scientific reports

Check for updates

OPEN

Genome-wide identification of selection signatures across altitudinal gradients in dairy sheep breeds

Slim Ben Jemaa[1,2], Salvatore Mastrangelo[1]✉, Federica Carta[1], Silvia Riggio[1], Corrado Dimauro[3], Christian Persichilli[4], Baldassare Portolano[1], Gabriele Senczuk[4] & Alberto Cesarani[3,5]

Deciphering the genetic mechanisms of adaptation in livestock species is important in ensuring sustainability of livestock production in the context of climatic change. Here, we aimed to identify signatures of selection within Sarda and Valle del Belice, the two major Italian dairy sheep breeds, as well as between their ecotypes raised at three different altitudinal gradients: plain, hill, and mountain. By joinly analyzing the three ecotypes from each breed, we found strong signatures of selection at the *KDM6A* gene (X chromosome), a key epigenetic regulator involved in modulating gene expression in response to environmental stress. In both breeds, the mountain ecotype showed a strong selection toward several pleiotropic genes with effects linked to the nervous system, body size, energy balance and muscular function and development. Our study also provides evidence that genes involved in neuroendocrine mechanisms are targets of environmental adaptation in the mountain ecotypes, suggesting their fundamental role in the regulation of several vital physiological adaptive responses to mountainous areas in sheep, including growth, energy expenditure, thermogenesis, sensory perception, and locomotor activity. Through the identification of compelling candidate selection targets, our study illustrates how natural selection has contributed to environmental adaptation of sheep to highlands.

An important goal in evolutionary biology is to uncover the genetic basis of local adaptation. This is relevant to climate change, animal production, and conservation of genetic resources[1]. With the advance of cost-effective methods for obtaining genome-scale data, the analysis of large datasets from domesticated species offers great opportunities to examine the genetics of local adaptation[2,3]. Exploring genomic patterns and mechanisms of local adaptation offers a way to detect selection signatures and identify the genes, pathways, and environmental factors shaping this crucial evolutionary process[4]. Most adaptive traits are polygenic and involve a few initial major-effect mutations, followed by numerous minor-effect mutations in smaller genes[5,6]. Hence, resolving the genetics of traits involved in local adaptation remains a challenging task because of their polygenic architecture.

Sheep provide a good example of a species whose genome has been shaped mainly by natural selection following integration into diverse environments and production systems since their domestication[7,8]. They have been exposed to a plethora of selective pressures and have adapted to thrive in a diverse range of environments, which makes them important reservoirs for resilience in the face of climate change.

Sarda (SAR) and Valle del Belice (VDB) are two economically important Italian dairy sheep breeds reared on the islands of Sardinia and Sicily, respectively. Within the two breeds, three different subpopulations (or ecotypes) are distinguished according to the geographical altitude of the farming area: plains, hills, and mountains. Animals raised in plain and hill regions usually present greater genetic value for milk production traits than those raised in mountainous areas do. This is particularly true for the Sarda breed[9,10]. Although no studies have investigated genomic differences among ecotypes in the Valle del Belice breed, only one study has reported variations in runs of homozygosity (ROH) patterns among Sarda sheep raised at different altitudes[11]. Consequently, significant knowledge gaps persist regarding the genomic signatures of environmental adaptation in these two prominent dairy sheep breeds. In the present study, we used genotyping data and two complementary extended haplotype

[1]University of Palermo, Palermo, Italy. [2]Institut National de la Recherche Agronomique de Tunisie, Université de Carthage, 2049 Ariana, Tunisia. [3]University of Sassari, Sassari, Italy. [4]University of Molise, Campobasso, Italy. [5]University of Georgia, Athens, USA. ✉email: salvatore.mastrangelo@unipa.it

homozygosity statistics to investigate (a) signatures of selection in the two breeds (by jointly analyzing the three ecotypes from each breed) and (b) evidence for local adaptation across different elevational gradients in both breeds, first by computing a within-ecotype test (iHS) and then by contrasting EHH patterns of the same haplotype between ecotypes (Rsb) within each breed. From the set of candidate genes found to be under strong selection, we outline their potential importance in the process of adaptation to mountainous regions in sheep.

Methods

Ethics declaration

All experimental procedures and sampling were approved by the Bioethics Committee of the University of Palermo: protocol code UNPA-CLE–203098. Blood samples were collected in compliance with European rules [Council Regulation (EC) No. 1/2005 and Council Regulation (EC) No. 1099/2009] during routine health control by the public veterinary service. The authors confirm that they have followed EU standards for the protection of animals used for scientific purposes.

Animals and genotyping

For this study, blood or nasal swab samples were collected from 653 ewes from the two breeds reared under different agroclimatic conditions to capture the representative genetic diversity within each breed. The individuals were classified by ecotype based on whether they were raised in plain areas (≤ 200 m above sea level, m.a.s.l.), hills (350–450 m.a.s.l.) or mountains (> 700 m.a.s.l.). Only farms where animals are born to ancestors raised under the same environmental conditions were included in the sampling process.

Genomic DNA was extracted from blood samples (using the commercial Illustra Blood Genomic Prep Mini Spin Kit) or from nasal swabs (using the MagMAX CORE nucleic acid purification kit). The samples were genotyped via the Illumina OvineSNP50 BeadChip v3 (Illumina, Inc., San Diego, CA, United States). Chromosomal coordinates for single nucleotide polymorphisms (SNPs) were updated to the ARS-UI_Ramb_v2.0 (*Ovis aries*) genome sequence assembly. The genotyping quality was determined via PLINK ver. 1.09[12]. Samples with genotyping rate below 90% were excluded, resulting in the removal of 27 SAR and 14 VDB individuals. We excluded SNP that had a low call rate (< 90%), a high rate of missing genotypes (> 10%) or minor allele frequency (MAF) < 5% and those that were out of Hardy–Weinberg Equilibrium (P value < 0.001). Accordingly, 4193 SNP were removed from the Sarda and 2249 from the Valle del Belice datasets. We then conducted a relationship-based filtering using the *–king-cutoff* option in PLINK ver. 2.0, with a threshold of 0.177 to exclude pairings between closely related individuals. This led to the removal of 33 SAR and 55 VDB additional individuals. After genotype data filtering, selection signature analyses were performed using 42,170 SNP from 266 SAR individuals (82 hill (SAR_C), 77 mountain (SAR_M), and 107 plain (SAR_P) ecotypes) and 44,114 SNP from 258 VDB individuals (69 hill (VDB_C), 77 mountain (VDB_M), and 112 plain (VDB_P) ecotypes).

Nucleotide diversity

Nucleotide diversity (Π) for the 26 autosomes and the X chromosome was estimated via VCFtools[13]. Prior to that, linkage disequilibrium (LD) pruning was applied: SNPs were excluded if the LD between each pair of SNPs was greater than 0.5 ($r^2 > 0.5$) in a window size of 50 SNPs moving 5 SNPs per window, leaving 37,526 and 40,731 variants in SAR and VDB, respectively. Π was first calculated for each chromosome in sliding windows of 500 kb and then averaged over the whole chromosome.

Detecting signatures of selection

We conducted two EHH-based tests, iHS[14] and Rsb[15], implemented in the *rehh* 2.0 R package[16], to detect genomic regions under putative selection for the Sarda and Valle del Belice breeds. Input files for selection signature detection were formatted in fastPHASE[17] format. For this purpose, the software was used to reconstruct haplotypes from the genotyped SNPs. For each breed, 10 random starts and 40 iterations of the EM algorithm were employed, with 50 haplotypes taken from posterior distributions. The number of clusters for cross-validation was set at 40. Two levels of haplotype extension patterns were investigated. First, we applied the iHS test within each dairy breed, considering the three ecotypes from each breed as a single population. The iHS statistic was subsequently calculated independently for each ecotype (plain, hill, and mountain), and the Rsb statistic was used to compare the three ecotypes from each breed. In iHS computation, information on the ancestral and derived allele state is needed for each SNP because this statistic is based on the ratio of the EHH associated with each allele. In our analysis, the ancestral allele was inferred as the most common allele in the dataset[18,19]. The Rsb statistic is defined as the natural logarithm of the ratio between the integrated EHHS (iES) values in two populations (in our case, it is about the three ecotypes) i.e., Rsb = ln(iES_ecotype1/iES_ecotype2), where iES represents the integrated site-specific extended haplotype homozygosity (EHHS) for each SNP within an ecotype. Rsb values are assumed to follow a standard normal distribution under neutrality. Therefore, a Z-test was applied to identify SNP showing significant differentiation in haplotype structure between ecotypes. Two-sided *p*-values were computed as $pRsb = -\log10[1 - 2|\Phi(Rsb) - 0.5|]$, where $\Phi(x)$ denotes the cumulative distribution function of the standard normal distribution. Selection signature detection was performed in 1-Mb sliding windows with a 50-kb overlapping step. A window is classified as putatively under selection when it contains at least 3 markers exceeding the significance threshold of $-\log10$ (p value) = 3[20].

