Received: 18 September 2023 | Accepted: 11 October 2024

DOI: 10.1111/age.13487

RESEARCH ARTICLE

ANIMAL GENETICS WILEY

Genome-wide scans for signatures of selection in North African sheep reveals differentially selected regions between fat- and thin-tailed breeds

Slim Ben-Jemaa[1] | **Ghazi Yahyaoui**[2] | **Samia Kdidi**[3] | **Afef Najjari**[4] | **Johannes A. Lenstra**[5] | **Salvatore Mastrangelo**[6] | **Semir B. S. Gaouar**[7] | **Joram M. Mwacharo**[8] | **Touhami Khorchani**[3] | **Mohamed H. Yahyaoui**[3]

[1]Laboratoire Des Productions Animales et Fourragères, Institut National de la Recherche Agronomique de Tunisie, Université de Carthage, Ariana, Tunisia

[2]Département de Biologie, Faculté Des Sciences de Tunis, Université Tunis El Manar, Tunis, Tunisia

[3]Laboratoire d'Elevage et Faune Sauvage, Institut Des Régions Arides, Medenine, Tunisia

[4]Laboratoire de Microbiologie et Biomolécules Actives LR03ES03, Faculté Des Sciences de Tunis, Université Tunis El Manar, Tunis, Tunisia

[5]Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

[6]Dipartimento di Scienze Agrarie, Alimentari e Forestali, University of Palermo, Palermo, Italy

[7]Applied Genetic in Agriculture, Ecology and Public Health Laboratory, University of Tlemcen, Tlemcen, Algeria

[8]Small Ruminant Genomics, International Centre for Agricultural Research in the Dry Areas (ICARDA), Addis Ababa, Ethiopia

Correspondence
Slim Ben-Jemaa, Laboratoire des Productions Animales et Fourragères, Institut National de la Recherche Agronomique de Tunisie, Université de Carthage, 2049 Ariana, Tunisia.
Email: benjemaaslim@gmail.com

[Correction added on 17 December 2024, after first online publication: The copyright line was changed.]

Abstract

North Africa counts several sheep breeds that can be categorized as fat- and thin-tailed. The former are well adapted to dryland environments. In this study, we used 50K genome-wide single nucleotide polymorphism profiles from 462 animals representing nine fat-tailed and 13 thin-tailed sheep breeds across North Africa to localize genomic regions putatively under differential selective pressures between the two types of breeds. We observed genetic clines from east to west and from north to south. The east–west cline separates the fat- and thin-tailed breeds, with the exception of the fat-tailed Algerian Barbarine, which is closely related to a genetically homogeneous cluster of Moroccan and Algerian thin-tailed breeds. Using a combination of three extended haplotype homozygosity tests, we detected seven candidate regions under divergent selection between fat- and thin-tailed sheep. The strongest selection signals reside on chromosomes 1 and 13, with the latter spanning the *BMP2* gene, known to be associated with the fat-tail phenotype. Overall, the candidate regions under selection in fat-tailed sheep overlap with genes associated with adaptation to desert-like environments including adipogenesis, as well as heat and drought tolerance. Our results confirm previously reported candidate genes known to be a target of fat-tail selection in sheep but also reveal novel candidate genes specifically under selection in North African populations.

KEYWORDS

candidate gene, extended haplotype homozygosity, fat tail, North African sheep, selection signature, single nucleotide polymorphism

INTRODUCTION

Sheep are important assets to local agrarian economies across North African countries (Egypt, Libya, Tunisia, Algeria and Morocco) and play a critical role as a primary source of income in rural farming systems (Dutilly-Diane, 2007; Gaouar, 2009). The more than 70 million North African sheep (Food and Agriculture Organization

of the United Nations, 2019) are either thin- or fat-tailed and several breeds are transboundary. For instance, the fat-tailed Barbarine, the major breed in Tunisia, is also found in Libya and Algeria and the thin-tailed Ouled–Djellal, D'man and Hamra (Algerian name)/Beni–Guil (Moroccan name) breeds occur in Morocco and Algeria (Belabdi et al., 2019). Thin-tailed were the earliest domestic sheep to be introduced to Africa from their domestication center in the Near east via the Sinai Peninsula, between 7500 and 7000 BP. They spread westwards following continental and maritime trading routes and reached West Africa by 3700 BP (Muigai & Hanotte, 2013). Several Algerian and Moroccan thin-tailed breeds are now genetically homogeneous (Belabdi et al., 2019; Gaouar et al., 2015), while the morphological differences between the breeds are maintained. In a second migration wave, fat-tailed sheep entered the African continent from the Arabian Peninsula, by which local thin-tailed sheep were introgressed and/or replaced (Muigai & Hanotte, 2013). Today, thin-tailed sheep are prevalent in northwestern Africa (Morocco and Algeria), whereas fat-tailed sheep are mostly found in northeastern (Egypt and Libya) as well as in Eastern and Southern Africa. Analysis of mitochondrial DNA revealed a close relationship of Iberian and Moroccan sheep (Kandoussi et al., 2020) suggesting close maternal ancestry. Genome-wide 50K SNP profiles suggest a major South-European influence and a minor southwestern Asian influence on North African sheep (Ben-Jemaa et al., 2019). East African fat-tailed sheep belong to a different gene pool without European influence (Mwacharo et al., 2017).

The fat-tail phenotype is an important adaptive trait and provides an energy reserve in harsh climates (Rocha et al., 2021). Nowadays, the high production of mutton by the modern sheep industry has considerably decreased the importance of the fat-tail phenotype with consumers showing an increasing preference for low-fat meat. Accordingly, farmers in several North African countries have attempted to lower the fat content of their sheep through repeated and unsupervised cross-breeding with thin-tailed breeds. For instance, the Algerian Barbarine has been extensively crossed with Ouled–Djellal (Abdelkader et al., 2017). Identification of putative genomic regions differentially selected between fat- and thin-tailed sheep would help breeders to select for reduced tail size in local sheep breeds without resorting to cross-breeding.

Using high-density genotyping and whole-genome sequencing, several reports have identified genomic regions involved in fat-tailed and thin-tailed traits in Asian and African sheep (Ahbara et al., 2019; Dong et al., 2020; Mastrangelo et al., 2019; Moradi et al., 2012; Mwacharo et al., 2017). Although the overlap between these studies is limited, *PDGFD* (Dong et al., 2020; Li et al., 2020; Yuan et al., 2017; Zhao et al., 2020) and *BMP2* (Baazaoui et al., 2021; Bedhiaf-Romdhani et al., 2023; Mastrangelo et al., 2019; Moioli et al., 2015; Pan et al., 2019; Xu et al., 2023; Yuan et al., 2017; Zhao et al., 2020; Zhu et al., 2021) are often found to be associated with the tail type.

In this study, we focus on the sheep breeds in North African countries (Egypt, Libya, Tunisia, Algeria and Morocco). Combining Illumina Ovine SNP50K array genotypes for three Tunisian populations and one Libyan sheep breed with publicly available genomic data, we (i) provide a detailed picture of the genomic landscape of North African sheep, and (ii) identify putative genomic regions associated with differential selection between fat- and thin-tailed sheep phenotypes.

MATERIALS AND METHODS

Data merging and SNP filtering

For the purpose of this study, we used Illumina Ovine SNP50K BeadChip genotypes of 462 individuals belonging to 22 sheep breeds raised in five North African countries (Table 1). Genotypes from the three Tunisian sheep breeds, Tunisian Barbarine (BART), Noire de Thibar (BT) and Sicilio-Sarde (SS) and the Libyan Barbarine (BARL) breed have been reported previously (Ben-Jemaa et al., 2019). We retrieved publicly available data from 18 North African breeds (Belabdi et al., 2019; Gaouar et al., 2017; Mastrangelo et al., 2019; Mwacharo et al., 2017; nextGen project: http://projects.ensembl.org/nextgen/). These 18 breeds include: i) six Algerian populations – Barbarine (BRBA, $N=6$), D'men (DMNA, $N=8$), Hamra (HAMA, $N=16$), Ouled–Djellal (OLDA, $N=6$), Sidaoun (SDNA, $N=36$) and Tazegzawth (TZGA, $N=6$); ii) six Moroccan populations – Beni–Guil (BIGM, $N=6$), D'men (DMNM, $N=30$), Local Population (MOR, $N=72$), Ouled–Djellal (OLDM, $N=8$), Sardi (SRDM, $N=27$) and Timahdite (TMHM, $N=16$); and iii) six Egyptian sheep populations – Aburamad–Halaieb–Shalateen (AHS, $N=26$), Barki (BARK, $N=35$), Farafra (FRF, $N=32$), Ossimi (OSSI, $N=8$), Saidi (SAID, $N=46$) and Souhagi (SOUH, $N=20$). Half of these 18 breeds (BRBA, AHS, BARK, FRF, OSSI, SAID, SOUH, BARL and BART) are fat-tailed while the remainder are thin-tailed. During quality control of the public dataset we used PLINK v.1.09 software (Purcell et al., 2007) and removed (i) SNPs with call rates <90% or minor allele frequencies <0.01 (six SNPs removed); (ii) individuals with marker call rates <90% (one individual removed); and (iii) SNPs in linkage disequilibrium (LD) as revealed by PLINK default parameters (SNP window size 50, step 5 SNPs, r^2 0.5) (1890 SNPs removed). Finally, 32 613 SNPs overlapping between the datasets remained for subsequent analyses. The median distance between consecutive SNPs is 54 kb while, on average, 64% of the consecutive markers are separated by less than 70 kb (Table S1).

