency in Iraqi goat samples. Haplogroup Y2B was found exclusively in Chinese domestic goats and bezoars (Supplementary Figure S6C). Based on SMC++ analysis, the divergence times among all modern goat populations predated the expected domestication time from archaeological evidence (~11 000 years ago) (Supplementary Figure S7).

### High genetic affinity between modern goat populations in SWA and East Africa

To determine possible genome-wide introgression in present-day domestic groups, we calculated the D and F-branch ( $f_b$ ) statistics using Dsuite v3 (Malinsky et al., 2021). Results showed strong signals of introgression in the IRQ-IRN ( $f_b=0.61$ ,  $Z=35.21$ ), EAFR-IRQ ( $f_b=0.419$ ,  $Z=14.32$ ), and EAFR-IRN ( $f_b=0.374$ ,  $Z=11.98$ ) populations (Figure 2A). Additional cross-drainage introgression was also detected among several trios, e.g., common ancestor of African and IRQ ( $f_b=0.58$ ,  $Z=12.32$ ) and IRN ( $f_b=0.51$ ,  $Z=9.58$ ) populations, but no evidence of gene flow was detected between the WAFR and SWA populations. These values are higher than those for EAFR-IRQ and EAFR-IRN, suggesting that the latter may reflect introgression of the common ancestor of EAFR and WAFR in the Levant region. The  $f_b$  results also highlighted excessive allele sharing among the CHN/BNG, PKS/IRN, and EUR/WAFR populations (Figure 2A). These results were supported by TreeMix analysis, which revealed signals of genomic introgression from EAFR goats into IRQ ( $m=1$ ) and IRN ( $m=2$ ), explaining 99.7% and 99.9% of variance, respectively (Figure 2B; Supplementary Figure S8).

Quantification of gene flow from EAFR goats to IRQ and IRN populations using Loter (Dias-Alves et al., 2018) showed that the proportion of introgression was not uniform across present-day goats in different SWA regions. The frequency of gene flow gradually decreased with the increase in distance from EAFR (Figures 2C, D). Average frequency was highest (~65%) and lowest (~36%) in samples located in IRQ and



**Figure 1 Phylogenetic relationships among goat populations**

A: Map depicting locations of modern and ancient domestic samples analyzed in this study. B: Principal component analysis (PCA) with modern globally distributed goats and ancient goats around the Fertile Crescent. C: ADMIXTURE results for  $K=2$  and  $K=5$ , showing low cross-validation error.

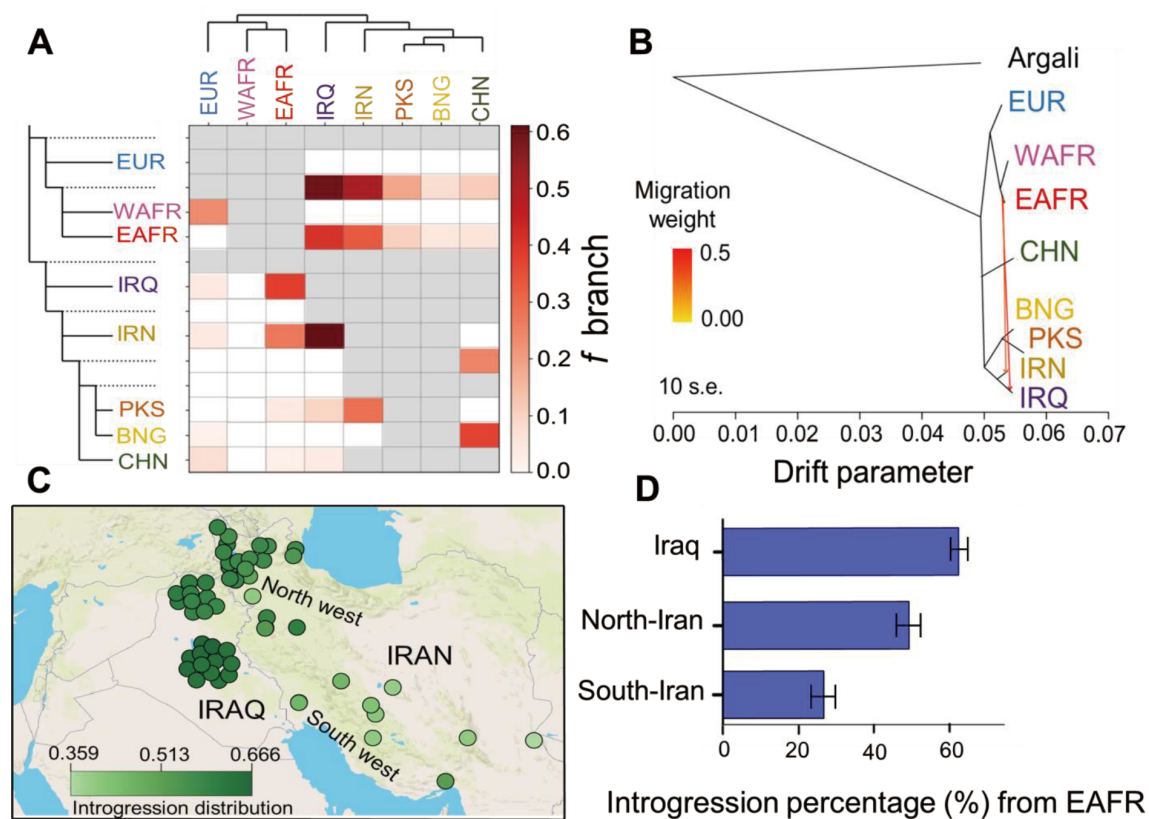
southern parts of IRN, respectively. Using *fd* (Malinsky et al., 2021), we further localized genomic regions with significant probability of introgression from EAFR goats into the IRQ and IRN populations (Supplementary Figure S9A, B). The length of the introgressed regions varied from several hundred to more than 200 000 base pairs. Selecting the top 1% of regions, we obtained 2 456 (2 218 genes) and 974 (966 genes) autosomal regions as high-confidence introgressed segments for the IRQ and IRN groups, respectively (Supplementary Tables S4, S5).