Functional impact of the variants under selective pressure

Ensembl Variant Effect Predictor (VEP)[21] was used to predict the statistically significant variants with the relevant genomic regions putatively under selection. VEP was used to determine the location of the variant (e.g., intronic, intergenic, in a coding sequence, in regulatory regions). The VEP also provides an impact rating (high, moderate, low, or modifier), indicating the severity of the consequences of the mutation.

Results

Genomic diversity analysis

To evaluate the level of genomic diversity in the two Italian dairy sheep breeds, we computed nucleotide diversity (π) in windows of 500 kb across the chromosomes. The X chromosome clearly presented the lowest nucleotide diversity in both breeds: 3.34×10^{-6} and 3.27×10^{-6} in the SAR and VDB, respectively. The average π value on the autosomes was almost double (6.24×10^{-6} and 6.37×10^{-6} for Sarda and Valle del Belice, respectively) (Supplementary Fig. S1).

Selection signature analysis in the Sarda breed

iHS

Examination of the iHS computed jointly for the three SAR ecotypes (SAR_C, SAR_M, and SAR_P) yielded 8 statistically significant regions, of which the top three iHS regions (those showing the highest proportion of outlier SNPs) were on sheep chromosome (OAR) 19 (one region: 29–40.45 Mb) and X (two regions: 40.4–45.9 Mb and 52.5–55.6 Mb) (Supplementary Fig. S2; Supplementary Table S1). These three candidate regions have at least 37% of their SNPs exceeding the significance threshold (Supplementary Table S1). Chromosome 19 had the largest peak (11.45 Mb in length), which included 77 outliers. Inside this region, the variants with the highest statistical significance (-log (p values) between 4.66 and 5.87) are concentrated within a 3.5-Mb region (35.79–39.3 Mb). All these variants fell within the intronic regions of four genes, *MAGI1, PRICKLE2, SYNPR*, and *PTPRG* (Supplementary Table S2), which are related to neuronal function and synaptic organization.

The first candidate region on the X chromosome (40.4–45.9 Mb) had the highest proportion of outliers (79% of the SNPs exceeded the significance threshold) and the highest iHS values. Within a ~747 kb region (between positions 43,984,832 and 44,733,729 bp), 6 SNPs had –log (p-values) ranging from 6.48 to 8.35. These SNPs are located either in intergenic regions or within the intronic regions of two genes, namely, *KDM6A* (2 markers) and *DIPK2B* (1 marker), and were classified as modifier impact variants based on the VEP-derived annotations (Supplementary Table S2). The second region on the X chromosome (52.5–55.6 Mb) includes 8 SNPs, which are consequently also predicted as "modifiers" by VEP. Most of these SNPs (5) are located within intronic regions of two protein-coding genes: *SHROOM4* and *DGKK* (Supplementary Table S2).

The iHS plots for each of the three Sarda ecotypes are in Fig. 1 and indicate for each ecotype the SNP above the selected significance threshold ($-\log_{10}$(p-value) = 3). We identified a candidate region on chromosome 3 (166.9–168.7 Mb), unique to the mountain ecotype, containing four significant SNP (Fig. 1; Supplementary Table S3). One variant (*rs401844951*) lies within an intron of *Ankyrin Repeat And Sterile Alpha Motif Domain Containing 1B* (*ANKS1B)* gene (Supplementary Table S4) and shows a negative iHS value, indicating extended haplotypes with the derived allele. Figure 2 shows bifurcation plots for *rs401844951* in the SAR_M and SAR_P ecotypes. In the mountain ecotype (Fig. 2a), the derived allele (SAR_M, red) is associated with a long, unbranched haplotype, consistent with recent positive selection. In contrast, the plain ecotype (Fig. 2b) displays rapid branching, indicating no strong selective sweep. Similarly, the decay of the site-specific extended haplotype homozygosity (EHHS) plotted against *rs401844951* also clearly shows a slower decay of the derived allele in the SAR_M ecotype than in the SAR_P ecotype (Supplementary Fig. S3).

Rsb

We subsequently investigated haplotype extension patterns via the standardized log ratio of integrated EHHS (iES) between pairs of populations (Rsb) analysis. Given that we are primarily interested in finding areas under selective pressure that are associated with environmental adaptation to mountainous environments, we focused on the comparisons between plain and mountain ecotypes in both breeds. In our situation, extremely positive values of Rsb would indicate longer haplotypes in the mountain ecotype and hence suggest that selection occurred in this ecotype. Each of the Rsb SAR_C/SAR_M and Rsb SAR_C/SAR_P statistics identified 12 outlier windows (Fig. 3a,b; Supplementary Table S5). The Rsb SAR_M/SAR_P test revealed a total of 14 candidate genomic regions (Fig. 3c; Supplementary Table S5), of which 11 were under selection in the mountain ecotype. Among these, we identified four highly significant windows where at least 20% of the outlier SNPs exceeded the significance threshold. Three of these four windows are located on chromosome 3 (Table 1). The first one, at position 166.2–174.3 Mb, overlaps with that revealed by the iHS test in the mountain ecotype (Supplementary Table S3), thus providing more evidence of the existence of selection toward this area in the Sarda ecotype reared in mountainous environments. Most of the 65 statistically significant SNPs located within this candidate region fell within introns of several protein-coding genes, such as *ANKS1B*, *SCYL2*, *SLC17A8*, *NR1H4*, *ANO4*, *UTP20*, *SPIC*, and *IGF-1*. We found that a single SNP (*rs159877897*) was located within a coding region of the *EP300 Interacting Inhibitor Of Differentiation 3* (*EID3*) gene. The highest Rsb value between the SAR_M and SAR_P ecotypes on chromosome 3 was in an intergenic region, 87.7 kb downstream of the *IGF-1* gene. The second relevant candidate region on OAR03 is located close to the first region (OAR03: 176.7–179.1 Mb) and includes 12 significant SNPs, of which 8 are in intronic regions of the *LARGE1* gene (modifier impact variants). The third relevant candidate region on OAR03, which is located between 186.65 and 190.75 Mb, includes 27 significant SNPs (Supplementary Table S6). Within this region, the strongest variants under selection fell within intronic regions of the *BMAL2* gene (also known as *ARNTL2*), which plays a role in regulating the circadian rhythm. The fourth relevant region under strong selection in SAR_M is located on OAR01 between 218.6 and 224.05 Mb and holds the genes *GOLIM4, SERPINI2 and SI* (Table 1).

We also identified a region under strong selection in the hill (SAR_C) ecotype on OAR01 where more than 42% of the variants in the interval exceeded the significance threshold in both Rsb SAR_C/SAR_M (123.4–133.55 Mb) and Rsb SAR_C/SAR_P tests (126–132.85 Mb) (Supplementary Table S5). Within this interval, seven shared protein-coding genes (between the two comparisons) included at least one statistically significant variant: *GRIK1, MAP3K7CL, RWDD2B, LTN1, N6AMT1, GABPA* and *APP*. The latter gene included a synonymous