TABLE 1 Sheep breeds analyzed in this study.

Origin	Breed	Code	*N*[a]	Tail	Ho	Homogeneous ADMIXTURE cluster	References
Algeria	Barbarine Algeria	BRBA	6	Fat	0.339		Belabdi et al. (2019)
Algeria	Algerian D'men	DMNA	8	Thin	0.381		Gaouar et al. (2017)
Algeria	Hamra	HAMA	16	Thin	0.391		Belabdi et al. (2019)
Algeria	Ouled–Djellal from Algeria	OLDA	6	Thin	0.390		Belabdi et al. (2019)
Algeria	Sidaoun	SDNA	36	Thin	0.348	Y	Belabdi et al. (2019)
Algeria	Tazegzawth	TZGA	6	Thin	0.321	Y	Belabdi et al. (2019)
Egypt	Aburamad-Halaieb-Shalateen	AHS	26	Fat	0.351	Y	Mwacharo et al. (2017)
Egypt	Egyptian Barki	BARK	35	Fat	0.374		Mwacharo et al. (2017)
Egypt	Farafra	FRF	32	Fat	0.335		Mwacharo et al. (2017)
Egypt	Ossimi	OSSI	8	Fat	0.361		Mastrangelo et al. (2019)
Egypt	Saidi	SAID	46	Fat	0.338		Mwacharo et al. (2017)
Egypt	Souhagi	SOUH	20	Fat	0.355	Y	Mwacharo et al. (2017)
Lybia	Lybian Barbarine	BARL	14	Fat	0.375		Mastrangelo et al. (2019) and Ben-Jemaa et al. (2019)
Morocco	Beni–Guil Morocco	BIGM	6	Thin	0.393		Belabdi et al. (2019)
Morocco	Moroccan D'men	DMNM	30	Thin	0.352		Belabdi et al. (2019)
Morocco	Local population	MOR	72	Thin	0.379		nextGen project
Morocco	Ouled–Djellal from Morocco	OLDM	8	Thin	0.385		Belabdi et al. (2019)
Morocco	Sardi	SRDM	27	Thin	0.382		Belabdi et al. (2019)
Morocco	Timahdite	TMHM	16	Thin	0.390		Belabdi et al. (2019)
Tunisia	Tunisian Barbarine	BART	16	Fat	0.388		Ben-Jemaa et al. (2019)
Tunisia	Noire de Thibar	BT	17	Thin	0.373	Y	Ben-Jemaa et al. (2019)
Tunisia	Sicilio-Sarde	SS	11	Thin	0.390		Ben-Jemaa et al. (2019)

Abbreviations: AHS, Aburamad–Halaieb–Shalateen; BARK, Egyptian Barki; BARL, Lybian Barbarine; BART, Tunisian Barbarine; BIGM, Beni–Guil Morocco; BRBA, Barbarine Algeria; BT, Noire de Thibar; DMNA, Algerian D'men; DMNM, Moroccan D'men; FRF, Farafra; HAMA, Hamra; Ho, average observed heterozygosity; MOR, Local Population; OLDA, Ouled–Djellal from Algeria; OLDM, Ouled–Djellal from Morocco; OSSI, Ossimi; SAID, Saidi; SDNA, Sidaoun; SOUH, Souhagi; SRDM, Sardi; SS, Sicilio-Sarde; TMHM, Timahdite; TZGA, Tazegzawth.

[a]Number of individuals typed.

Population structure and genetic relationship analyses

Principal component analysis (PCA) was performed using the dudi.pca function from the ade4 package (Dray & Dufour, 2007). Prior to this step, allele frequencies were scaled using the scaleGen function implemented in the adegenet R package (Jombart, 2008). We ran ADMIXTURE 1.23 (Alexander et al., 2009) to estimate the number of underlying populations through cross-validation for values of K from 3 to 22 (the number of populations). Default input parameters were considered (i.e. quasi-Newton convergence acceleration method and termination criterion of $<10^{-4}$ for the log-likelihood increase between successive iterations). DISTRUCT software (Rosenberg, 2004) was used to produce plots of ancestry proportions for K ancestral populations. The pairwise fixation index (F_{ST}) between populations was estimated using GENEPOP 4.6 (Rousset, 2008).

Identification of selection signatures

Three extended haplotype homozygosity (EHH)-based tests, *Rsb* (Tang et al., 2007), *XP-EHH* (Sabeti et al., 2007) and the integrated haplotype score (*iHS*; Voight et al., 2006), were implemented to detect signatures of selection using the rehh package (Gautier & Vitalis, 2012). In the first two tests, EHH patterns of the same haplotype were contrasted between fat- and thin-tailed sheep populations. For the *Rsb* and *XP-EHH* computation, haplotypes were reconstructed using FASTPHASE 1.4 software (Scheet & Stephens, 2006) with 15 haplotype clusters (K) for 5% of the masked data as being the optimal value inferred by the imputeqc R package (Khvorykh & Khrunin, 2020). Considering a normal distribution of *Rsb* and *XP-EHH* values, a Z-test was applied to identify significant SNPs under selection. Two-sided p-values were derived as $pRsb = -\log_{10}[1 - 2|\Phi(Rsb) - 0.5|]$ and $pXP\text{-}EHH = -\log_{10}[1 - 2|\Phi(XP\text{-}EHH) - 0.5|]$, where Φ

(x) is the Gaussian distribution function. For *iHS* computations, the most common allele in a panel of 170 sheep breeds was considered as the ancestral allele state. The *iHS* scores for each SNP were transformed into two-sided p-values: $piHS = -\log_{10}[1 - 2|\Phi(iHS) - 0.5|]$, where $\Phi(iHS)$ represents the Gaussian distribution function for the *iHS* statistic. For each of the three tests, we used 1.5 Mb sliding windows, in which adjacent windows overlap by 50 kb. Because selective sweeps tend to produce clusters of extreme scores across the sweep region (Voight et al., 2006), windows with six or more SNPs exceeding the threshold of $-\log_{10}$ (p-value) = 2.5 were considered as selection signatures. In order to check if the window size can impact the results of detection, we performed a new analysis setting the window size to 1 Mb.

Gene identification and functional enrichment analysis

The genomic regions putatively under selection (those containing at least six markers exceeding the significance threshold of $-\log_{10}$ (p-value) = 2.5) identified by each EHH-based test were interrogated for genes annotated to the Oar_v4.0 reference genome (International Sheep Genomics Consortium et al., 2010).

Functional enrichment analysis was performed for the genes overlapping the candidate regions that were jointly identified by *Rsb* and *XP-EHH* tests using DAVID (Database for Annotation, Visualization and Integrated Discovery, https://david.ncifcrf.gov/), version 6.8. DAVID uses thousands of annotation terms in several annotation categories, such as molecular function, biological process, cellular component and gene functional summaries. An adjusted Benjamini-corrected p-value of 0.05 was used as the criterion for statistical significance of over-enrichment of genes in each category.

RESULTS

Breeds analyzed in this study

Figure 1 shows the locations of the fat- and thin-tailed breeds used in this study with Tunisia and eastern Algeria separating the thin- and fat-tailed sheep. The two thin-tailed Tunisian breeds (Merino-crossbred BT and SS) show European ancestry. The observed heterozygosities (Ho) of all thin-tailed breeds with the exception of the isolated TZGA, DMNM and the southernmost SDNA, are as high as those observed in several Iberian and Balkan breeds (Ciani et al., 2020; Kijas et al., 2012) and range from 3.7 to 3.9. In contrast, the Ho was lower for the easternmost cluster composed of Egyptian fat-tailed breeds (Ho ranging from 3.3 to 3.7; Table 1).