#### Historical evidence of gene flow from Levant into SWA

We next calculated D statistics to explore the Neolithic origin of all modern domestic goats, based on previously published ancient genomic data from the Near East. Notably, among all Neolithic goat groups from the Fertile Crescent, samples from the Levant region exhibited higher genetic affinity with modern African goats (Figure 3A), consistent with previous research (Daly et al., 2018). Furthermore, among the Asian modern

goat groups, the IRQ and IRN samples showed high allele sharing with Neolithic Levant goats. However, goat samples from South Asia (BNG and PKS) showed high genetic affinity with ancient samples from the Neolithic eastern group. To estimate the possible time of gene flow from Levant to SWA, we performed outgroup  $f_3$ -statistics (Ancient IRN, Ancient Levant; Outgroup) using ADMIXTOOLS software (Alexander et al., 2009). Results indicated that among all ancient samples from IRN, those from the Chalcolithic period (~4 200 BC) showed the closest genetic affinity with different ancient Levant populations (Figure 3B), thus dating the admixture event. These results were confirmed by D statistics in the form of (Ancient IRN, Neolithic-East; EAFR, Outgroup) and (Ancient IRN, Neolithic-East; Ancient Levant, Outgroup) (Supplementary Figure S10), again illustrating the high level of genetic continuity between Iranian goats with EAFR and different Levant populations in the Chalcolithic period.





**Figure 2 Gene flow from EAFR goats to SWA goat populations**

A:  $f_b$  summary of introgression among eight goat populations. ML tree was used as a basis for the branch statistic. Each cell shows  $f_b$  statistic between tree nodes (rows) and terminal nodes (column). Tree nodes in rows represent terminal nodes or (dotted lines) internal nodes within the tree. Gray cells in the matrix correspond to tests not consistent with the phylogeny. B: Inferred goat tree of mixture events deduced by TreeMix. Branch length is proportional to the drift of each population and scale bar shows 10 times the average standard error (SE) of entries in the sample covariance matrix. Argali was used as the outgroup to root the tree. C, D: Distribution of introgressed segments from EAFR, serving as proxy of the source of introgression from Levant, into IRQ and IRN goat populations.