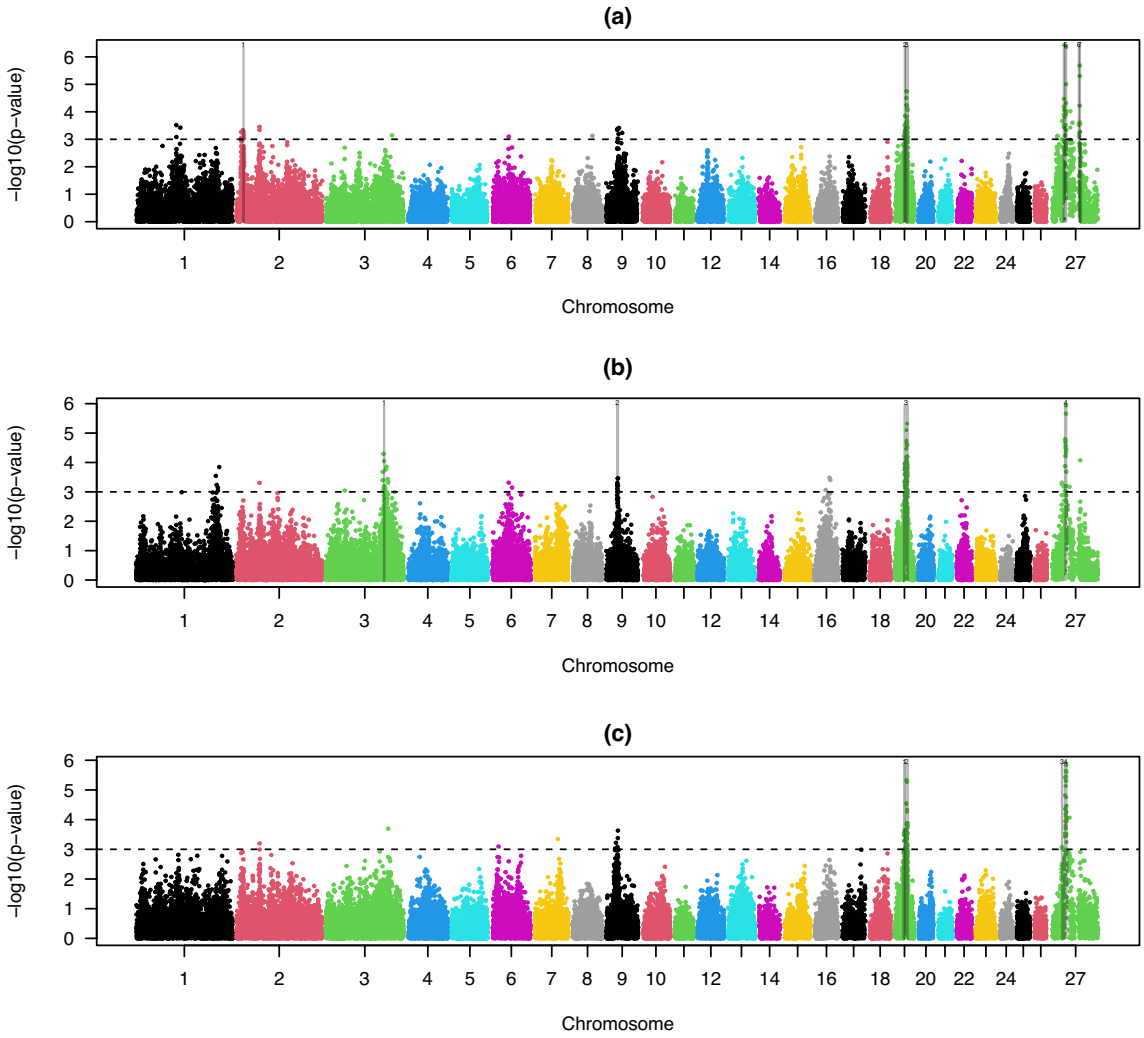

Fig. 1. Manhattan plots of the genome-wide iHS test within each of the three Sarda ecotypes. (**a**) iHS test in SAR_C (hill). (**b**) iHS test in SAR_M (mountain). (**c**) iHS test in SAR_P (plain). The horizontal dashed lines mark the significance threshold applied to detect the outlier SNPs (−log (p values) = 3.

mutation (*rs412371953*) with a low impact on the protein in both comparisons (Supplementary Table S7). *APP* (*amyloid precursor protein*) is a type 1 membrane glycoprotein that plays an important role in a range of biological activities, including neuronal development, signaling, intracellular transport, and other aspects of neuronal homeostasis[22].

Selection signature analysis in the Valle del Belice breed

iHS

A total of 13 candidate regions were detected via the iHS test, which was jointly applied to the three Valle del Belice ecotypes (Supplementary Fig. S4), five of which are located on chromosome 3 (184.45–211.4 Mb). The other eight regions are distributed across chromosomes 8 (2 regions), 10 (3 regions), 11, 23 and X (1 region each) (Supplementary Table S1). The region with the highest proportion of outlier SNPs represented only 23% of the variants that exceeded the significance threshold. This region is located on the X chromosome at 43.25–45.4 Mb, which overlaps with the region revealed in the Sarda breed (Supplementary Table S1).

The within-ecotype iHS analysis revealed 1, 11, and 14 genomic regions potentially under selection in VDB_C, VDB_M, and VDB_P, respectively (Fig. 4; Supplementary Table S3). Among these, three candidate regions were specific to the mountain ecotype, as they were detected exclusively in VDB_M. These windows are located on OAR03 (108.3–109.7 Mb and 115.2–116.4 Mb) and on OAR06 (80.3–82.8 Mb). The statistically significant SNPs in these three regions were in introns of *Synaptotagmin 1* (*SYT1*), *Thyrotropin Releasing Hormone Degrading Enzyme* (*TRHDE*) for OAR03 and *Ephrin Type-A Receptor* (*EPHA5*) for OAR06 (Supplementary Table S8).

Rsb

Rsb analysis identified 17, 15, and 28 candidate genomic regions in the VDB_C/VDB_M, VDB_C/VDB_P, and VDB_M/VDB_P comparisons, respectively. The 28 outlier windows revealed by the Rsb analysis between

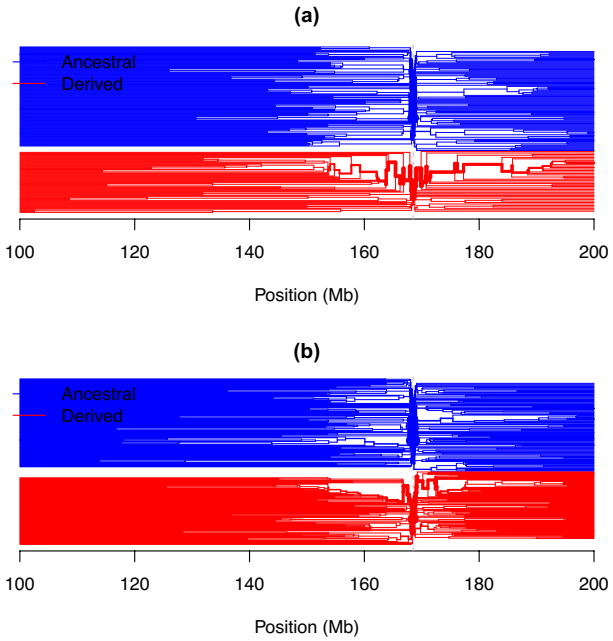

Fig. 2. Haplotype bifurcation diagrams of ancestral and derived alleles around rs401844951, located within ANKS1B gene. (**a**) Mountain ecotype of Sarda. (**b**) Plain ecotype of Sarda.

mountain and plain ecotypes of Valle del Belice are distributed over 11 chromosomes (Fig. 5; Supplementary Table S5). We identified 13 highly significant windows, all of which are under selection in the VDB_M ecotype (Table 1). Notably, some of these windows overlap with the candidate regions on chromosomes 3 and 6 that were previously highlighted by the iHS test as being specific to the mountain ecotype. Chromosomes 6 and 4 presented the largest peaks, with 14.6 (OAR06, 70.85–85.45 Mb) and 10.6 Mb (OAR04, 86.95–97.55 Mb) in length. The former candidate region presented the second-highest number of statistically significant variants (~60%) (Table 1). Notably, a group of 18 highly significant variants on OAR06 (76.7–82 Mb; $7.7 < -\log$ (p value) < 9.9) clustered in the intronic regions of three genes: *ADGRL3*, *TECRL*, and *EPHA5*. In addition, a second cluster of 5 significant SNPs ($3.89 < -\log$ (p value) < 5.03) flanking the gonadotropin-releasing hormone receptor (*GNRHR*) gene was identified. On chromosome 4, the SNPs with the highest statistical significance ($-\log$ (p value) > 5) are in the introns of the *CADPS2* gene. At the genome-wide level, the strongest signal was found on chromosome 3 between 106.6 and 114.75 Mb, with more than 63% of the SNPs exceeding the significance threshold. The most significant SNPs ($7.74 < -\log$ (p value) < 9.48) were concentrated within a 126-kb region (113,459,553–113,585,231 bp). All of them fell within intronic regions of the *neuron navigator 3* (*NAV3*) gene. Another chromosome that showed highly significant regions only between VDB_M and VDB_P was OAR09. Within this chromosome, four outlier windows were detected between 55.35 and 68.85 Mb (Supplementary Table S5). Most of the SNPs were located in the introns of the following genes: *TPD52, TRPS1, CSMD3, NUDCD1*, and *TRHR*. The variants with the highest Rsb values clustered in a 172-kb region that encompasses *KCNV1*, a voltage-gated potassium (Kv) channel gene. Voltage-gated potassium (Kv) channels have diverse functions, including regulating neurotransmitter release, heart rate, insulin secretion, neuronal excitability, epithelial electrolyte transport, smooth muscle contraction, and cell volume[23]. According to VEP annotations, most of the statistically significant variants under selection in the mountain ecotype of Valle del Belice fell either within intergenic regions or within introns of a hundred protein-coding genes. We found four synonymous mutations within the *CSMD2, LIPT1, OVCH1*, and *RTF1* genes with a predicted low impact on the protein. We also identified two additional missense variants (having a moderate impact on the protein) within *ELF5* (a key regulator of mammary gland alveologenesis) and *PKHD1L1*, a gene marker of the embryonic epicardial tissue that affects hearing sensitivity (Supplementary Table S9).