Population structure analysis

In the PCA (Figure 2a), the percentage of the variance explained by the first two principle components (PC1 and PC2) is lower than that generally seen for livestock breeds, presumably because the genetic cline is weak relative to the genetic development within several breeds as is apparent from the long branch lengths of several breeds (Figure 2b). Nevertheless, PC1 and PC2 show a striking correlation of genetic differentiation and geographical distance. PC1 captures an east–west cline from the Egyptian (SOUH, FRF, SAID, OSSI and BARK), Libyan (BARL) and Tunisian (BART) fat-tailed breeds to the thin-tailed Algerian and Moroccan breeds and then to the two Tunisian composite breeds, BT and SS with a large proportion of Southern European ancestry (Ben-Jemaa et al., 2019). Because the fat- and thin-tailed breeds occur in northeast and northwest Africa, respectively, PC1 also separates the breeds according to their

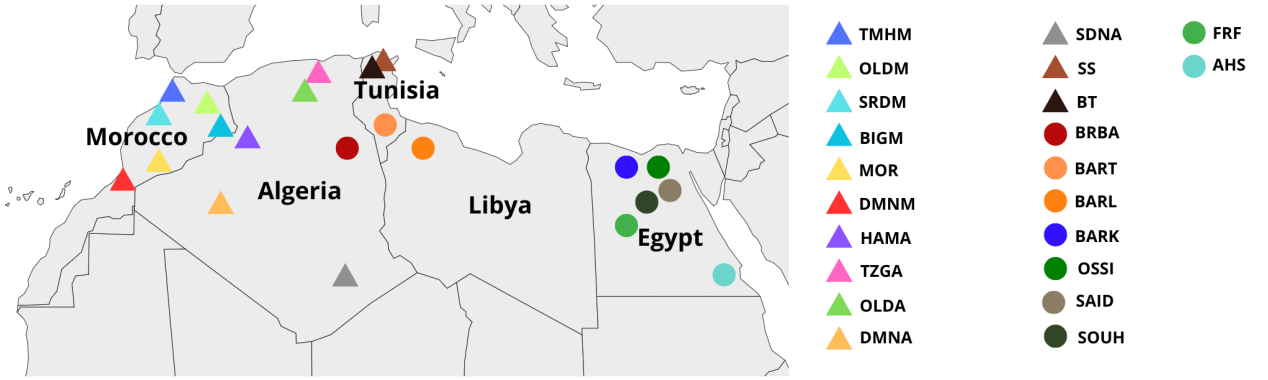

FIGURE 1 Map showing the sampling locations for this study. Locations of thin-tailed sheep are represented by colored triangles. Those of fat-tailed sheep are represented by colored circles. AHS, Aburamad–Halaieb–Shalateen; BARK, Egyptian Barki; BARL, Lybian Barbarine; BART, Tunisian Barbarine; BIGM, Beni–Guil Morocco; BRBA, Barbarine Algeria; BT, Noire de Thibar; DMNA, Algerian D'men; DMNM, Moroccan D'men; FRF, Farafra; HAMA, Hamra; MOR, Local Population; OLDA, Ouled–Djellal from Algeria; OLDM, Ouled–Djellal from Morocco; OSSI, Ossimi; SAID, Saidi; SDNA, Sidaoun; SOUH, Souhagi; SRDM, Sardi; SS, Sicilio-Sarde; TMHM, Timahdite; TZGA, Tazegzawth.

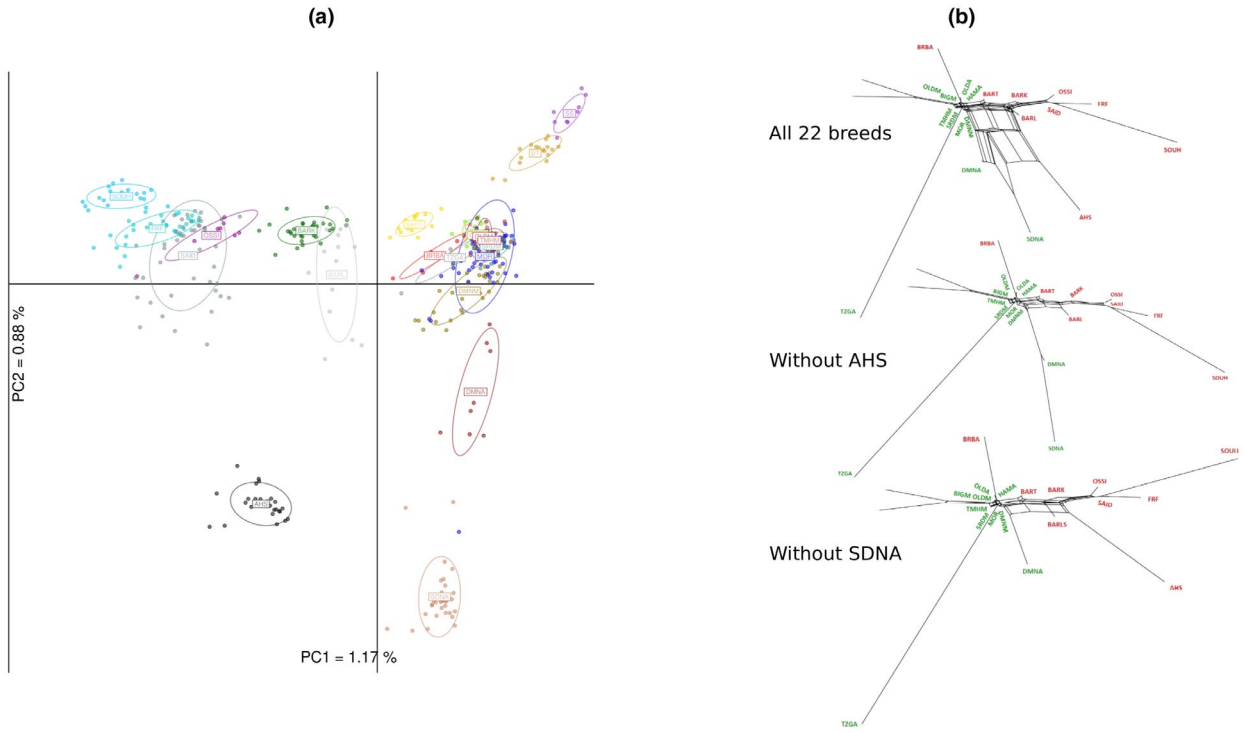

FIGURE 2 Principal component analysis (PCA) and phylogenetic results showing the relationship between North African sheep populations. (a) The PCA results of allele frequencies obtained from 32 613 SNPs genotyped in 462 individuals. (b) The phylogenetic tree obtained from pairwise F_{ST} values using the same number of SNPs as PCA. From top to bottom, three configurations are shown: (i) the phylogenetic tree constructed from all the 22 populations; (ii) the phylogenetic tree constructed when excluding AHS population; and (iii) phylogenetic tree constructed when excluding the Sidaoun population.

tail type (see below). PC2 runs from the south to the southernost breeds SDNA and AHS via DMNA to the thin- and fat-tailed breeds near the Mediterranean coast and then to the breeds with recent European ancestry.

This phylogeographical pattern is confirmed by visualizing F_{ST} genetic distances in a NeighborNet phylogenetic network (Figure 2b), reproducing the east–west cline, which separates thin- and fat-tailed breeds. It also shows the intermediate position of DMNA between SDNA and the breeds near the coast.

As reported previously (Belabdi et al., 2019), there is no genetic differentiation between the six Moroccan breeds, BIGM, DMNM, MOR, OLDM, TMHM and SRDM (mean $F_{ST} \approx 0$), hardly any differentiation of these breeds and the two Algerian thin-tailed breeds (HAMA and OLDA, $F_{ST} < 0.005$), and a genetic distance of these breeds from the fat-tailed Tunisian Barbarine of only ~0.008. The very low F_{ST} value (0.007) that separates the Lybian BARL and the Egyptian BARK confirms the Libyan origin of Barki sheep (Mahrous et al., 2016) (Table S2). Remarkably, the NeighborNet graph (Figure 2b) shows a reticulation that depends on the presence of both SDNA and AHS, suggesting gene flow between these two populations.

The ADMIXTURE results for ancestral populations $K = 3$–7 and $K = 10$ are shown in Figure 3. The cross-validation error suggests that four source populations can model the genetic diversity adequately (Figure S1). The most homogeneous inferred clusters correspond to the isolated TZGA, to the southernmost breeds (SDNA and AHS) and to BT. However, this does not imply that these breeds are ancestral (Lenstra et al., 2012). At $K = 3$, the cross-bred Egyptian-SDNA origin of AHS is confirmed. Likewise, in agreement with the PCA and the NeighborNet graph, the $K = 3$ pattern also shows the east–west genetic structure and an intermediate position of DMNA between SDNA and the other Algerian breeds. However, the ADMIXTURE analysis shows a gradual genetic cline, whereas the tail phenotype shows a clear subdivision.