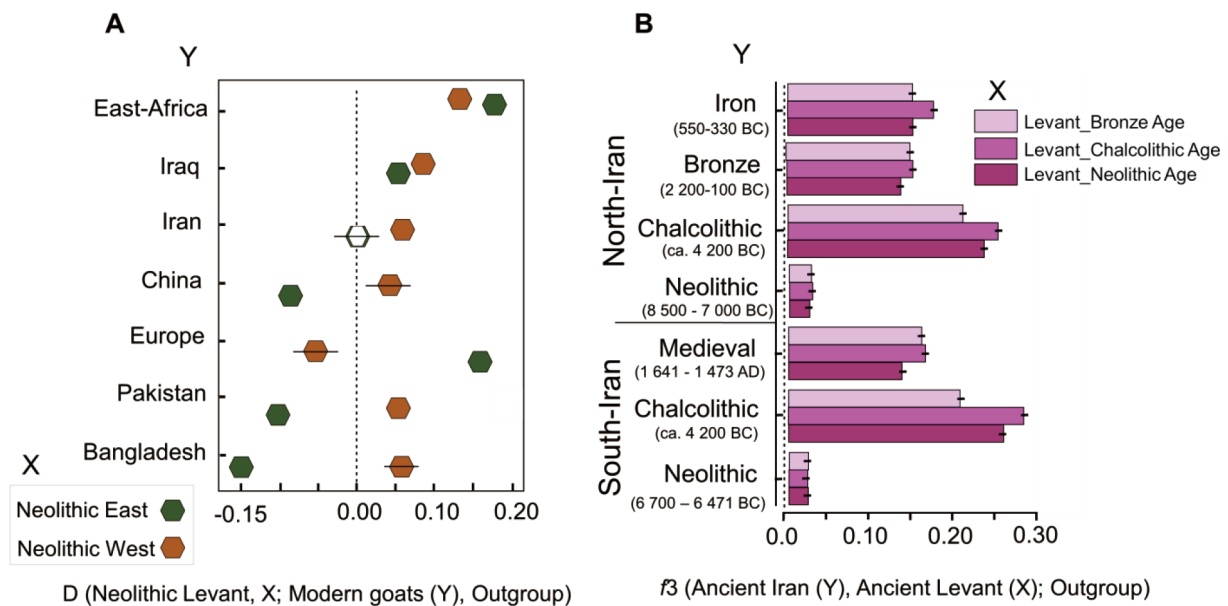
### Local adaptation by selective sweeps

Adaptation of indigenous livestock species in the SWA region to hot and dry desert ecosystems (Middleton, 1986; Vaghefi et al., 2019) may have influenced hair and skin traits, which are directly related to heat loss capacity (Berihulay et al., 2019; Deng & Xu, 2018). For example, most SWA indigenous sheep breeds have coarse wool coats, whereas most local goat breeds in the cold climates of northern Asia produce cashmere (China National Commission of Animal Genetic Resources, 2011). To gain insight into adaptations to desert climates, we compared native goats from SWA with those from Northeast Asia using pairwise genetic differentiation ( $F_{ST}$ ) and differences in nucleotide diversity ( $\pi$  In-ratio Northeast Asia/SWA). After merging consecutive windows with high values, we found 46 windows with the top 1% of values in both  $F_{ST}$  and In-ratio, spanning 72 candidate genes (Figure 4A; Supplementary Table S6). The top  $F_{ST}$  outlier window contained introgressed signals related to the *KITLG* gene on chromosome 5 (18 075–18 160 kb), which is associated with hair follicle traits (Guenther et al., 2014). The low  $\pi$  and Tajima's  $D$  values and high composite likelihood ratios (CLR) indicated strong positive selection at this locus in SWA goats (Figure 4B). Previous studies have shown that *KITLG* is an

important modulator of skin pigmentation in animals, playing a vital role in follicular survival and development of melanocyte lineage in adult skin (Yang et al., 2018). Melanin pigments play a direct role in protection against ultraviolet (UV) radiation and thermoregulation, which can pleiotropically influence a variety of behavioral and physiological traits in mammals (Deng & Xu, 2018; Sulem et al., 2007; Yang et al., 2018). We found no non-synonymous mutations within the coding region, suggesting a role for regulatory mutations.

### *KITLG*-selected region was possibly introgressed from the Nubian ibex into domestic goats

The strongest selection signal at the *KITLG* locus (18 116–18 161 kb) was found in the introgression segments introduced into the SWA population from EAFR goats. We constructed a haplotype map of the *KITLG*-selected region to explore the relationships among all populations. Interestingly, results showed two contrasting haplotypes in this genomic region that varied among the different *Capra* species, supporting potential relevance to adaptation (Figure 4C). The haplotype showing high frequency in Africa and Southwest and South Asia (hereafter AFR-ASA) had the highest similarity to ibexes, particularly the Nubian ibex (*C. nubiana*), whereas



