BMC Genomics

RESEARCH Open Access

Genomic signals on the X chromosome reveal local adaptations in Ethiopian cattle

Wondossen Ayalew[1], Getinet M. Tarekegn[2,3*], Wu Xiaoyun[4,5], Min Chu[4,5], Rakan Naboulsi[6], Tesfaye S. Tessema[3], Erik Bongcam-Rudloff[7], Enyew Negussie[8], Yan Ping[4,5,9*] and Zhe Zhang[1*]

Abstract

Background In livestock, understanding the genetic basis of adaptation to the environment is essential for enhancing resilience to climate change and sustaining productivity in diverse environments. Indigenous Ethiopian cattle represent an ideal model for such studies, as they have evolved across a wide range of environments from the cool, oxygen-limited highlands to the hot, pathogen-rich lowlands. These environmental gradients imposed intense selective pressures, shaping their genomic landscape. In this study, we performed the first comprehensive analysis of X-linked adaptive signatures in Ethiopian indigenous cattle using whole-genome sequencing data.

Results Population structure analysis revealed clear genetic differentiations between Abigar and Barka cattle, while the remaining populations showed substantial shared ancestry and admixtures. Pairwise fixation index (F_{ST}) estimates, runs of homozygosity (ROH) patterns, and linkage disequilibrium (LD) decay further supported historical isolation and stronger selection pressure in Barka, contrasting with the greater diversity and faster LD decay in Gojjam Highland cattle. Complementary selection signature detections (F_{ST}, XP-EHH, and iHS) revealed population-specific and shared genomic regions under selection on the X chromosome. Notably, signals associated with high-altitude adaptation were detected near the *RBM3*, *RPS4X*, and *TSC22D3* loci. Additional signals were observed in genes related to thermoregulation and oxidative stress response (*EDA*, *SUV39H1*, and *HDAC8*), as well as immune regulation (*IRAK1*, *BDA20*, and *IL1RAPL1*), suggesting adaptation to hot and pathogen-rich environments. Functional enrichment analysis highlighted genes involved in extracellular matrix organization and immune signaling pathways, underscoring their roles in environmental adaptation.

Conclusions This study provides the first genome-wide evidence of X-linked adaptive divergence in the Ethiopian cattle. The findings highlight the contribution of the X chromosome to heat tolerance, hypoxia adaptation, and immune resilience, offering valuable genomic insights for breeding programs aimed at enhancing productivity and climate adaptability in tropical cattle.

Keywords Adaptation, Ethiopia, Indigenous cattle, and X-linked signature

*Correspondence:
Getinet M. Tarekegn
Getinet.Tarekegn@sruc.ac.uk
Yan Ping
pingyanlz@163.com
Zhe Zhang
zhezhang@scau.edu.cn

Full list of author information is available at the end of the article

Background

African cattle represent the most genetically diverse and environmentally resilient livestock populations on the continent. They are evolved across a wide range of agroecological zones, from high-altitude plateaus and arid deserts to humid lowlands. As a consequence, they have developed unique adaptations that enable them to survive and thrive under harsh conditions, including extreme heat, limited water and feed availability, endemic diseases, and vector-borne infections [1, 2]. These remarkable potentials reflect complex evolutionary histories shaped by natural selection and introgression from different ancestral lineages, including taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle [3, 4].

Ethiopia, in particular, represents a unique hotspot of cattle diversity, harboring the largest cattle population in Africa [5]. Owing to its location at the crossroads of historical cattle migration routes, Ethiopian cattle have experienced extensive admixture and exhibit remarkable adaptive and productive capacities [6, 7]. Autosomal studies have identified numerous genomic regions under selection, associated with key adaptive traits such as trypanotolerance, heat tolerance, immune response, resistance to endemic parasites and vector-borne diseases, and hypoxia [8–13]. These efforts have significantly advanced our understanding of the adaptive genetic architecture of Ethiopian cattle; however, they have disproportionately focused on the autosomal genome, mainly overlooking the contribution of sex chromosomes. A single Y-chromosome study in Ethiopian cattle has revealed substantial paternal genetic diversity, emphasizing the role of Y-lineages in tracing domestication and migration history [7]. Addressing this gap will provide a more complete picture of genetic diversity, population history, and the evolution of adaptive attributes in Ethiopian cattle populations.