Discussion

In this work, we relied on widely used LD-based tests, centered around the extended haplotype homozygosity concept[24], to detect selection signatures in the genomes of Sarda and Valle del Belice sheep ecotypes raised across different altitudinal gradients. We performed a two-level detection analysis. First, we conducted EHH-based tests to detect regions under putative selection for each dairy breed (i.e., by jointly considering the three ecotypes from each of the Sarda and Valle del Belice). Second, we conducted analyses to determine whether any genomic regions were significantly associated with a specific ecotype, particularly the one raised in a mountainous environment. In both Sarda and Valle del Belice breeds, there is almost no use of artificial insemination, and the exchange of rams among farms is limited; when it does occur, it typically happens between neighboring farms within the same geographic area. Furthermore, during sampling, we ensured that the selected animals had ancestors born and raised in the same environment, allowing us to minimize recent gene flow across ecotypes.

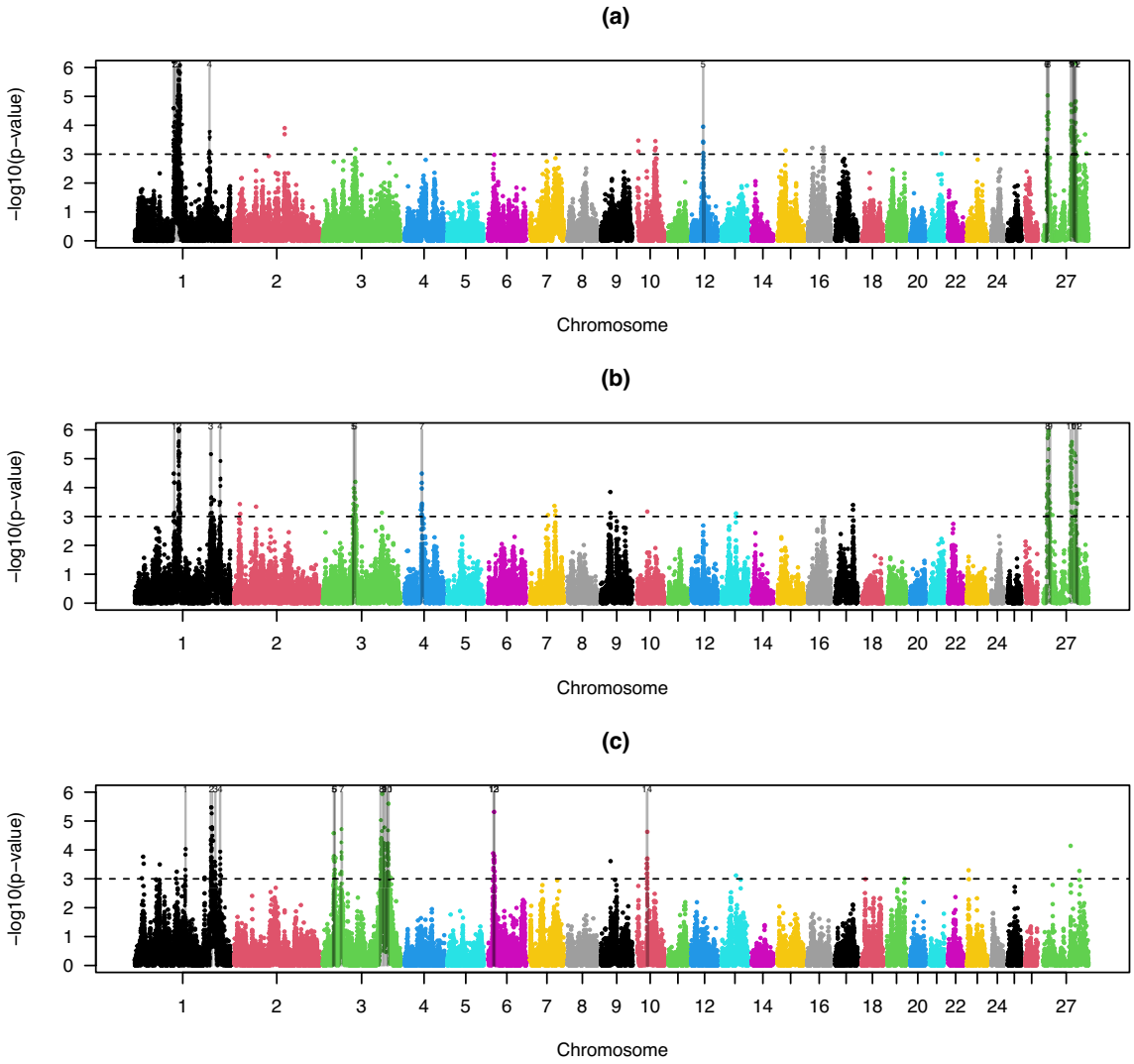

Fig. 3. Manhattan plots of the genome-wide Rsb test between the three ecotypes of Sarda. (**a**) Rsb SAR_C/SAR_M. (**b**) Rsb SAR_C/SAR_P. (**c**) Rsb SAR_M/SAR_P. The horizontal dashed lines mark the significance threshold applied to detect the outlier SNPs (−log (p values) = 3.

These breeding and management practices contribute to maintaining genetic separation among ecotypes over several generations.

Within-breed signatures of selection detection

Compared to Valle del Belice, the Sarda breed exhibited stronger within-population selection signals, with over 24% of variants in the top four candidate regions exceeding the significance threshold. This likely reflects both demographic and selective factors: as Italy's largest sheep breed (~ 80% of the national stock)[25], Sarda may harbor more standing genetic variation and have undergone a longer history of selection, thereby increasing the power to detect iHS signals.

Most variants under selection in both breeds lie in intronic or intergenic regions. While these regions do not directly affect protein-coding sequences, they can influence gene regulation through mechanisms such as alternative splicing or chromatin remodeling.

The region on the X chromosome between 40 and 45 Mb had the highest proportion of outlier variants and the highest iHS values (the latter criterion is particularly true for Sarda) where two genes, *KDM6A* and *DIPK2B*, harboured significant variants in both breeds. This region overlaps with a conserved, low-diversity X-linked segment identified in global sheep populations[26]. Our data also consistently revealed clearly lower nucleotide diversity on the X chromosome compared to the autosomes (Supplementary Fig. S1). Thus, one wonders whether the selective sweeps observed on the X chromosome are caused primarily by a low recombination rate, as observed in Drosophila[27]. This would generate long haplotype blocks that can mimic the effect of a selective sweep. Although this hypothesis may lead to a cautious interpretation of the strong selection signal on the X chromosome, we cannot rule out the presence of genes under selective pressure in this region because sex chromosomes are subject to sex-specific selective evolutionary forces[28]. Interestingly, we found that the most

Comparison	Region	% ext_mar	Genes
SAR_M/SAR_P	1: 218.6–224.05	42	*GOLIM4, SERPINI2, SI*
SAR_M/SAR_P	3: 166.2–174.3	46	*ANKS1B, ANO4, APAF1, EID3, GLT8D2, IGF1, NR1H4, NT5DC3, PAH, SCYL2, SLC17A8, SPIC, STAB2, TDG, TXNRD1, U2, UTP20*
SAR_M/SAR_P	3: 176.7–179.1	24	*LARGE1*
SAR_M/SAR_P	3: 186.65–190.75	37	*KLHL42, MANSC4, MRPS35, PPFIBP1, BMAL2, STK38L, TM7SF3, ITPR2, SSPN, IRAG2*
VDB_M/VDB_P	2: 108.15–112.8	22	*GALNTL6, MFAP3L, SH3RF1*
VDB_M/VDB_P	2: 115.95–122.1	39	*HS6ST1, UGGT1, SAP130, POLR2D, WDR33, MYO7B, IWS1, MAP3K2, NAB1, MFSD6, HIBCH, COL5A2, COL3A1, GULP1, TFPI*
VDB_M/VDB_P	3: 54.35–56.65	22	–
VDB_M/VDB_P	3: 100.05–103.15	28	*NMS, AFF3, EIF5B, MITD1, LIPT1, MGAT4A, UNC50*
VDB_M/VDB_P	3: 104–106.05	27	*MGAT4A, BUB1, ACOXL, BCL2L11*
VDB_M/VDB_P	3: 106.6–114.75	63	*LGR5, TRHDE, ATXN7L3B, KCNC2, CAPS2, GLIPR1L1, KRR1, PHLDA1, OSBPL8, ZDHHC17, CSRP2, E2F7, NAV3*
VDB_M/VDB_P	3: 184–190.65	56	*IPO8, TMTC1, OVCH1, ERGIC2, FAR2, CCDC91, KLHL42, MANSC4, MRPS35, PPFIBP1, U6, BMAL2, STK38L, TM7SF3, FGFR1OP2, ITPR2, BHLHE41, RASSF8*
VDB_M/VDB_P	4: 86.95–97.55	35	*PTPRZ1, CADPS2, POT1, LOW, GRM8, SND1, RBM28, IMPDH1, IRF5, TNPO3, TSPAN33, AHCYL2, MKLN1*
VDB_M/VDB_P	4: 103.75–106.25	22	*DGKI, HIPK2, TBXAS1, PARP12, KDM7A, SLC37A3, MKRN1, DENND2A, TMEM178B*
VDB_M/VDB_P	6: 70.85–85.45	60	*SRD5A3, NMU, EXOC1, CEP135, CRACD, IGFBP7, ADGRL3, TECRL, EPHA5, UBA6, GNRHR, TMPRSS11A, TMPRSS11F*
VDB_M/VDB_P	7: 30.45–35.85	28	*MEIS2, TMCO5A, FSIP1, EIF2AK4, U6, INO80, CHP1, EXD1*
VDB_M/VDB_P	9: 64.05–66.25	29	*CSMD3*
VDB_M/VDB_P	9: 66.35–68.85	24	*PKHD1L1, NUDCD1, TRHR*