Signatures of selection

To identify genomic regions putatively under selection in North African fat-tailed sheep, we used the *iHS*, which measures how the EHH decays around the derived allele in relation to the ancestral allele in all the fat-tailed study breeds. Two candidate regions located on OAR01 (at position 18 050 000–20 600 000 bp) and OAR13 (at position 47 450 000–49 650 000 bp) were revealed by the *iHS* test (Figure 4a and Table S3). We also used the *Rsb* and *XP-EHH* tests, to contrast the EHH profiles between 406 fat-tailed and 518 thin-tailed haplotypes. In

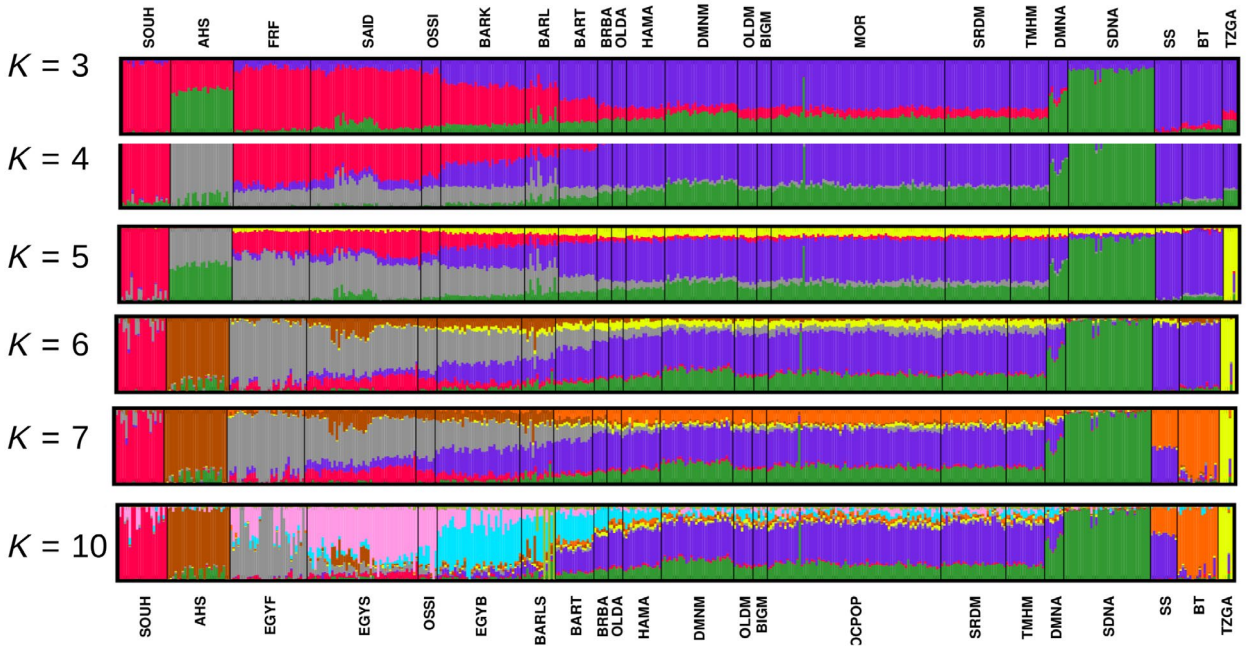

FIGURE 3 ADMIXTURE results of the 462 individuals from the 22 North African sheep populations. Results for K (number of clusters) = 3, to K = 7 and 10 are shown. Individuals are grouped by population. Each individual is represented by a vertical bar. The proportion of the bar in each of k colors corresponds to the average posterior likelihood that the individual is assigned to the cluster indicated by that color. Populations are separated by black lines.

total, with the 1.5 Mb window size (median number of SNPs = 20), 13 and seven outlier windows were detected by the *Rsb* and *XP-EHH* tests, respectively (Figure 4b,c). All but one candidate region (the one located on OAR01; 88 150 000–90 100 000 bp) were also detected with the 1 Mb window size (median number of SNPs = 14; Table S4). All seven candidate regions detected by *XP-EHH* using the 1.5 Mb window size were also revealed by *Rsb*. Six among these seven windows contain at least one marker among the top 1% SNPs with the highest F_{ST} values between fat- and thin-tailed breeds (Figure S2). A further result is that *Rsb* values were positive for the five candidate regions located on OAR01 (two regions), 4, 7 and 13, suggesting that selection in these regions occurred in fat-tailed breeds. The seven candidate regions spanned 326 genes (Table S3), of which 220 are characterized (i.e. with an Ensembl ID). Among these, 190 could be mapped using DAVID Bioinformatics resources. Glutathione *S*-transferase, Mu class, belonging to InterPro protein functional group IPR003081 was the only significantly enriched functional class (Benjamin-corrected p-value < 0.05), which was identified for the annotation cluster 1 (Table S5).

The strongest signals jointly detected by *Rsb* and *XP-EHH* reside within 248 and 287 kb regions located on OAR01 and OAR13, respectively (Figure 4b,c). Interestingly, these two regions overlap with the two outlier windows revealed by *iHS* (Table S3) and display the highest number of SNPs belonging to the top 1% of markers with the highest F_{ST} values between fat- and thin-tailed breeds (Figure S2). The 248 kb signal on OAR01 (19 014 192–19 262 028 bp) includes six highly significant SNPs [10 < log (p-value) < 21 in *Rsb* test] and includes 10 identified genes: *BEST4*, *BTBD19*, *C1H1orf228*, *KIF2C*, *PLK3*, *PTCH2*, *RNF220*, *RPS8*, *TCTEX1D4* and *TMEM53*. The second strong signal on OAR13 (48 193 040–48 480 264 bp) contains five highly significant SNPs (11 < log (p-value) < 25 in *Rsb* test) and includes only the *BMP2* gene. Other selection signal peaks (and associated candidate genes) jointly detected by *Rsb* and *XP-EHH* include the following regions: OAR01, 86 313 617–86 352 115 bp (*AHCYL1*); OAR04, 68 649 197–69 367 542 bp (*EVX1*, *HOXA13*, *HOXA11*, *HOXA10*, *HOXA9*, *HOXA3*, *HOXA5*, *HOXA4*, *HOXA2*, *HOXA1*, *SKAP2*); OAR07, 55 640 042–55 863 181 bp (*SCG3* and *DMXL2*); OAR12, 53 402 994–54 123 627 bp (*RABGAP1L* and *GPR52*); and OAR24, 27 407 352–27 524 277 bp (*ITGAX*, *ITGAD*, *ARMC5*, *TGFB1I1*, *SLC5A2* and *C24H16orf58*).

DISCUSSION

The past decades have witnessed the development of numerous statistical methods to detect genome-wide selective sweeps. These methods capture various modes of selection acting at different time scales, thus allowing the exploration of different intervals of a population's history. Several studies on the genomic architecture of sheep tail phenotypes have been conducted, mostly including breeds from Eurasia (e.g. Dong et al., 2020; Mastrangelo et al., 2019; Moioli et al., 2015; Pan