The X chromosome, which diverged from ancestral autosomes around 300 million years ago [14], plays a crucial role in shaping genetic diversity and adaptive evolution in mammals [15]. Its unique inheritance pattern, reduced effective population size, and hemizygosity in males confer a distinct evolutionary dynamics [16]. Enriched for genes involved in reproduction, immune function, and other fitness-related traits, the X chromosome is particularly sensitive to selection pressures. It provides a valuable genomic target for uncovering novel selection signatures that may be overlooked in autosome-focused analyses [16–18]. X-linked genes can also experience accelerated evolution due to differential selective pressures in males and females, in contrast to the more balanced selection acting on autosomal genes [19]. Given these characteristics, analyzing patterns of X-linked genetic variation offers a unique opportunity to detect selection signatures and adaptive variants that would remain hidden in analyses restricted to autosomal markers.

Presence of important selection footprints on the X chromosome has been reported in several mammals, including cattle [15], horse [18], human [19], mastiff [20], pigs [21], and sheep [22], highlighting its pivotal role in driving adaptation and trait variation across diverse taxa. Despite these efforts, the X chromosome remains unexplored, particularly in Ethiopian cattle, leaving a significant gap in understanding the full landscape of their genetic diversity and adaptive evolution in these populations. In this study, we characterized patterns of genetic diversity and signals of selection on the X chromosome using whole-genome sequencing data from seven indigenous Ethiopian cattle populations reported in our previous study [23]. Using X-linked variants, we identify genomic regions shaped by adaptive evolution and historical domestication pressures. The results reveal population-specific and shared adaptive footprints, highlighting the unique role of the X chromosome in shaping traits associated with local adaptations in tropical environments. These findings provide valuable genomic insights to support sustainable genetic improvement strategies of cattle facing accelerating climatic and disease challenges.

Materials and methods

Cattle populations, variants, and quality control

We used the publicly available genomic data in variant call format (VCF) for seven Ethiopian cattle populations deposited in the European Variation Archive (EVA) under Project PRJEB75238 [23]. The dataset comprised Abigar (ABI), Barka (BAR), Boran (BOR), Fellata (FEL), Fogera (FOG), Gojam Highland (GH), and Horro (HOR), representing animals sampled across diverse agro-ecological zones in Ethiopia (Additional file 1: Table S1). All sampled individuals were females, and X chromosome genotypes were therefore treated as diploid in downstream allele-frequency and haplotype-based analyses. Detailed descriptions of sampling locations, population background, and SNP-calling procedures are provided in our previous study [23]. From the full variant dataset, SNPs located on the X chromosome were extracted, resulting in a final set of 919,141 variants (Fig. 1A and B). Quality control (QC) was performed using PLINK v2.0 [24] by removing SNPs with a minor allele frequency (MAF) < *0.01*, a genotype call rate < 90%, and deviating from Hardy–Weinberg equilibrium at $p < 1 \times 10^{-6}$. Individuals with more than 10% missing genotypes were excluded from the analysis. After QC, 784,822 SNPs remained for downstream analyses.