Table 1. Genes containing significant variants within the relevant candidate regions revealed by Rsb mountain/plain ecotypes. % ext_mar = percentage of markers in the interval that exceeded the significance threshold.

significant variants, both in Sarda and Valle del Belice, clustered within *KDM6A*, a gene known for its female-biased expression[29]. Mounting evidence suggests that selection pressure on X-linked genetic variation may preferentially favor alleles that benefit females[30]. The presence of strong signals of selection on *KDM6A* gene in both breeds suggests a possible case of parallel adaptation, where similar environmental pressures may have independently driven selection on the same gene in these two genetically distinct populations. The *KDM6A* gene controls luminal milk-secreting cells in the mammary gland[31]. Polymorphisms in this gene, observed in the plain ecotypes of Sarda and Valle del Belice, may influence milk volume and composition by affecting alveolar development, potentially contributing to the higher productivity observed in these ecotypes. The presence of selective signals in *KDM6A* in mountain ecotypes suggests an additional adaptive role. As an epigenetic regulator, *KDM6A* modulates gene expression in response to environmental stress; under energy- or oxygen-limited conditions, its inhibition suppresses differentiation pathways, maintaining cells in an undifferentiated state, potentially enhancing cellular plasticity under hypoxic stress[32]. In addition, *KDM6A* is directly recruited to myogenic targets during differentiation[33] and is required for skeletal muscle repair and regeneration[34]. Adaptive variants in *KDM6A* could thus contribute to improved endurance and mobility in challenging mountainous environments.

In Sarda, chromosome 19 showed the highest proportion of significant SNP. The candidate region on this chromosome was also shown to be under divergent selection between the Sarda and the Sardinian Ancestral Black sheep breeds[35]. The most notable variants lie within the *MAGI1*, a gene involved in epithelial cell integrity[36] and mammary function[37], likely relevant for dairy traits. In Valle del Belice, three previously reported candidate regions on chromosomes 3 (189.15–191.1 Mb), 10 (45.35–47.45 Mb), and 11 (20.65–21.8 Mb) were confirmed[38]. Significant variants fell within genes such as *LMNTD1* and *BCAT1* (Supplementary Table S10), both linked to muscle development, consistent with secondary selection for growth traits in this breed.

Candidate genes under selection in mountain ecotypes

Over generations, populations living at high altitude levels have developed genetic adaptations to environmental stressors such as cold temperatures and low oxygen levels[39]. A common trait among mountain-dwelling animals is smaller body size and slower growth[40], features reported in sheep breeds like Tibetan[41] and Sarda[11], and thought to aid in coping with high-altitude conditions. In this study, we identified several genomic regions consistently associated with the mountain ecotype. Notably, in Sarda sheep, a 1.8 Mb region on chromosome 3 (166.9–168.7 Mb), identified by the iHS test, was exclusive to the mountain population and overlapped with a broader divergent window detected only in the SAR_M/SAR_P comparison, supporting its role in altitude adaptation.

We found a high number of genes under selective pressure in the mountain ecotype of SAR and VDB related to nervous system development, energy metabolism, growth, and muscle function, key pathways for physiological and behavioral adaptation to mountainous environments. The nervous system likely plays a central role, as mountain sheep face greater temperature fluctuations, requiring efficient thermoregulation. Focusing first on variants with moderate predicted effects, we detected two missense variants under selection in *ELF5*

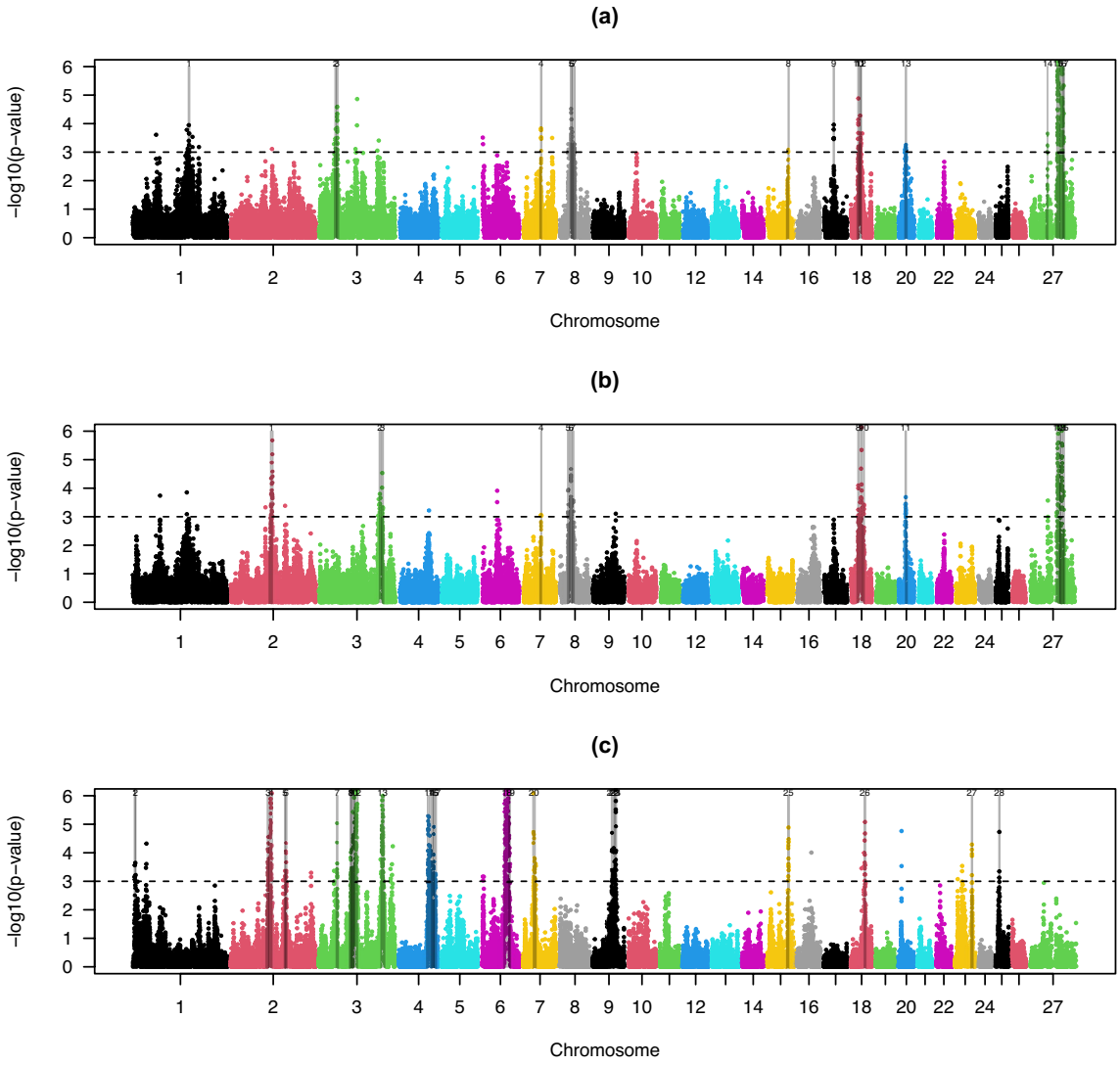

Fig. 4. Manhattan plots of the genome-wide iHS test within each of the three Valle del Belice ecotypes. (**a**) iHS test in VDB_C (hill). (**b**) iHS test in VDB_M (mountain). (**c**) iHS test in VDB_P (plain). The horizontal dashed lines mark the significance threshold applied to detect the outlier SNPs (− log (p values) = 3.

and *PKHD1L1* in the VDB_M ecotype. The *ELF5* gene, involved in prolactin signaling and mammary gland development[42], may enhance lactation under environmental stress. The *PKHD1L1* could reflect dual adaptation: improved sensory function and cardiovascular resilience, both crucial for acclimatization to high altitudes[43–45].