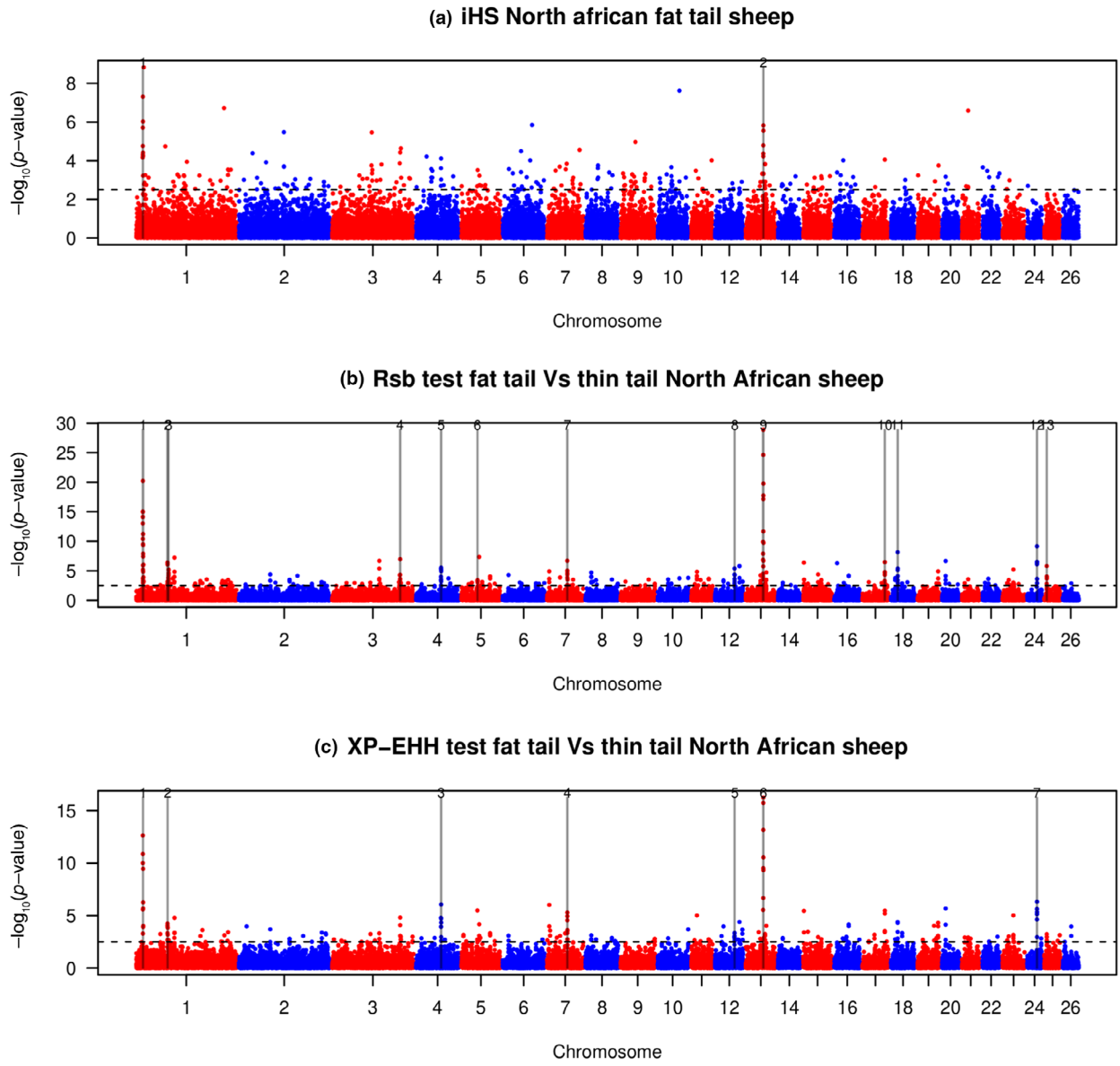

FIGURE 4 Manhattan plots showing the results of extended haplotype homozygosity-based tests. (a) The integrated haplotype score (*iHS*) test computed within fat-tailed breeds. (b) The *Rsb* test fat- vs. thin-tailed breeds (c) The *XP-EHH* test fat- vs. thin-tailed breeds. Horizontal dashed lines mark the significance threshold applied to detect the outlier SNPs ($-\log_{10}$ (*p*-value)=2.5).

et al., 2019) and Eastern Africa (e.g. Ahbara et al., 2019; Mwacharo et al., 2017). Only a few large-scale genomic studies have been performed to identify genes associated with tail phenotype in North African sheep (Fonseca et al., 2024). Baazaoui et al. (2024) relied on differences in allele frequencies to identify SNPs putatively under differential selection between the Tunisian Barbarine and six other thin-tailed sheep breeds. Recently, we used a subset of the present data to identify selection signatures in two Algerian (Hamra and Ouled Djellal) and two Tunisian breeds (Barbarine and Noire de Thibar) (Yahyaoui et al., 2024). In the present study, we used a larger panel of populations across a broad geographic region in North Africa and employed LD-based methods to identify genomic regions under divergent selection between fat- and thin-tailed breeds. Our choice of LD methods is motivated by the fact that they have optimal detection rates in a range from low beneficial allele frequency up to close to fixation (Weigand & Leese, 2018). Precisely, we implemented two complementary approaches, *Rsb* and *XP-EHH*, which are more sensitive to selective sweeps in which the corresponding allele has approached or achieved fixation (Sabeti et al., 2007; Tang et al., 2007), and the *iHS* test, which has more power to detect incomplete sweeps (Voight et al., 2006). We identified seven relevant genomic regions putatively under divergent selection between nine fat-tailed and 13 thin-tailed North African sheep breeds with different geographical origins using these three EHH-based tests. The

geographical diversity of the sampled breeds of both types makes it unlikely that the identified selection footprints result from random demographic changes. Rather, we expect that most of our candidate regions are true targets of divergent selection between North African fat- and thin-tailed sheep populations. Our analysis of population structure showed that despite the fact that most of the study breeds could be split into two main groupings, where the first one includes the Moroccan–Algerian thin-tailed cluster and the second grouping is composed of Egyptian fat-tailed populations, several other breeds have different sets of ancestral populations (such as SDNA, TZGA from the thin-tailed group and AHS from the fat-tailed group), thus reflecting different population history. A further result is that although the Moroccan–Algerian thin-tailed cluster and the fat-tailed Algerian and Tunisian Barbarine show clear morphological differences and are expected to have different ancestral components, these two groups display similar patterns of ancestry and low genetic differentiation. This result is consistent with previous studies that reported a large-scale genetic homogenization of sheep breeds within and between neighboring North African countries regardless of their tail type (Belabdi et al., 2019; Ben-Jemaa et al., 2019; Gaouar et al., 2017). Genetic similarity between fat-tailed and thin-tailed breeds from the same geographic area has been also reported in Iran and Italy (Ciani et al., 2020; Mastrangelo et al., 2019). Based on this observation, it has been suggested that it is likely that only a limited number of genes is involved in fat-tail deposition in sheep (Moradi et al., 2022). Consistently, examining genes located near selection signal peaks of the candidate regions jointly detected by *Rsb* and *XP-EHH* revealed few genes that are involved in adipogenesis. Among the candidate regions, the 248 and 287 kb windows located on OAR01 and OAR13, respectively, may be the most biologically relevant ones, owing to their overlap with results from the *iHS* test and extreme peaks (in all three EHH-based tests and in locus-specific F_{ST}). The region on OAR13 spanned *BMP2* gene, previously identified by several other studies (Ahbara et al., 2019; Lu et al., 2020; Mastrangelo et al., 2019; Zhao et al., 2020). By performing a genome scan using whole genome sequence data from a few dozen Chinese sheep, Pan et al. (2019) demonstrated that fixation of fat tails in domestic sheep is associated with a selective sweep near a retro-transposable hotspot on OAR13, which specifically affects the expression of *BMP2*.

PTCH2 is the second interesting candidate associated with fat deposition, possibly under divergent selection between fat- and thin-tailed North African sheep. *PTCH2* was identified by examining the nearest gene to the strong signal peak on OAR01 (19 014 192–19 262 028 bp). This gene is the receptor for Hedgehog signaling which exhibits high activity in the *Drosophila* fat body. Moreover, fasting conditions increased the level of Hedgehog protein in the fat body (Zhang et al., 2020) which may suggest that *PTCH2* has play an essential role in adipogenesis regulation in North African fat-tailed sheep during extended periods of food scarcity (through Hedgehog signal transduction). At the same time, *PTCH2* is found at high levels in the skin and spermatocytes (Carpenter et al., 1998) and controls the growth and morphogenesis of hair follicle epithelium in mice (Nieuwenhuis et al., 2006). The dual role played by *PTCH2* in both adipogenesis regulation and hair characteristics (linked to thermal adaptability) suggests that selection occurred on genes involved in adaptability to multiple environmental stressors in North African sheep.

Our results also indicate that another gene, *SCG3*, on OAR07, may be involved in fat storage in fat-tailed North African sheep. *SCG3* is located only 35 kb downstream the selection signal peak (SNP rs403325628 at position 55 676 739 bp, $-\log_{10}$ (p-value) = 6.71). Mutations in this gene are associated with obesity in humans (Tanabe et al., 2007) while Derks et al. (2021) found a missense mutation in *SCG3* that is likely to affect backfat and growth rate in pig.

Notably, genes located in our candidate regions were particularly enriched for a gene family, the glutathione *S*-transferases (GSTs) found in peroxisomes of most organisms from fungi to animals (Ferreira et al., 2023) and mostly known to catalyze fatty acid conjugation (Listowsky et al., 1988). In addition to the conjugation reactions, GSTs have a role in several other catalytic functions such as the catalysis of the reduction of phospholipids, fatty acids and DNA hydroperoxides produced by lipid peroxidation and oxidative damage to DNA (Hayes et al., 2005). Also, it has been shown that *GSTA4* (a class of GSTs) is downregulated in adipose tissue of obese mice and in insulin-resistant humans (Curtis et al., 2010). Genes with potential functions in fatty acid oxidation have been found to be differentially expressed between fat- and thin-tailed sheep from Iran (Bakhtiarizadeh et al., 2019) and China (Chao et al., 2017).

A further result is the identification of two statistically significant SNPs on OAR04 near the *HOXA* gene cluster (*HOXA13*, *HOXA11*, *HOXA10*, *HOXA9*, *HOXA3*, *HOXA5*, *HOXA4*, *HOXA2* and *HOXA1*). The Hox gene family has 39 members and can be divided into 13 paralogous groups (Mallo et al., 2010). These genes are key players in regulating morphologies along the anteroposterior axis of mammalian embryos (Kostic & Capecchi, 1994). Importantly, the *HOXA11* paralogous group regulates the development of the sacrum (Mallo et al., 2010) and is required for proper patterning of the first few caudal vertebrae (Wellik & Capecchi, 2003). The *HoxA* gene cluster was shown to be under strong selection in Mongolian (Wang et al., 2019) and South West Asian sheep (Fariello et al., 2014) and goats (Bertolini et al., 2018).