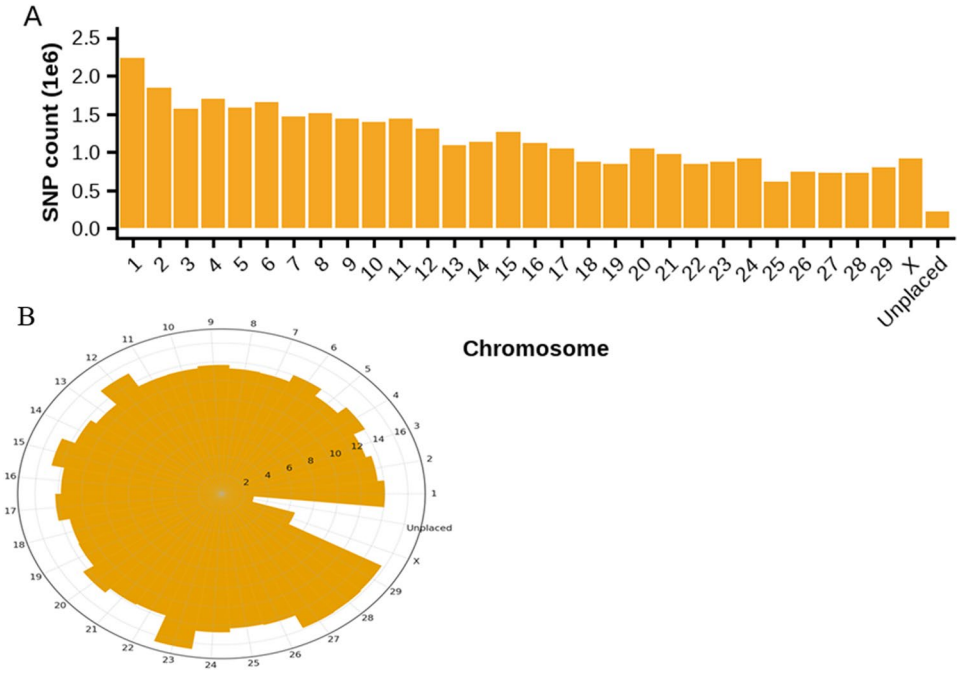

Fig. 1 Whole-genome sequence SNP count and density across all chromosomes in seven Ethiopian cattle. (A) SNP count and (B) SNP density per kb. These plots were adapted from the tabular analysis of Ayalew et al. [23] to include SNP distribution on the X chromosome alongside autosomes

Genetic relationship and population structure

After QC, to minimize highly linked variants for population differentiation analysis, LD pruning was performed in PLINK v2.0 using the parameters *--indep-pairwise* 50 5 0.2, resulting in 68,029 SNPs retained for subsequent population structure analyses. Principal component analysis (PCA) was conducted using the same software. A model-based clustering approach implemented in ADMIXTURE v1.3.0 [25] was then applied to infer ancestral population structure. The number of genetic clusters (K) was tested from 2 to 7, and the optimal value was determined based on the lowest cross-validation (CV) error. PC and admixture results were visualized using the R package "ggplot2" [26]. To further explore phylogenetic relationships, a neighbor-joining (NJ) tree was constructed based on pairwise genetic distances calculated in PLINK v2.0. The NJ tree was generated in MEGA X [27] with 1,000 bootstrap replicates and visualized using iTOL v6 [28].

Genomic diversity and differentiation

To explore the genomic variation within and between populations, four complementary statistics, including nucleotide diversity (π), LD decay, runs of homozygosity (ROH), and pairwise genetic differentiation (F_{ST}) were calculated based on high-quality SNPs from the X chromosome. Pairwise F_{ST} and π were estimated using VCFtools v0.1.16 [29] with a 50 kb window and a 25 kb sliding step. LD decay was computed as the squared correlation coefficient (r^2) between SNP pairs using PLINK v2.0. ROH were detected using PLINK v2.0 to identify continuous homozygous segments across the genome and to assess recent inbreeding patterns. ROH segments were classified into three length categories: short (0.5–1 Mb), medium (1–5 Mb), and long (> 5 Mb) [30].

Selective sweep analysis

To identify genomic regions under selection on the X chromosome, we employed complementary approaches capturing both allele frequency differentiation and haplotype-based signals. The analysis focused on contrasting populations adapted to distinct ecological conditions: GH cattle from high-altitude environments (~ 4,000 m above sea level (m.a.s.l.)) and the lowland ABI and BAR populations, which inhabit hot, arid regions below ~ 895 m.a.s.l [23]. Population differentiation was estimated using Weir and Cockerham's F_{ST} [31], implemented in VCFtools [29]. The analysis was conducted in sliding windows of 50 kb with a 25 kb step size across the X chromosome for the pairwise comparisons GH vs. ABI and GH vs. BAR. For haplotype-based analyses, genotype data for each population were phased and imputed using Beagle v5.1 [32]. Two complementary haplotype-based statistics were then calculated using selscan [33]: the Integrated Haplotype Score (iHS), which detects partial selective sweeps within populations based on extended haplotype homozygosity around focal selection sweeps [34], and the Cross-Population Extended Haplotype Homozygosity (*XP-EHH*), which contrasts haplotype decay between populations to identify population-specific sweeps [35].