**Figure 3 Genetic affinity between domestic goat populations**

A: Allele sharing between modern and Neolithic groups around the Fertile Crescent. Positive D statistic indicates higher level of allele sharing between Neolithic Levant goats and modern goat populations. Standard errors are shown with bars. Statistics with  $|Z \text{ score}| < 3$  are presented as unfilled symbols. B: Outgroup  $f_3$  statistics for shared genetic history between different ancient samples from Iran and Levant groups. Higher  $f_3$  values represent a closer genetic affinity of Iranian goats with Levant samples.

the bezoar haplotype was close to the *KITLG* regions in the European and Chinese goats. The frequencies of these introgressed loci were higher in EAFR (90%), IRQ (88%), WAFR (79%), IRN (77%), and PKS (72%), but much lower in EUR (2%) and CHN (5%) (Figure 4D), indicating a gradual decrease from Africa to SWA and South Asia. In addition, the haplotype network of the *KITLG*-selected region showed two highly divergent haplogroups in the worldwide domestic goats (Figure 4E). Of the SWA individuals, <75% belonged to the AFR-ASA haplogroup. To further assess the origin and diffusion of the *KITLG* allele in domestic goats, we carried out an in-depth genetic survey of ancient remains from the Fertile Crescent region. We detected an ancient animal with the *KITLG*-selected region in the Bronze age in Levant (dating to ~3500 YBP) (Supplementary Figure S11). Furthermore, one Iranian goat sample from the Medieval age (1641–1473 cal AD) also carried this genomic region. These findings are consistent with the hypothesis that gene flow from Levant goat groups into the SWA region occurred prior to the Medieval age.