Because the fat-tail phenotype has been proposed to have arisen several millennia ago in arid regions of Asia

(Dong et al., 2020), it is tempting to speculate that some of the genes found to be under divergent selection between fat- and thin-tailed sheep are associated with adaptive processes relating to multiple stressors prevailing in desert-like environments. Accordingly, in addition to the previously mentioned *PTCH2* gene, the selection signal on OAR01 includes two other genes, *PLK3* and *RNF220*, directly related to thermal stress tolerance. *PLK3* is a key microRNA target in regulating skin hair follicle progenitors' growth competency (Liu et al., 2021) while *RNF220* has been associated with heart rate recovery in humans (Van de Vegte et al., 2019; Verweij et al., 2018). It has previously been shown that maximum heart rate and its rate-limiting temperature during acute warming are significantly higher in fish that have evolved in a desert climate compared with a montane climate (Chen et al., 2018). The presence of selective pressure where these two genes are located could be indicative of a region involved in heat stress response in North African fat-tailed sheep.

The relatively short history and the common origin of fat-tailed sheep (Dong et al., 2020) is expected to lead to several common candidate regions between studies aiming to detect selective sweeps between fat- and thin-tailed breeds. Accordingly, many of our candidate regions were also identified by other studies. This is particularly true for the two candidate regions showing the strongest selection pressure on OAR01 (18 400 000–20 750 000 bp) and OAR13 (47 450 000–49 650 000 bp). The genomic region on OAR01 overlaps with a candidate region found to be highly differentiated between Ethiopian fat-rump and Sudanese thin-tail sheep (Ahbara et al., 2019) and contains one SNP (OAR1_100921673.1) among seven others, that were shown to be highly diverged between 18 fat- and 14 thin-tailed breeds from around the world (Dong et al., 2020). Likewise, it has been strongly highlighted that our candidate region on OAR13 (including the *BMP2* gene) is under divergent selection between fat-tailed and thin-tailed sheep worldwide (Ahbara et al., 2019; Dong et al., 2020; Li et al., 2020; Mastrangelo et al., 2019; Moioli et al., 2015; Zhao et al., 2017, 2020). Our results further support previously reported selection signatures on OAR04 (Dong et al., 2020; Li et al., 2020; Yuan et al., 2017; Zhao et al., 2017, 2020) and OAR07 (Li et al., 2020; Lv et al., 2014).

It is worth mentioning that the number of SNPed in the present study resulted in moderate marker coverage, which may have led to reduced information for some relevant genomic regions harboring selection signatures. Clearly, the use of SNP chips with higher marker density or whole genome sequencing would help in detecting additional selection signatures at higher resolution. Because selective sweeps produce clusters of candidate SNPs in the vicinity of selection targets (Voight et al., 2006), a common approach to limit the number of false-positives is to consider only windows harboring a minimum number of SNPs exceeding a predefined threshold as selection signatures. In the present study, we set this minimum number to six, which is higher than in other studies using a similar SNP density (e.g. Ben-Jemaa et al., 2020; Gautier & Naves, 2011; Lukic et al., 2023; Mastrangelo et al., 2023; Persichilli et al., 2023). By using such a conservative criterion in the definition of selection signatures it is likely that some selection signals will be missed. This is particularly true in high recombining regions where hitchhiking effects of selection are limited.

In conclusion, we identified seven relevant genomic regions under differential selection between nine fat-tailed and 13 thin-tailed sheep populations originating from diverse geographical regions in North Africa. We found that selection mainly occurred in fat-tailed breeds and our results support a direct link between *BMP2* and the fat-tail phenotype in sheep. Our results give also novel insights into other promising candidate genes associated with fat deposition as well as other features related to heat stress adaptation.

AUTHOR CONTRIBUTIONS

Slim Ben-Jemaa: Conceptualization; formal analysis; funding acquisition; investigation; project administration; supervision; writing – original draft. **Ghazi Yahyaoui:** Formal analysis; writing – original draft. **Samia Kdidi:** Data curation; writing – review and editing. **Afef Najjari:** Supervision. **Johannes A. Lenstra:** Formal analysis; investigation; writing – original draft. **Salvatore Mastrangelo:** Data curation; writing – review and editing. **Semir B. S. Gaouar:** Writing – review and editing. **Joram M. Mwacharo:** Writing – review and editing. **Touhami Khorchani:** Funding acquisition; resources. **Mohamed H. Yahyaoui:** Resources; writing – review and editing.

ACKNOWLEDGEMENTS

We acknowledge IFS grant B/5478-2 from the International Foundation for Science and we thank all donors to the CGIAR Trust Fund which globally support its Small Ruminant Genomics work through the SAPLING Initiative.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All genotyping data used in the present study is publicly available.

ORCID

Slim Ben-Jemaa https://orcid.org/0000-0002-7103-3315
Johannes A. Lenstra https://orcid.org/0000-0003-1745-7739
Salvatore Mastrangelo https://orcid.org/0000-0001-6511-1981

REFERENCES

Abdelkader, A.A., Ata, N., Benyoucef, M.T., Djaout, A., Azzi, N., Yilmaz, O. et al. (2017) New genetic identification and characterisation of 12 Algerian sheep breeds by microsatellite markers. *Italian Journal of Animal Science*, 17, 38–48. Available from: https://doi.org/10.1080/1828051X.2017.1335182

Ahbara, A., Bahbahani, H., Almathen, F., Al, A.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. *Frontiers in Genetics*, 9, 699. Available from: https://doi.org/10.3389/fgene.2018.00699

Alexander, D.H., Novembre, J. & Lange, K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. Available from: https://doi.org/10.1101/gr.094052.109

Baazaoui, I., Bedhiaf-Romdhani, S., Mastrangelo, S. & Ciani, E. (2021) Genome-wide analyses reveal population structure and identify candidate genes associated with tail fatness in local sheep from a semi-arid area. *Animal*, 15, 100193. Available from: https://doi.org/10.1016/j.animal.2021.100193

Baazaoui, I., Bedhiaf-Romdhani, S., Mastrangelo, S., Lenstra, J.A., Da Silva, A., Benjelloun, B. et al. (2024) Refining the genomic profiles of north African sheep breeds through meta-analysis of worldwide genomic SNP data. *Frontiers in Veterinary Science*, 11, 1339321. Available from: https://doi.org/10.3389/fvets.2024.1339321

Bakhtiarizadeh, M.R., Salehi, A., Alamouti, A.A., Abdollahi-Arpanahi, R. & Salami, S.A. (2019) Deep transcriptome analysis using RNA-seq suggests novel insights into molecular aspects of fat-tail metabolism in sheep. *Scientific Reports*, 9, 9203. Available from: https://doi.org/10.1038/s41598-019-45665-3

Bedhiaf-Romdhani, S., Baazaoui, I., Dodds, K.G., Brauning, R., Anderson, R.M., Van Stijn, T.C. et al. (2023) Efficiency of genotyping by sequencing in inferring genomic relatedness and molecular insights into fat tail selection in Tunisian sheep. *Animal Genetics*, 54, 389–397. Available from: https://doi.org/10.1111/age.13296

Belabdi, I., Ouhrouch, A., Lafri, M., Gaouar, S.B.S., Ciani, E., Benali, A.R. et al. (2019) Genetic homogenization of indigenous sheep breeds in Northwest Africa. *Scientific Reports*, 9, 7920. Available from: https://doi.org/10.1038/s41598-019-44137-y

Ben-Jemaa, S., Kdidi, S., Gdura, A.M., Dayhum, A.S., Eldaghayes, I.M., Boussaha, M. et al. (2019) Inferring the population structure of the Maghreb sheep breeds using a medium-density SNP chip. *Animal Genetics*, 50, 526–533. Available from: https://doi.org/10.1111/age.12831

Ben-Jemaa, S., Mastrangelo, S., Lee, S.-H., Lee, J.H. & Boussaha, M. (2020) Genome-wide scan for selection signatures reveals novel insights into the adaptive capacity in local North African cattle. *Scientific Reports*, 10, 19466. Available from: https://doi.org/10.1038/s41598-020-76576-3

Bertolini, F., Servin, B., Talenti, A., Rochat, E., Kim, E.S., Oget, C. et al. (2018) Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genetics Selection Evolution*, 50, 57. Available from: https://doi.org/10.1186/s12711-018-0421-y

Carpenter, D., Stone, D.M., Brush, J., Ryan, A., Armanini, M., Frantz, G. et al. (1998) Characterization of two patched receptors for the vertebrate hedgehog protein family. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 13630–13634. Available from: https://doi.org/10.1073/pnas.95.23.13630