For iHS, the statistic was computed as the standardized natural logarithm of the ratio between the integrated haplotype homozygosity (iHS) of the derived and ancestral alleles, following genome-wide normalization as described by Voight et al. [34]. Absolute $|iHS|$ values were averaged in non-overlapping 50 kb windows, and windows in the top 0.5% were considered candidate regions under selection for each studied cattle populations (i.e., ABI, BAR, GH). Using the same contrasting cattle populations, *XP-EHH* was computed as the standardized log-ratio of integrated haplotype homozygosity between two populations ($iHHA/iHHB$), with average |*XP-EHH*| values [35] calculated for each 50 kb window. The upper 0.5% of windows was used to define genomic regions showing strong inter-population selection signals. Finally, genome-wide distributions of $|iHS|$ and |*XP-EHH*| scores were visualized in Manhattan plots using ggplot2 [26]. Shared signals were identified and visualized in a Venn diagram using ggplot2 [26]. Finally, *Tajima's D* and F_{ST} statistics were calculated for candidate genes identified by at list two selection signature method using VCFtools [29].

Functional annotation of candidate regions

To explore the biological relevance of genomic regions under selection, the top 0.5% of candidate SNPs identified across all selection signature analyses were functionally annotated. SNP coordinates were aligned to the latest *Bos taurus* reference genome (ARS-UCD1.2) [36] using BEDTools v2.25.0 [37]. Overlapping genes were merged, and unique loci were retained for downstream analyses. Functional enrichment of these genes was evaluated using the Database for Annotation, Visualization and Integrated Discovery (DAVID, v6.8) [38] to provide additional insights into associated biological processes, cellular components, and molecular functions. Gene enrichment p-values were adjusted using the Benjamini–Hochberg procedure to control the false discovery rate (FDR < 0.05).

Results

Genetic relationship and population structure

The first two principal components (PC1 and PC2) together explained 5.43% of the total genetic variance, with PC1 accounting for 3.04% and PC2 accounting for 2.39%, respectively (Fig. 2A). PC1 clearly separated the ABI population from the BAR population, reflecting their distinct geographic or ancestral origins. In contrast, the remaining five populations clustered closely in the central region of the PCA plot, indicating a high degree of admixture. Consistent with the PCA result, the admixture analysis at K = 2 showed a similar pattern (Fig. 2B), as supported by the lowest cross-validation error (Additional file 2: Figure S1). Likewise, the neighbor-joining tree revealed clustering patterns similar to those observed in PCA and admixture results (Fig. 2C). This pattern suggests shared ancestry or ongoing gene flow among these populations.

Genetic diversity among the population

Nucleotide diversity (π) exhibited population-specific distribution patterns, but these differences were not significant. All seven populations had low median π values, with no clear distinctions in overall genetic diversity