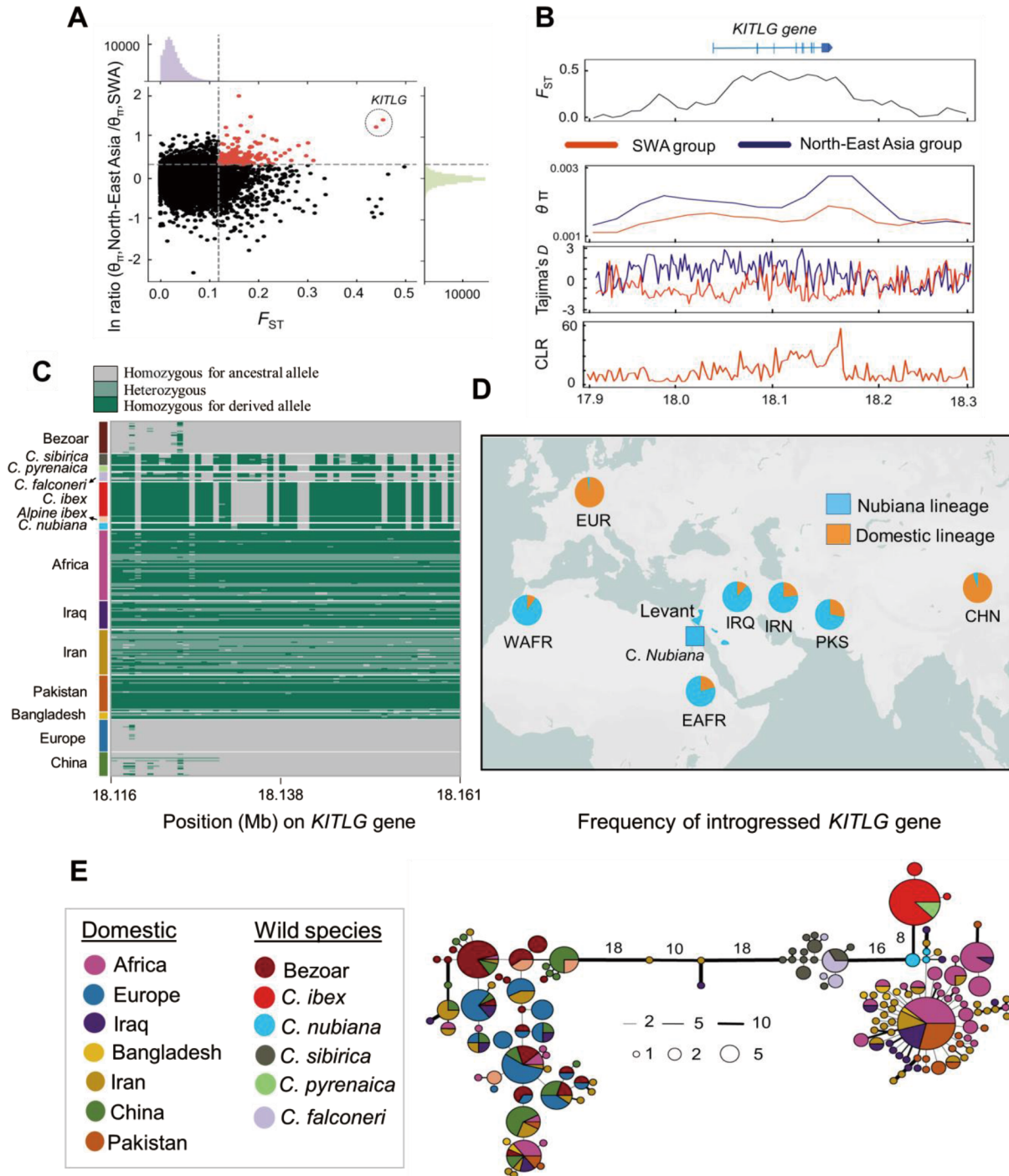
## DISCUSSION

We used worldwide ancient and modern goat genomes to elucidate the evolutionary events that shaped present-day SWA goat populations and characterize adaptive introgression and genetic changes after domestication in the region. Descriptive findings from ADMIXTURE, PCA, and TreeMix analyses indicated excessive allele sharing between the SWA and EAFR populations. Furthermore, patterns of allele sharing based on outgroup  $f_3$ , D, and  $f_b$  statistics confirmed that the SWA samples had a higher genetic affinity to EAFR than to other Asian goat groups. By exploring the genetic affinities of

ancient samples from northern and southern parts of IRN to different period samples from Levant and present-day EAFR, we observed a remarkable genetic continuity between individuals from the Chalcolithic period in IRN (7100–5700 YBP) and Levant/EAFR. The geographic distance between EAFR and SWA is considerable, but the EAFR influence may well reflect the influence of more proximate goats from the Levant region or Saudi-Arabian Peninsula, which were not sampled in this study.

The high level of allele sharing between ancient samples from IRN and Levant in the Chalcolithic period coincides with known transitions to more arid conditions in the Middle East (6.5–5 kyr BP) (Sun et al., 2021). These results are also consistent with the high genetic continuity and reduced genetic differentiation among farmers from southern Levant (Israel and Jordan) and the Zagros Mountains (Iran) (Harney et al., 2018; Lazaridis et al., 2016). Thus, we propose extensive gene flow from Levant lineages into SWA populations after domestication, dating between ~7000 and 5000 YBP, suggesting a northeastern path of gene flow from Levant into eastern parts of the Fertile Crescent. Given that the timing of this gene flow coincides with a wet-to-dry transition in western IRN regions, such as Lake Zeribar and Lake Mirabad (Sun et al., 2021), we hypothesize that changing precipitation and temperature patterns may have favored introgression of adaptive genes from the Levant populations to the SWA region. The same climate transition is proposed to have driven the migration of heat-resistant indicine cattle to SWA after 4200 BP (Verdugo et al., 2019).

Recent studies support a major role of adaptive introgression in livestock animal evolution, acting during periods of range expansion and contraction (Bangs et al., 2018; Zheng et al., 2020). For example, introgression signals



**Figure 4** Genomic regions with selection signals in SWA group

A: Distribution of pairwise fixation index ( $F_{ST}$ ) (x-axis) and  $\pi$  In-ratio (y-axis) between SWA and Northeast Asian goats. B: Putative sweep region (Chr. 5, 17 901 to 18 300 kb) was additionally validated by  $F_{ST}$ ,  $\pi$ , Tajima's  $D$ , and CLR tests. C: Patterns of 32 465 high frequency SNPs in variable *KITLG* region (18 116–18 161 kb) are shown. D: Frequency of introgressed *KITLG* haplotypes in domestic goat populations. E: Haplotype network based on pairwise differences within selective sweep region (chr5: 1 8116–18 161 kb) in *KITLG*.

of alleles in olfactory-related genes from wild relatives are associated with climatic adaptation in domestic sheep (Cao et al., 2021). Furthermore, an introgressed locus harboring the *MUC6* gene from West Caucasian tur-like species into domestic goats is reported to confer pathogen resistance

(Zheng et al., 2020).