Putative candidate genes related to neurons, growth and muscle

We observed that numerous, statistically significant, modifier impact variants are located within genes with pleiotropic effects on the nervous system, body size, and muscular development. The most striking results involve those on chromosome 3 (both in Sarda and Valle del Belice). In Sarda, the region under strong selection on chromosome 3 (166.2–174.3 Mb) carries the *ANKS1B* gene because: (1) it includes the highest number of statistically significant variants (12) with the Rsb SAR_M/SAR_P test; (2) the statistically significant variant located within this gene (in the iHS test) presented a high-derived allele frequency and long-term haplotype homozygosity in the mountain ecotype but not in the plain ecotype (Fig. 2; Supplementary Fig. S3). The *ANKS1B* gene encodes AIDA-1, a protein that is enriched at neuronal synapses and regulates synaptic plasticity and social behavior in humans[46]. *ANKS1B* has also been shown to be associated with adiposity and body weight in humans[47–49] and sheep[50]. These functions suggest that mutations in the *ANKS1B* gene may contribute to mountain adaptation through mechanisms involving both energy metabolism and behavioral plasticity. In mountainous environments, animals face challenges such as limited food availability and lower temperatures. Therefore, genetic variants in the *ANKS1B* gene, that influence fat deposition and body weight, could support improved energy storage and thermoregulation under these harsh conditions.

Within the same interval, two significant SNPs, revealed by the Rsb SAR_M/SAR_P test, were in intronic and downstream regions of *insulin-like growth factor 1* (*IGF-1*), which is essential for embryonic and adult

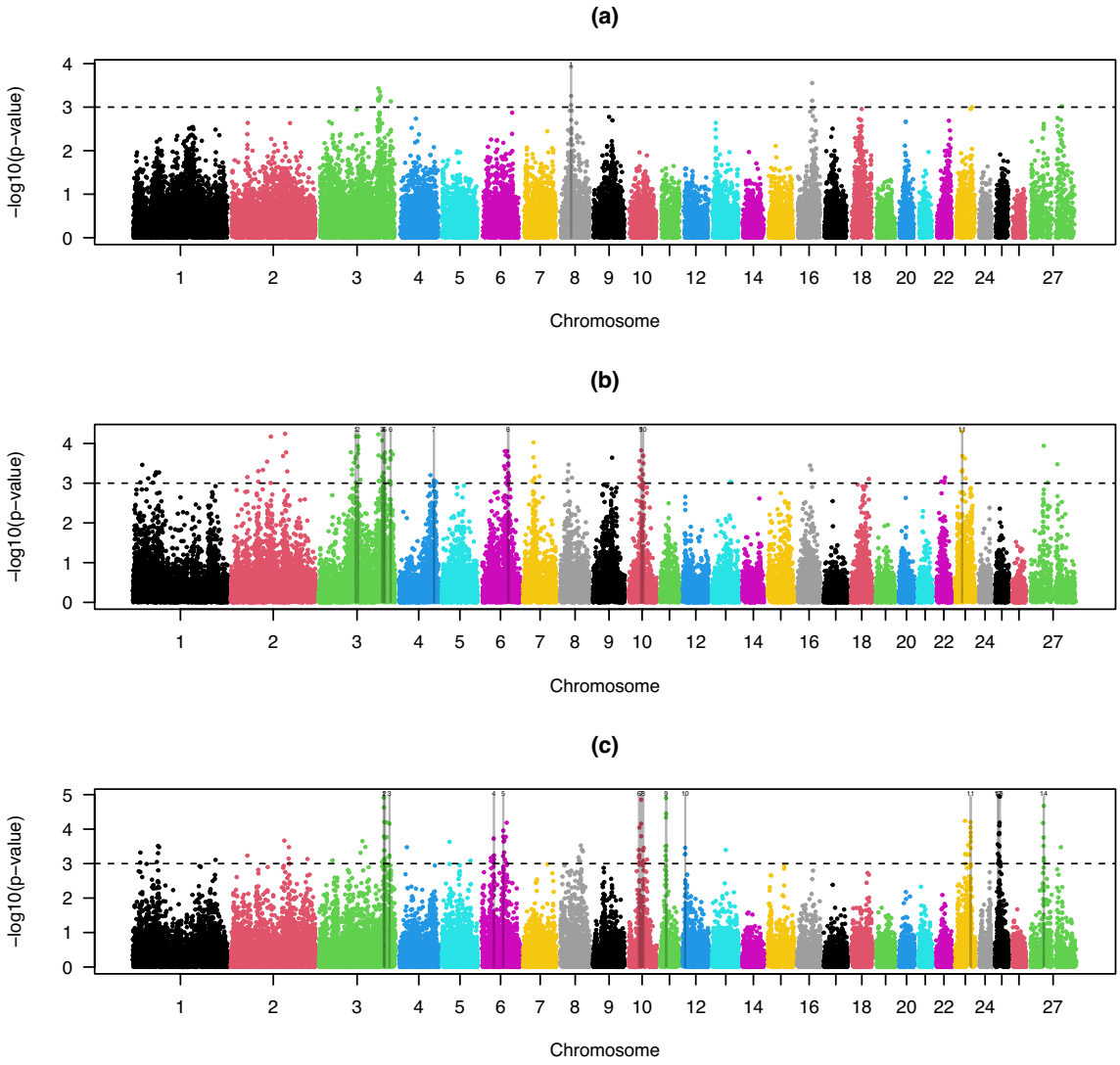

Fig. 5. Manhattan plots of the genome-wide Rsb test between the three ecotypes of the Valle del Belice breed. (**a**) Rsb VDB_C/VDB_M. (**b**) Rsb VDB_C/VDB_P. (**c**) Rsb VDB_M/VDB_P. The horizontal dashed lines mark the significance threshold applied to detect the outlier SNPs (−log (p values) = 3.

neurogenesis, neuronal plasticity, angiogenesis, and metabolic functions[51]. *IGF-1* is a major determinant of small size in dogs and one of the foremost regulators of skeletal muscle metabolism[52,53]. Additionally, within the same region, we observed a synonymous mutation within the *EID3* gene. The protein encoded by this gene interacts with EP300, a histone acetyltransferase encoding a hub protein, p300, which interacts with hundreds of other proteins[54].

P300 is crucial for hypoxia-induced gene activation, neuron survival[55], and skeletal muscle function[56]. The reduction in oxygen availability at 700–1000 m is mild. At this range of altitude, the partial pressure of oxygen is approximately 90% of that at sea level[57]. However, chronic exposure over generations, especially when combined with other environmental stressors such as cold temperatures, rugged terrain, and limited forage, may lead to mild hypoxic stress that contributes to adaptive responses in the mountain ecotype of Sarda sheep. On chromosome 3, most significant variants mapped in *LARGE1*, a gene linked to reduced body weight, growth retardation[58], and muscle function[59]. On OAR01, two growth-related genes, *SERPINI2* and *SI*, were also detected. Mutations in these genes cause growth defects in mice and humans[60,61].

In Valle del Belice, *SYT1*, under selection only in the mountain ecotype, encodes a calcium sensor protein involved in neurotransmitter release and synaptic vesicle trafficking[62]. The release of neurotransmitters in the brain is the main physiological mechanism that organizes the response to stress[63]. The strongest genome-wide signal was found on chromosome 3 (at position 113,459,553–113,585,231 bp) within an intronic variant in *NAV3*, a neuronal gene involved in morphogenesis and neuromuscular function[64]. Additionally, on chromosome 2, 13 SNPs under strong selection were in genes linked to muscle (e.g., *MYO7B*, *HS6ST1*), neural development (*MFSD6*), and collagen structure (*COL5A2, COL3A1*). Differential expression of *COL5A2* and *COL3A1* genes in yaks across altitudes suggests a possible role in adaptation to mountainous environments[65].

On chromosome 4, the SNPs with the highest statistical significance (−log (p value) > 5) are in the introns of the *CADPS2* gene, which is expressed at the highest levels in the brain. *CADPS2* regulates the release of two neurotrophins, brain-derived neurotrophic factor (*BDNF*) and neurotrophin-3 (*NT3*)[66]. *BDNF* is a critical regulator of neuroplasticity, cognitive function, and neuronal survival, with its expression known to be modulated by environmental stressors such as hypoxia, extreme temperature fluctuations, and physical exertion[67–69]. Accordingly, the chronic exposure to reduced oxygen availability, thermal stress, and increased locomotor demands associated with rugged mountainous terrain may stimulate *BDNF*-mediated pathways that enhance neurological resilience, sensory integration, and overall stress adaptation in Valle del Belice sheep living in such environments.