Chao, T., Wang, G., Ji, Z., Liu, Z., Hou, L., Jin, W. et al. (2017) Transcriptome analysis of three sheep intestinal regions reveals key pathways and hub regulatory genes of large intestinal lipid metabolism. *Scientific Reports*, 7, 5345. Available from: https://doi.org/10.1038/s41598-017-05551-2

Chen, Z., Farrell, A.P., Matala, A., Hoffman, N. & Narum, S.R. (2018) Physiological and genomic signatures of evolutionary thermal adaptation in redband trout from extreme climates. *Evolutionary Applications*, 11, 1686–1699. Available from: https://doi.org/10.1111/eva.12672

Ciani, E., Mastrangelo, S., Da Silva, A., Marroni, F., Ferenčaković, M., Ajmone-Marsan, P. et al. (2020) On the origin of European sheep as revealed by the diversity of the Balkan breeds and by optimizing population-genetic analysis tools. *Genetics Selection Evolution*, 52, 25. Available from: https://doi.org/10.1186/s12711-020-00545-7

Curtis, J.M., Grimsrud, P.A., Wright, W.S., Xu, X., Foncea, R.E., Graham, D.W. et al. (2010) Downregulation of adipose glutathione S-transferase A4 leads to increased protein carbonylation, oxidative stress, and mitochondrial dysfunction. *Diabetes*, 59, 1132–1142. Available from: https://doi.org/10.2337/db09-1105

Derks, M.F.L., Gross, C., Lopes, M.S., Reinders, M.J.T., Bosse, M., Gjuvsland, A.B. et al. (2021) Accelerated discovery of functional genomic variation in pigs. *Genomics*, 113, 2229–2239. Available from: https://doi.org/10.1016/j.ygeno.2021.05.017

Dong, K., Yang, M., Han, J., Ma, Q., Han, J., Song, Z. et al. (2020) Genomic analysis of worldwide sheep breeds reveals *PDGFD* as a major target of fat-tail selection in sheep. *BMC Genomics*, 21, 800. Available from: https://doi.org/10.1186/s12864-020-07210-9

Dray, S. & Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22, 1–20. Available from: https://doi.org/10.18637/jss.v022.i04

Dutilly-Diane, C. (2007) Pastoral economics and marketing in North Africa: a literature review. *Nomadic Peoples*, 11, 69–90.

Fariello, M.-I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., Consortium, I.S.G. et al. (2014) Selection signatures in worldwide sheep populations. *PLoS One*, 9, e103813. Available from: https://doi.org/10.1371/journal.pone.0103813

Ferreira, M.J., Rodrigues, T.A., Pedrosa, A.G., Silva, A.R., Vilarinho, B.G., Francisco, T. et al. (2023) Glutathione and peroxisome redox homeostasis. *Redox Biology*, 67, 102917. Available from: https://doi.org/10.1016/j.redox.2023.102917

Fonseca, P.A.S., Suárez-Vega, A., Arranz, J.J. & Gutiérrez-Gil, B. (2024) Integration of selective sweeps across the sheep genome: understanding the relationship between production and adaptation traits. *Genetics Selection Evolution*, 56, 40. Available from: https://doi.org/10.1186/s12711-024-00910-w

Food and Agriculture Organization of the United Nations. (2019) *Crops and livestock products*. Available from: https://www.fao.org/faostat/en/#data/QCL [Accessed 10th September 2021].

Gaouar, S.B.S. (2009) *Biodiversity study: analysis of the genetic variability of Algerian sheep breeds and their phylogenetic relationships using microsatellites – Etude de la biodiversité: Analyse de la variabilité génétique des races ovines algériennes & de leurs relations phylogénétiques par l'utilisation de microsatellites* [Doctoral dissertation, Institute of Natural Sciences, University of Essénia, Oran, Algeria]. https://doi.org/10.13140/RG.2.1.1997.5204

Gaouar, S.B.S., Lafri, M., Djaout, A., El-Bouyahiaoui, R., Bouri, A., Bouchatal, A. et al. (2017) Genome-wide analysis highlights genetic dilution in Algerian sheep. *Heredity*, 118, 293–301. Available from: https://doi.org/10.1038/hdy.2016.86

Gaouar, S.B.S., Silva, A.D., Ciani, E., Kdidi, S., Aouissat, M., Dhimi, L. et al. (2015) Admixture and local breed marginalization threaten Algerian sheep diversity. *PLoS One*, 10, e0122667. Available from: https://doi.org/10.1371/journal.pone.0122667

Gautier, M. & Naves, M. (2011) Footprints of selection in the ancestral admixture of a new world creole cattle breed. *Molecular Ecology*, 20, 3128–3143. Available from: https://doi.org/10.1111/j.1365-294X.2011.05163.x

Gautier, M. & Vitalis, R. (2012) Rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, 28, 1176–1177. Available from: https://doi.org/10.1093/bioinformatics/bts115

Hayes, J.D., Flanagan, J.U. & Jowsey, I.R. (2005) Glutathione transferases. *Annual Review of Pharmacology and Toxicology*, 45, 51–88. Available from: https://doi.org/10.1146/annurev.pharmtox.45.120403.095857

International Sheep Genomics Consortium, Archibald, A.L., Cockett, N.E., Dalrymple, B.P., Faraut, T., Kijas, J.W. et al.

(2010) The sheep genome reference sequence: a work in progress. *Animal Genetics*, 41, 449–453. Available from: https://doi.org/10.1111/j.1365-2052.2010.02100.x

Jombart, T. (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. Available from: https://doi.org/10.1093/bioinformatics/btn129

Kandoussi, A., Boujenane, I., Auger, C., Serranito, B., Germot, A., Piro, M. et al. (2020) The origin of sheep settlement in Western Mediterranean. *Scientific Reports*, 10, 10225. Available from: https://doi.org/10.1038/s41598-020-67246-5

Khvorykh, G.V. & Khrunin, A.V. (2020) Imputeqc: an R package for assessing imputation quality of genotypes and optimizing imputation parameters. *BMC Bioinformatics*, 21, 304. Available from: https://doi.org/10.1186/s12859-020-03589-0

Kijas, J.W., Lenstra, J.A., Hayes, B., Boitard, S., Porto Neto, L.R., San, C.M. et al. (2012) Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, 10, e1001258. Available from: https://doi.org/10.1371/journal.pbio.1001258

Kostic, D. & Capecchi, M.R. (1994) Targeted disruptions of the murine *Hoxa-4* and *Hoxa-6* genes result in homeotic transformations of components of the vertebral column. *Mechanisms of Development*, 46, 231–247. Available from: https://doi.org/10.1016/0925-4773(94)90073-6

Lenstra, J.A., Groeneveld, L.F., Eding, H., Kantanen, J., Williams, J.L., Taberlet, P. et al. (2012) Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Animal Genetics*, 43, 483–502. Available from: https://doi.org/10.1111/j.1365-2052.2011.02309.x

Li, X., Yang, J., Shen, M., Xie, X.-L., Liu, G.-J., Xu, Y.-X. et al. (2020) Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. *Nature Communications*, 11, 2815. Available from: https://doi.org/10.1038/s41467-020-16485-1

Listowsky, I., Abramovitz, M., Homma, H. & Niitsu, Y. (1988) Intracellular binding and transport of hormones and xenobiotics by glutathiones-transferases. *Drug Metabolism Reviews*, 19, 3–4. Available from: https://doi.org/10.3109/03602538808994138

Liu, F., Zhang, X., Peng, Y., Zhang, L., Yu, Y., Hua, P. et al. (2021) *miR-24* controls the regenerative competence of hair follicle progenitors by targeting *Plk3*. *Cell Reports*, 35, 109225. Available from: https://doi.org/10.1016/j.celrep.2021.109225

Lu, Z., Liu, J., Han, J. & Yang, B. (2020) Association between *BMP2* functional polymorphisms and sheep tail type. *Animals*, 10, 739. Available from: https://doi.org/10.3390/ani10040739

Lukic, B., Curik, I., Drzaic, I., Galić, V., Shihabi, M., Vostry, L. et al. (2023) Genomic signatures of selection, local adaptation and production type characterisation of East Adriatic sheep breeds. *Journal of Animal Science and Biotechnology*, 14, 142. Available from: https://doi.org/10.1186/s40104-023-00936-y

Lv, F.-H., Agha, S., Kantanen, J., Colli, L., Stucki, S., Kijas, J.W. et al. (2014) Adaptations to climate-mediated selective pressures in sheep. *Molecular Biology and Evolution*, 31, 3324–3343. Available from: https://doi.org/10.1093/molbev/msu264

Mahrous, K.F., Hassanane, M.S., Shafey, H.I., Abdel, M.M. & Rushdi, H.E. (2016) Association between single nucleotide polymorphism in ovine calpain gene and growth performance in three Egyptian sheep breeds. *Journal of Genetic Engineering and Biotechnology*, 14, 233–240. Available from: https://doi.org/10.1016/j.jgeb.2016.09.003