In the current study, the strongest signals of adaptive introgression were associated with a skin pigmentation-related gene (*KITLG*) on chromosome 5 (Deng & Xu, 2018; Yang et al., 2018). This gene encodes the ligand for the receptor

tyrosine kinase protein KIT, which plays essential roles in the development and differentiation of different cell types in the body, such as melanocytes, blood cells, and germ cells (Morrison-Graham & Takahashi, 1993). A regulatory region of human *KITLG* contains an enhancer that regulates hair follicle expression (Guenther et al., 2014). As the frequency of the adaptive *KITLG* allele was fixed in the Nubian ibex and nearly fixed in the AFR-ASA populations, we believe this gene may have been introgressed from this wild species. The Nubian ibex is distributed around northern Africa and the Arabian Peninsula, geographically close to the goat domestication center in SWA. It inhabits a dry desert environment with intense solar radiation and extreme ambient temperatures (Castelló et al., 2016; Chebii et al., 2020). As several hair/wool characteristics are directly related to heat loss and improve the adaptive flexibility of livestock to extreme environmental conditions (McManus et al., 2020), we hypothesize that the *KITLG* gene may have played a role in the evolution of goat lineages in SWA during history.

Based on haplotype analysis of the *KITLG*-selected region, we suggest two possible scenarios to explain the high similarity between haplotypes in AFR-ASA and Nubian ibexes. The first hypothesis is introgression from the Nubian ibex to ancestral goat populations around Northeast Africa and Levant regions. This scenario is conceivable as we identified extensive ancient gene flow from Levant populations into the SWA region. Furthermore, in this scenario, the high frequency of the adaptive *KITLG* allele in West African goats could be explained by previously documented gene flow from the northeast to the western and central parts of Africa during the Holocene, closely related to the expansion of herders (Schlebusch, 2019). A second possibility may be that multiple separate introgression events occurred between Nubian ibexes and domestic goats in different geographical regions. However, due to the limited sample size, our data do not allow any strong conclusions.

## CONCLUSIONS

By analyzing genome-wide data from worldwide domestic goats and wild caprids, we provide the first glimpse into the genomic structure of native goat populations in SWA. Results indicated that the genetic pool of the SWA goat populations was affected by gene flow from Levant lineages in the Chalcolithic period. Furthermore, we found evidence of adaptive introgression of the *KITLG* allele from the Nubian ibex into domestic goats, which showed a strong signature of selection in goat populations located in desert climates. This study provides new insights into the evolutionary history of SWA goats and the importance of introgression in adaptation to past environmental challenges in this region.

## DATA AVAILABILITY

Raw sequencing data that support the findings of this study have been deposited in the NCBI database (BioProjectID PRJNA801057). Furthermore, whole-genome sequencing data have been submitted to the Genome Sequence Archive (<https://ngdc.cncb.ac.cn/gsa>) and Science Data Bank (<https://www.scidb.cn/en>) under accession No. CRA007867

and 10.57760/sciencedb.j00139.00032, respectively.

## SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

Y.J. and X.W. led the project and designed and conceived the study. H.A.N., Y. C., and J.Y.W. performed the data analysis. A.A., T.H., and A.E. collected the local goat samples. A.A. provided critical information about Iraqi goats. H.A.N. prepared the first draft of the manuscript. Y.J., X.W., J.A.L., W.W.F., and N.Y.X. reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

## ACKNOWLEDGMENTS

We thank Northwest A&F University for providing the high-performance computing platform and computing resources. We express our thanks to the owners of the goats for donating local samples.