Putative Candidate genes involved in neuroendocrine mechanisms

Many of the markers under strong selection that we discovered are located within genes that play key roles in the regulation of neuroendocrine functions, such as the hypothalamus-pituitary-adrenal (HPA) axis. In the mountain ecotype of Valle del Belice, the significant SNPs within the outlier window on chromosome 3 common to both the iHS VDB_M and Rsb VDB_M/VDB_P tests (OAR03: 108.3–109.7 Mb) fell within introns of the *thyroidropin-releasing hormone-degrading enzyme (TRHDE)*. This gene encodes an enzyme that specifically cleaves and inactivates a neuropeptide, thyrotropin-releasing hormone (TRH), which is closely interconnected to the HPA axis through neuroendocrine crosstalk. *TRH* plays a critical role in mediating changes in metabolism, energy balance and thermogenesis[70]. A mutation in *TRHDE* gene may help maintain higher TRH levels, enhancing cold tolerance and metabolic flexibility in the mountain ecotype of Valle del Belice. This gene lies within a CNV region previously linked to altitude adaptation in cattle[71] and sheep[72]. Additional HPA axis–related genes under selection in this ecotype include *SYT1, SLCO1C1, EPHA5, GNRHR*, and *TRHR*. Since the HPA axis is key to thermal adaptation, selection on these genes may improve metabolic regulation and homeostasis in cold, mountainous environments.

In Sarda sheep, several genes under selection on chromosome 3 (*IGF-1*, *SLC17A8*, and *NR1H4*) are linked to neuroendocrine functions. Among them, *IGF-1*, essential for fetal and placental growth[73], has been associated with high-altitude adaptation in humans[74]. Selection on *IGF-1* may reflect an adaptive response to prenatal stress in mountain environments, promoting smaller, more resilient offspring. Epigenetic regulation of *IGF-1* under stress supports its role in growth-related adaptation[75]. Additionally, strong selection signals were found in *BMAL2*, a key circadian regulator. The *BMAL2* gene interacts with hypoxia-inducible factors, linking circadian rhythms to hypoxia responses[76]. Selection on *BMAL2* may thus help synchronize circadian and hypoxia responses in mountain ecotypes.

These findings are consistent with and expand our knowledge of the mechanisms underlying the environmental adaptations of sheep to mountain environments, wherein the neuroendocrine system is a key regulator of complex homeostatic processes in the body, including growth, metabolism, energy expenditure and thermogenesis. This probably had an impact on the neuroendocrine control of reproductive function in mountain ecotypes and could account for the stronger reproductive seasonality observed in individuals dwelling at higher altitudes[77].

Limits of the study and future directions

Although our findings identify multiple regions under putative selection in the two major dairy sheep breeds in Italy and provide plausible biological hypotheses, it is important to acknowledge that these signals have not yet been cross-validated. Some of our reported outlier regions may represent false positives. To address this, future work should aim to replicate these signals in independent cohorts from the same breeds. Furthermore, increasing marker density, using high density BeadChip or whole-genome resequencing, would increase statistical power and refine the candidate regions. Transcriptomic analyses, such as RNA-seq or qPCR, could test whether the candidate genes identified in the present study are differentially expressed across ecotypes, particularly in tissues relevant to their putative function (e.g., mammary gland, muscle, brain, adipocytes).

In conclusion, our study provides some of the first clues into the genetic basis of local adaptation of sheep to highlands on the basis of comparisons of ecotypes from the same breed. We identified various genome regions putatively under selection in the mountain ecotypes of Sarda and Valle. The genes under selective pressure are associated with neuroendocrine functions, metabolism, growth, neural and muscle development. Beyond the identification of selection signatures, our results have practical implications for breeding and conservation. Genes involved in traits such as thermoregulation, milk synthesis, and endurance, if validated, could be targeted in selection programs aiming to improve climate resilience. Genomic selection schemes could incorporate adaptive haplotypes to produce better animals suited to harsh and variable environments, such as those predicted under future climate change scenarios. This approach is especially relevant for low-input or marginal systems where genetic adaptation is key to long-term sustainability. Furthermore, preserving the genetic integrity of mountain ecotypes through in situ conservation and controlled gene flow could help preserve alleles that confer environmental robustness, which may be lost under intensive selection for productivity traits alone.

Data availability

All data produced in this study are included in the supplementary information.

Received: 11 April 2025; Accepted: 4 August 2025
Published online: 08 August 2025

References

1. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**, 807–820 (2013).
2. Fariello, M.-I. et al. Selection signatures in worldwide sheep populations. *PLoS ONE* **9**, e103813 (2014).
3. Rochus, C. M. et al. Revealing the selection history of adaptive loci using genome-wide scans for selection: An example from domestic sheep. *BMC Genom.* **19**, 71 (2018).
4. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
5. Orr, H. A. The genetic theory of adaptation: A brief history. *Nat. Rev. Genet.* **6**, 119–127 (2005).
6. Olson-Manning, C. F., Wagner, M. R. & Mitchell-Olds, T. Adaptive evolution: Evaluating empirical support for theoretical predictions. *Nat. Rev. Genet.* **13**, 867–877 (2012).
7. Kim, E.-S. et al. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity* **116**, 255–264 (2016).
8. Serranito, B. et al. Search for selection signatures related to trypanosomosis tolerance in African Goats. *Front. Genet.* **12**, 715732 (2021).
9. Macciotta, N. P., Cappio-Borlino, A. & Pulina, G. Analysis of environmental effects on test day milk yields of Sarda dairy ewes. *J. Dairy Sci.* **82**, 2212–2217 (1999).
10. Correddu, F. et al. Effect of altitude of flock location, season of milk production and ripening time on the fatty acid profile of Pecorino Sardo cheese. *Int. Dairy J.* **113**, 104895 (2021).
11. Cesarani, A., Gaspa, G., Correddu, F., Dimauro, C. & Macciotta, N. P. P. Unravelling the effect of environment on the genome of Sarda breed ewes using Runs of Homozygosity. *J. Anim. Breed. Genet.* **139**, 292–306 (2022).
12. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
13. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
14. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
15. Tang, K., Thornton, K. R. & Stoneking, M. A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biol.* **5**, e171 (2007).
16. Gautier, M., Klassmann, A. & Vitalis, R. rehh 20: A reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol. Ecol. Resources.* **17**, 78–90 (2017).
17. Scheet, P. & Stephens, M. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**, 629–644 (2006).
18. Ben-Jemaa, S. et al. Genome-wide scans for signatures of selection in North African sheep reveals differentially selected regions between fat- and thin-tailed breeds. *Anim. Genet.* **56**, e13487 (2025).
19. Deng, T., Ma, X., Duan, A., Lu, X. & Abdel-Shafy, H. Genomic insights into selection signatures and candidate genes for milk production traits in buffalo population. *Animal* https://doi.org/10.1016/j.animal.2025.101427 (2025).
20. Flori, L. et al. Adaptive admixture in the West African bovine hybrid zone: Insight from the Borgou population. *Mol. Ecol.* **23**, 3241–3257 (2014).
21. McLaren, W. et al. The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
22. Chen, G. et al. Amyloid beta: Structure, biology and structure-based therapeutic development. *Acta Pharmacol. Sin.* **38**, 1205–1235 (2017).
23. KCNV1 potassium voltage-gated channel modifier subfamily V member 1 [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/27012.
24. Sabeti, P. C. et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002).
25. Pulina, G. et al. Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. *J. Dairy Sci.* **101**, 6715–6729 (2018).
26. Chen, Z.-H. et al. Contrasting patterns of genomic diversity reveal accelerated genetic drift but reduced directional selection on X-chromosome in wild and domestic sheep species. *Genome Biol. Evol.* **10**, 1282–1297 (2018).
27. Vicoso, B. & Charlesworth, B. Recombination rates may affect the ratio of X to autosomal noncoding polymorphism in African populations of drosophila melanogaster. *Genetics* **181**, 1699–1701 (2009).
28. Vallender, E. J. & Lahn, B. T. How mammalian sex chromosomes acquired their peculiar gene content. *BioEssays* **26**, 159–169 (2004).
29. Ma, W. et al. Sex-biased and parental allele-specific gene regulation by KDM6A. *Biol. Sex Differ.* **13**, 40 (2022).
30. Meisel, R. P., Malone, J. H. & Clark, A. G. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res.* **22**, 1255–1265 (2012).
31. Yoo, K. H. et al. Histone demethylase KDM6A controls the mammary luminal lineage through enzyme-independent mechanisms. *Mol. Cell Biol.* **36**, 2108–2120 (2016).
32. Chakraborty, A. A. et al. Histone demethylase KDM6A directly senses oxygen to control chromatin and cell fate. *Science* **363**, 1217–1222 (2019).
33. Seenundun, S. et al. UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J.* **29**, 1401–1411 (2010).
34. Faralli, H. et al. UTX demethylase activity is required for satellite cell–mediated muscle regeneration. *J. Clin. Invest.* **126**, 1555–1565 (2016).
35. Cesarani, A. et al. Investigation of genetic diversity and selection signatures between Sarda and Sardinian Ancestral black, two related sheep breeds with evident morphological differences. *Small Rumin. Res.* **177**, 68–75 (2019).
36. Laura, R. P., Ross, S., Koeppen, H. & Lasky, L. A. MAGI-1: A Widely expressed, alternatively spliced tight junction protein. *Exp. Cell Res.* **275**, 155–170 (2002).
37. Kantar, D. et al. MAGI1 inhibits the AMOTL2/p38 stress pathway and prevents luminal breast tumorigenesis. *Sci. Rep.* **11**, 5752 (2021).
38. Ben Jemaa, S. et al. A genome-wide comparison between selected and unselected Valle del Belice sheep reveals differences in population structure and footprints of recent selection. *J. Animal Breed. Genet.* **140**, 558–567 (2023).
39. Cheviron, Z. A. & Brumfield, R. T. Genomic insights into adaptation to high-altitude environments. *Heredity* **108**, 354–361 (2012).
40. Sun, Y.-B. et al. Species groups distributed across elevational gradients reveal convergent and continuous genetic adaptation to high elevations. *Proc. Natl. Acad. Sci.* **115**, E10634–E10641 (2018).
41. Yang, J. et al. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Mol. Biol. Evol.* **33**, 2576–2592 (2016).
42. Zhou, J. et al. Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. *EMBO J.* **24**, 635–644 (2005).
43. Yu, L. et al. Genomic analysis of snub-nosed monkeys (Rhinopithecus) identifies genes and processes related to high-altitude adaptation. *Nat. Genet.* **48**, 947–952 (2016).
44. Hu, X.-J. et al. The genome landscape of Tibetan sheep reveals adaptive introgression from Argali and the history of early human settlements on the Qinghai-Tibetan Plateau. *Mol. Biol. Evol.* **36**, 283–303 (2019).
45. Li, X. et al. Whole-genome resequencing to investigate the genetic diversity and mechanisms of plateau adaptation in Tibetan sheep. *J. Animal Sci. Biotechnol.* **15**, 164 (2024).