Mallo, M., Wellik, D.M. & Deschamps, J. (2010) *Hox* genes and regional patterning of the vertebrate body plan. *Developmental Biology*, 344, 7–15. Available from: https://doi.org/10.1016/j.ydbio.2010.04.024

Mastrangelo, S., Bahbahani, H., Moioli, B., Ahbara, A., Abri, M.A., Almathen, F. et al. (2019) Novel and known signals of selection for fat deposition in domestic sheep breeds from Africa and Eurasia. *PLoS One*, 14, e0209632. Available from: https://doi.org/10.1371/journal.pone.0209632

Mastrangelo, S., Ben-Jemaa, S., Perini, F., Cendron, F., Biscarini, F., Lasagna, E. et al. (2023) Genome-wide mapping of signatures of selection using a high-density array identified candidate genes for growth traits and local adaptation in chickens. *Genetics Selection Evolution*, 55, 20. Available from: https://doi.org/10.1186/s12711-023-00790-6

Moioli, B., Pilla, F. & Ciani, E. (2015) Signatures of selection identify loci associated with fat tail in sheep. *Journal of Animal Science*, 93, 4660–4669. Available from: https://doi.org/10.2527/jas.2015-9389

Moradi, M.H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K.G., Brauning, R. & McEwan, J.C. (2022) Fine mapping of candidate regions associated with fat deposition in thin and fat tail sheep breeds suggests new insights into molecular aspects of fat tail selection. *Animals*, 12, 1423. Available from: https://doi.org/10.3390/ani12111423

Moradi, M.H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K.G. & McEwan, J.C. (2012) Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, 13, 10. Available from: https://doi.org/10.1186/1471-2156-13-10

Muigai, A. & Hanotte, O. (2013) The origin of African sheep: archaeological and genetic perspectives. *African Archaeological Review*, 30, 39–50. Available from: https://doi.org/10.1007/s10437-013-9129-0

Mwacharo, J.M., Kim, E.-S., Elbeltagy, A.R., Aboul-Naga, A.M., Rischkowsky, B.A. & 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 Reports*, 7, 1–10. Available from: https://doi.org/10.1038/s41598-017-17775-3

Nieuwenhuis, E., Motoyama, J., Barnfield, P.C., Yoshikawa, Y., Zhang, X., Mo, R. et al. (2006) Mice with a targeted mutation of *Patched2* are viable but develop alopecia and epidermal hyperplasia. *Molecular and Cellular Biology*, 26, 6609–6622. Available from: https://doi.org/10.1128/MCB.00295-06

Pan, Z., Li, S., Liu, Q., Wang, Z., Zhou, Z., Di, R. et al. (2019) Rapid evolution of a retro-transposable hotspot of ovine genome underlies the alteration of *BMP2* expression and development of fat tails. *BMC Genomics*, 20, 261. Available from: https://doi.org/10.1186/s12864-019-5620-6

Persichilli, C., Senczuk, G., Mastrangelo, S., Marusi, M., van Kaam, J.-T., Finocchiaro, R. et al. (2023) Exploring genome-wide differentiation and signatures of selection in Italian and north American Holstein populations. *Journal of Dairy Science*, 106, 5537–5553. Available from: https://doi.org/10.3168/jds.2022-22159

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575. Available from: https://doi.org/10.1086/519795

Rocha, J.L., Godinho, R., Brito, J.C. & Nielsen, R. (2021) Life in deserts: the genetic basis of mammalian desert adaptation. *Trends in Ecology & Evolution*, 36, 637–650. Available from: https://doi.org/10.1016/j.tree.2021.03.007

Rosenberg, N.A. (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137–138. Available from: https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rousset, F. (2008) genepop'007: a complete re-implementation of the genepop software for windows and Linux. *Molecular Ecology Resources*, 8, 103–106. Available from: https://doi.org/10.1111/j.1471-8286.2007.01931.x

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C. et al. (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449, 913–918. Available from: https://doi.org/10.1038/nature06250

Scheet, P. & Stephens, M. (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics*, 78, 629–644. Available from: https://doi.org/10.1086/502802

Tanabe, A., Yanagiya, T., Iida, A., Saito, S., Sekine, A., Takahashi, A. et al. (2007) Functional single-nucleotide polymorphisms in the secretogranin III (*SCG3*) gene that form secretory granules with appetite-related neuropeptides are associated with obesity. *The Journal of Clinical Endocrinology and Metabolism*, 92, 1145–1154. Available from: https://doi.org/10.1210/jc.2006-1808

Tang, K., Thornton, K.R. & Stoneking, M. (2007) A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, 5, e171. Available from: https://doi.org/10.1371/journal.pbio.0050171

Van de Vegte, Y.J., Tegegne, B.S., Verweij, N., Snieder, H. & van der Harst, P. (2019) Genetics and the heart rate response to exercise. *Cellular and Molecular Life Sciences*, 76, 2391–2409. Available from: https://doi.org/10.1007/s00018-019-03079-4

Verweij, N., van de Vegte, Y.J. & van der Harst, P. (2018) Genetic study links components of the autonomous nervous system to heart-rate profile during exercise. *Nature Communications*, 9, 898. Available from: https://doi.org/10.1038/s41467-018-03395-6

Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. (2006) A map of recent positive selection in the human genome. *PLoS Biology*, 4, e72. Available from: https://doi.org/10.1371/journal.pbio.0040072

Wang, W., Zhang, X., Zhou, X., Zhang, Y., La, Y., Zhang, Y. et al. (2019) Deep genome resequencing reveals artificial and natural selection for visual deterioration, plateau adaptability and high prolificacy in Chinese domestic sheep. *Frontiers in Genetics*, 10, 300. Available from: https://doi.org/10.3389/fgene.2019.00300

Weigand, H. & Leese, F. (2018) Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society*, 184, 528–583. Available from: https://doi.org/10.1093/zoolinnean/zly007

Wellik, D.M. & Capecchi, M.R. (2003) *Hox10* and *Hox11* genes are required to globally pattern the mammalian skeleton. *Science*, 301, 363–367. Available from: https://doi.org/10.1126/science.1085672

Xu, Y.-X., Wang, B., Jing, J.-N., Ma, R., Luo, Y.-H., Li, X. et al. (2023) Whole-body adipose tissue multi-omic analyses in sheep reveal molecular mechanisms underlying local adaptation to extreme environments. *Communications Biology*, 6, 1–15. Available from: https://doi.org/10.1038/s42003-023-04523-9

Yahyaoui, G., Ben-Jemaa, S., Kdidi, S., Gaouar, S.B.S. & Yahyaoui, M.H. (2024) Identification of selection signatures in Algero-Tunisian sheep breeds using medium-density SNP chips. *Genetics & Biodiversity Journal*, 8, 59–75. Available from: https://doi.org/10.46325/gabj.v8i1.390

Yuan, Z., Liu, E., Liu, Z., Kijas, J.W., Zhu, C., Hu, S. et al. (2017) Selection signature analysis reveals genes associated with tail type in Chinese indigenous sheep. *Animal Genetics*, 48, 55–66. Available from: https://doi.org/10.1111/age.12477

Zhang, J., Liu, Y., Jiang, K. & Jia, J. (2020) Hedgehog signaling promotes lipolysis in adipose tissue through directly regulating *Bmm/ATGL lipase*. *Developmental Biology*, 457, 128–139. Available from: https://doi.org/10.1016/j.ydbio.2019.09.009

Zhao, F., Deng, T., Shi, L., Wang, W., Zhang, Q., Du, L. et al. (2020) Genomic scan for selection signature reveals fat deposition in Chinese indigenous sheep with extreme tail types. *Animals*, 10, 773. Available from: https://doi.org/10.3390/ani10050773

Zhao, Y.-X., Yang, J., Lv, F.-H., Hu, X.-J., Xie, X.-L., Zhang, M. et al. (2017) Genomic reconstruction of the history of native sheep reveals the peopling patterns of nomads and the expansion of early pastoralism in East Asia. *Molecular Biology and Evolution*, 34, 2380–2395. Available from: https://doi.org/10.1093/molbev/msx181

Zhu, C., Li, N., Cheng, H. & Ma, Y. (2021) Genome wide association study for the identification of genes associated with tail fat deposition in Chinese sheep breeds. *Biology Open*, 10, bio054932. Available from: https://doi.org/10.1242/bio.054932

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ben-Jemaa, S., Yahyaoui, G., Kdidi, S., Najjari, A., Lenstra, J.A., Mastrangelo, S. et al. (2025) Genome-wide scans for signatures of selection in North African sheep reveals differentially selected regions between fat- and thin-tailed breeds. *Animal Genetics*, 56, e13487. Available from: https://doi.org/10.1111/age.13487