## REFERENCES

- Alberto FJ, Boyer F, Orozco-terWengel P, Streeter I, Servin B, Villemereuil PD, et al. 2018. Convergent genomic signatures of domestication in sheep and goats. *Nature Communications*, **9**(1): 813.
- Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**(9): 1655–1664.
- Al-Thuwaini TM, Al-Shuhaib MBS, Hussein ZM. 2019. Heat shock protein 70 polymorphism associated with physio-biochemical parameters of Awassi and Arabi Iraqi sheep. *Euphrates Journal of Agriculture Science*, **11**(4): 84–95.
- Arnold ML, Kunte K. 2017. Adaptive genetic exchange: a tangled history of admixture and evolutionary innovation. *Trends in Ecology & Evolution*, **32**(8): 601–611.
- Asadollahpour Nanaei H, Kharrati-Koopaei H, Esmailizadeh A. 2022. Genetic diversity and signatures of selection for heat tolerance and immune response in Iranian native chickens. *BMC Genomics*, **23**(1): 224.
- Ashraf S, Nazemi A, AghaKouchak A. 2021. Anthropogenic drought dominates groundwater depletion in Iran. *Scientific Reports*, **11**(1): 9135.
- Bangs MR, Douglas MR, Musmann SM, Douglas ME. 2018. Unraveling historical introgression and resolving phylogenetic discord within *Catostomus* (Osteichthys: Catostomidae). *BMC Evolutionary Biology*, **18**(1): 86.
- Berihulay H, Abied A, He XH, Jiang L, Ma YH. 2019. Adaptation mechanisms of small ruminants to environmental heat stress. *Animals*, **9**(3): 75.
- Bertolini F, Servin B, Talenti A, Rochat E, Kim ES, Oget C, et al. 2018. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genetics Selection Evolution*, **50**(1): 57.
- Cao YH, Xu SS, Shen M, Chen ZH, Gao L, Lv FH, et al. 2021. Historical introgression from wild relatives enhanced climatic adaptation and resistance to pneumonia in sheep. *Molecular Biology and Evolution*, **38**(3): 838–855.

- Castelló JR, Huffman B, Groves C. 2016. *Bovids of the World: Antelopes, Gazelles, Cattle, Goats, Sheep, and Relatives*. Princeton: Princeton University Press.
- Chebii VJ, Oyola SO, Kotze A, Entfellner JBD, Mutuku JM, Agaba M. 2020. Genome-Wide analysis of Nubian ibex reveals candidate positively selected genes that contribute to its adaptation to the desert environment. *Animals*, **10**(11): 2181.
- Chen NB, Cai YD, Chen QM, Li R, Wang K, Huang YZ, et al. 2018. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nature Communications*, **9**(1): 2337.
- China National Commission of Animal Genetic Resources. 2011. *Animal Genetic Resources in China: Sheep and Goats*. Beijing: China Agriculture Press. (in Chinese)
- Clarke J, Brooks N, Banning EB, Bar-Matthews M, Campbell S, Clare L, et al. 2016. Climatic changes and social transformations in the Near East and North Africa during the 'long' 4th millennium BC: a comparative study of environmental and archaeological evidence. *Quaternary Science Reviews*, **136**: 96–121.
- Colli L, Milanese 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. *Genetics Selection Evolution*, **50**(1): 58.
- Daly KG, Delser PM, Mullin VE, Scheu A, Mattiangeli V, Teasdale MD, et al. 2018. Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Science*, **361**(6397): 85–88.
- Daly KG, Mattiangeli V, Hare AJ, Davoudi H, Fathi H, Doost SB, et al. 2021. Herded and hunted goat genomes from the dawn of domestication in the Zagros mountains. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(25): e2100901118.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.
- Decker JE, McKay SD, Rolf MM, Kim J, Molina Alcalá A, Sonstegard TS, et al. 2014. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics*, **10**(3): e1004254.
- Deng L, Xu SH. 2018. Adaptation of human skin color in various populations. *Hereditas*, **155**: 1.
- Dias-Alves T, Mairal J, Blum MGB. 2018. Loter: a software package to infer local ancestry for a wide range of species. *Molecular Biology and Evolution*, **35**(9): 2318–2326.
- Gooki FG, Mohammadabadi M, Fozi MA, Soflaei M. 2019. Association of Biometric Traits with growth hormone gene diversity in Raini cashmere goats. *Walailak Journal of Science and Technology (WJST)*, **16**(7): 499–508.
- Gooki FG, Mohammadabadi MR, Asadi Fozi M. 2018. Polymorphism of the growth hormone gene and its effect on production and reproduction traits in goat. *Iranian Journal of Applied Animal Science*, **8**(4): 653–659.
- Guenther CA, Tasic B, Luo LQ, Bedell MA, Kingsley DM. 2014. A molecular basis for classic blond hair color in Europeans. *Nature Genetics*, **46**(7): 748–752.
- Harney É, May H, Shalem D, Rohland N, Mallick S, Lazaridis I, et al. 2018. Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nature Communications*, **9**(1): 3336.
- Kaniewski D, Van Campo E, Weiss H. 2012. Drought is a recurring challenge in the Middle East. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(10): 3862–3867.
- Kingston SE, Parchman TL, Gompert Z, Buerkle CA, Braun MJ. 2017. Heterogeneity and concordance in locus-specific differentiation and introgression between species of towhees. *Journal of Evolutionary Biology*, **30**(3): 474–485.
- Köchl S, Niederstätter H, Parson W. 2005. DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. *Methods in Molecular Biology*, **297**: 13–30.
- Korneliusson TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*, **15**(1): 356.
- Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of population structure using dense haplotype data. *PLoS Genetics*, **8**(1): e1002453.
- Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, **536**(7617): 419–424.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, **26**(5): 589–595.
- Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, Taberlet P. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(10): 5927–5932.
- Malinsky M, Matschiner M, Svardal H. 2021. Dsuite-Fast *D*-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, **21**(2): 584–595.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297–1303.
- McManus CM, Faria DA, Lucci CM, Louvandini H, Pereira SA, Paiva SR. 2020. Heat stress effects on sheep: are hair sheep more heat resistant?. *Theriogenology*, **155**: 157–167.
- Medugorac I, Graf A, Grohs C, Rothhammer S, Zagdsuren Y, Gladyr E, et al. 2017. Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics*, **49**(3): 470–475.
- Middleton NJ. 1986. A geography of dust storms in South-West Asia. *Journal of Climatology*, **6**(2): 183–196.
- Moghbeli SM, Barazandeh A, Vatankhah M, Mohammadabadi M. 2013. Genetics and non-genetics parameters of body weight for post-weaning traits in Raini Cashmere goats. *Tropical Animal Health and Production*, **45**(7): 1519–1524.
- Morrison-Graham K, Takahashi Y. 1993. Steel factor and c - Kit receptor: from mutants to a growth factor system. *BioEssays*, **15**(2): 77–83.
- Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan YP, et al. 2012. Ancient admixture in human history. *Genetics*, **192**(3): 1065–1093.
- Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genetics*, **2**(12): e190.
- Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, **8**(11): e1002967.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**(3): 559–575.
- Razgour O, Taggart JB, Manel S, Juste J, Ibáñez C, Rebelo H, et al. 2018. An integrated framework to identify wildlife populations under threat from climate change. *Molecular Ecology Resources*, **18**(1): 18–31.