46. Cho, C. H. et al. ANKS1B encoded AIDA-1 regulates social behaviors by controlling oligodendrocyte function. *Nat. Commun.* **14**, 8499 (2023).
47. Parker, C. C. et al. Fine-mapping alleles for body weight in LG/J × SM/J F2 and F34 advanced intercross lines. *Mamm. Genome* **22**, 563–571 (2011).
48. Costa-Urrutia, P. et al. Genome-wide association study of body mass index and body fat in Mexican-Mestizo Children. *Genes* **10**, 945 (2019).
49. Lee, S.-I., Kim, S.-K. & Kang, S.-W. Genetic variants associated with body mass index changes in Korean Adults: The Anseong and Ansan Cohorts of the Korean Genome and Epidemiology Study. *Curr. Issues Mol. Biol.* **46**, 9074–9081 (2024).
50. Li, D. et al. Genomic selection and WssGWAS of sheep body weight and milk yield: imputing low-coverage sequencing data with similar genetic background panels. *J. Dairy Sci.* **S0022–0302**(24), 01440–01441. https://doi.org/10.3168/jds.2024-25681 (2025).
51. Arjunan, A., Sah, D. K., Woo, M. & Song, J. Identification of the molecular mechanism of insulin-like growth factor-1 (IGF-1): A promising therapeutic target for neurodegenerative diseases associated with metabolic syndrome. *Cell Biosci.* **13**, 16 (2023).
52. Sutter, N. B. et al. A Single IGF1 allele is a major determinant of small size in dogs. *Science* **316**, 112–115 (2007).
53. Yoshida, T. & Delafontaine, P. Mechanisms of IGF-1-mediated regulation of skeletal muscle hypertrophy and atrophy. *Cells* **9**, 1970 (2020).
54. Bedford, D. C., Kasper, L. H., Fukuyama, T. & Brindle, P. K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9–15 (2010).
55. Tan, X. L. et al. p300 expression is induced by oxygen deficiency and protects neuron cells from damage. *Brain Res.* **1254**, 1–9 (2009).
56. Svensson, K. et al. p300 and cAMP response element-binding protein-binding protein in skeletal muscle homeostasis, contractile function, and survival. *J. Cachexia. Sarcopenia Muscle* **11**, 464–477 (2020).
57. Center for Wilderness Safety. Oxygen levels at altitude. https://wildsafe.org/resources/ask-the-experts/altitude-safety-101/oxygen-levels/ (accessed 2 Jul 2025).
58. Yonekawa, T. et al. Large1 gene transfer in older MYD mice with severe muscular dystrophy restores muscle function and greatly improves survival. *Sci. Adv.* **8**, eabn0379 (2022).
59. Kanagawa, M. et al. Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* **117**, 953–964 (2004).
60. Loftus, S. K. et al. Acinar cell apoptosis in Serpini2-deficient mice models pancreatic insufficiency. *PLoS Genet.* **1**, e38 (2005).
61. Karakoyun, M. et al. Our cases with sucrase isomaltase deficiency. *J. Gastrointestinal Digest. Syst.* **5**, 1–4 (2015).
62. Yoshihara, M. & Montana, E. S. The synaptotagmins: Calcium sensors for vesicular trafficking. *Neuroscientist* **10**, 566–574 (2004).
63. Joëls, M. & Baram, T. Z. The neuro-symphony of stress. *Nat. Rev. Neurosci.* **10**, 459–466 (2009).
64. Ghaffar, A. et al. Variants of NAV3, a neuronal morphogenesis protein, cause intellectual disability, developmental delay, and microcephaly. *Commun. Biol.* **7**, 831 (2024).
65. Li, J. et al. Changes of collagen content in lung tissues of plateau yak and its mechanism of adaptation to hypoxia. *PeerJ* **12**, e18250 (2024).
66. Sadakata, T. et al. The secretory granule-associated protein CAPS2 regulates neurotrophin release and cell survival. *J. Neurosci.* **24**, 43–52 (2004).
67. Gold, S. M. et al. Basal serum levels and reactivity of nerve growth factor and brain-derived neurotrophic factor to standardized acute exercise in multiple sclerosis and controls. *J. Neuroimmunol.* **138**, 99–105 (2003).
68. Van Cutsem, J. et al. The influence of a mild thermal challenge and severe hypoxia on exercise performance and serum BDNF. *Eur. J. Appl. Physiol.* **115**, 2135–2148 (2015).
69. Das, S. K. et al. High altitude with monotonous environment has significant impact on mood and cognitive performance of acclimatized lowlanders: Possible role of altered serum BDNF and plasma homocysteine level. *J. Affect. Disord.* **237**, 94–103 (2018).
70. Nillni, E. A. Regulation of the hypothalamic thyrotropin releasing hormone (TRH) neuron by neuronal and peripheral inputs. *Front. Neuroendocrinol.* **31**, 134–156 (2010).
71. Ayalew, W. et al. Whole-genome sequencing of copy number variation analysis in Ethiopian cattle reveals adaptations to diverse environments. *BMC Genom.* **25**, 1088 (2024).
72. Edea, Z., Dadi, H., Dessie, T. & Kim, K.-S. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes Genom.* **41**, 973–981 (2019).
73. Stremming, J. et al. IGF-1 infusion to fetal sheep increases organ growth but not by stimulating nutrient transfer to the fetus. *Am. J. Physiol. Endocrinol. Metab.* **320**, E527–E538 (2021).
74. Sharma, V., Varshney, R. & Sethy, N. K. Human adaptation to high altitude: A review of convergence between genomic and proteomic signatures. *Hum. Genom.* **16**, 21 (2022).
75. Montoya-Williams, D. et al. Associations between maternal prenatal stress, methylation changes in IGF1 and IGF2, and birth weight. *J. Dev. Orig. Health Dis.* **9**, 215–222 (2018).
76. Hogenesch, J. B. et al. The basic helix-loop-helix-PAS protein MOP9 is a brain-specific heterodimeric partner of circadian and hypoxia factors. *J. Neurosci.* **20**, RC83 (2000).
77. Körner, C. The use of 'altitude' in ecological research. *Trends Ecol. Evol.* **22**, 569–574 (2007).

Author contributions

Conceptualization, Investigation, S.M., G.S., and A.C.; Methodology, S.B.J., A.C., and S.M.; Software, S.B.J.; Validation, S.M., G.S., and A.C.; Formal analysis, S.B.J., A.C., and S.M.; Resources, S.M., G.S., and A.C.; Data curation, F.C., S.R., C.D., C.P., and B.P.; writing—original draft preparation, S.B.J., S.M.; writing—review and editing, F.C., S.R., C.D., C.P., B.P., G.S., and A.C.; supervision, S.M., and A.C.; project ad-ministration, A.C.; funding acquisition, S.M., G.S., and A.C.

Funding

This study was supported by the project "Genomic Study of the resilience background in Mediterranean sheep breeds" (GEREMES, CUP J53D23013860001) within the "Finanziamento dell'Unione Europea—NextGenerationEU—missione 4, componente 2, investimento 1.1

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-14767-6.

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

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

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