- Saadatabadi LM, Mohammadabadi M, Amiri Ghanatsaman Z, Babenko O, Stavetska R, Kalashnik O, et al. 2021. Signature selection analysis reveals candidate genes associated with production traits in Iranian sheep breeds. *BMC Veterinary Research*, **17**(1): 369.
- Schlebusch C. 2019. Population migration and adaptation during the African Holocene: a genetic perspective. In: Sahle Y, Reyes-Centeno H, Bentz C. Modern Human Origins and Dispersal. Tübingen: Kerns Verlag, 261–283.
- Spolaore E, Wacziarg R. 2009. The diffusion of development. *The Quarterly Journal of Economics*, **124**(2): 469–529.
- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, et al. 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics*, **39**(12): 1443–1452.
- Sun WY, Wang B, Zhang Q, Chen DL, Lu GN, Liu J. 2021. Middle east climate response to the saharan vegetation collapse during the mid-holocene. *Journal of Climate*, **34**(1): 229–242.
- Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, **49**(2): 303–309.
- Vaghefi SA, Keykhai M, Jahanbakhshi F, Sheikholeslami J, Ahmadi A, Yang H, et al. 2019. The future of extreme climate in Iran. *Scientific Reports*, **9**(1): 1464.
- Verdugo MP, Mullin VE, Scheu A, Mattiangeli V, Daly KG, Delsler PM, et al. 2019. Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent. *Science*, **365**(6449): 173–176.
- Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**(6): 1358–1370.
- Yang ZH, Shi H, Ma PC, Zhao SL, Kong QH, Bian TH, et al. 2018. Darwinian positive selection on the pleiotropic effects of *KITLG* explain skin pigmentation and winter temperature adaptation in Eurasians. *Molecular Biology and Evolution*, **35**(9): 2272–2283.
- Zeder MA. 2008. Animal domestication in the Zagros: an update and directions for future research. In: Vila E, Gourichon L, Choyke A, Buitenhuis H. Archaeozoology of the Near East VII. Lyon: Maison de l'Orient et de la Méditerranée, 243–277.
- Zeder MA, Hesse B. 2000. The initial domestication of goats (*Capra hircus*) in the Zagros mountains 10, 000 years ago. *Science*, **287**(5461): 2254–2257.
- Zheng ZQ, Wang XH, Li M, Li YJ, Yang ZR, Wang XL, et al. 2020. The origin of domestication genes in goats. *Science Advances*, **6**(21): eaaz5216.