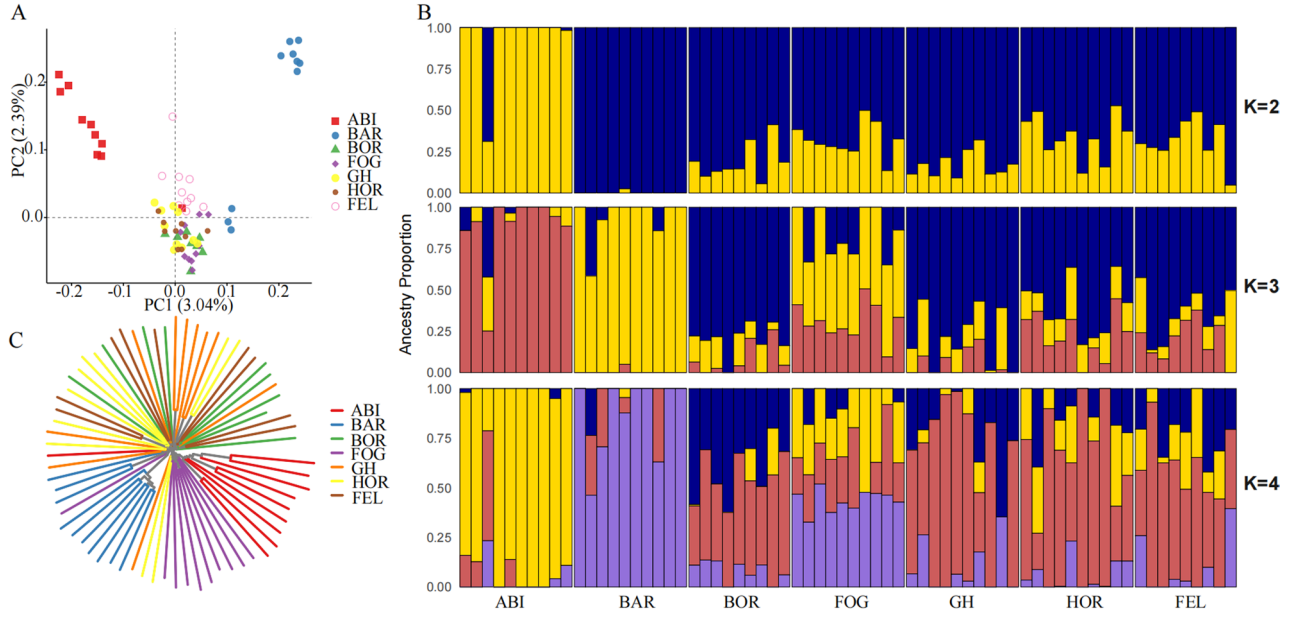

Fig. 2 Genetic diversity and population structure of the Ethiopian cattle population. **A** PCA plot, **B** Admixture and ancestry proportions, and **C** Neighbor-joining tree

(Fig. 3A). Genetic differentiation between populations is presented in Fig. 3B. BAR was highly differentiated from ABI (mean F_{ST} = 0.087), which is consistent with our population structure results. The patterns of total ROH length varied across populations (Fig. 3C), reflecting differences in inbreeding and selection history. BAR had the longest total ROH, dominated by segments of 1–5 Mb. In contrast, FEL had the shortest total ROH, with a higher proportion of 0.5–1 Mb segments. The LD decay on the X chromosome revealed population-specific patterns (Fig. 3D). BAR maintained the slowest LD decay, with r^2 remaining above 0.13 even at 500 kb, indicating long-range LD and potential recent selection or low effective population size. In contrast, GH and FEL showed faster LD decay, with r^2 dropping below 0.12 at 500 kb.

X-linked selection signatures in Ethiopian cattle

Genome-wide selection scans targeting the X chromosome were performed to identify genomic regions under selection between highland and lowland cattle populations. ABI was treated as a low-altitude cattle population, while the GH was treated as a high-altitude cattle population. Between ABI and GH cattle, a total of 49 genomic regions, 31 candidate genes were detected using F_{ST} and 35 regions with 26 protein-coding genes by *XP-EHH* (Additional file 1: Tables S2-S3). The within-population analysis (iHS) revealed 76 genomic regions, 26 candidate genes under selection in ABI, and 66 candidate regions, as well as 20 protein-coding genes in GH (Additional file 1: Tables S4-S5). The Manhattan plots for the three selection signature methods revealed multiple distinct selection signals distributed across the X chromosome (Fig. 4A–D). Several candidate genes were overlapped at least in two selection scan methods and located near the top SNPs, including *EDA*, *BDA20*, *RPS4X*, *RBM3*, *IRAK1*, *SH2D1A*, *IL1RAPL1*, and *IL1RAPL2*, which may be associated with local environmental adaptations (Fig. 5A, Additional file 1: Tables S2–S5). F_{ST} and *Tajima's D* statistics revealed that the *IRAK1* gene is differentiated between ABI and GC (Fig. 5C).

Likewise. the selection sweep test was extended to BAR cattle, a population that thrives in the hot, arid lowlands of northwestern Ethiopia, in contrast to the GH population. Genome-wide selection scans identified 61

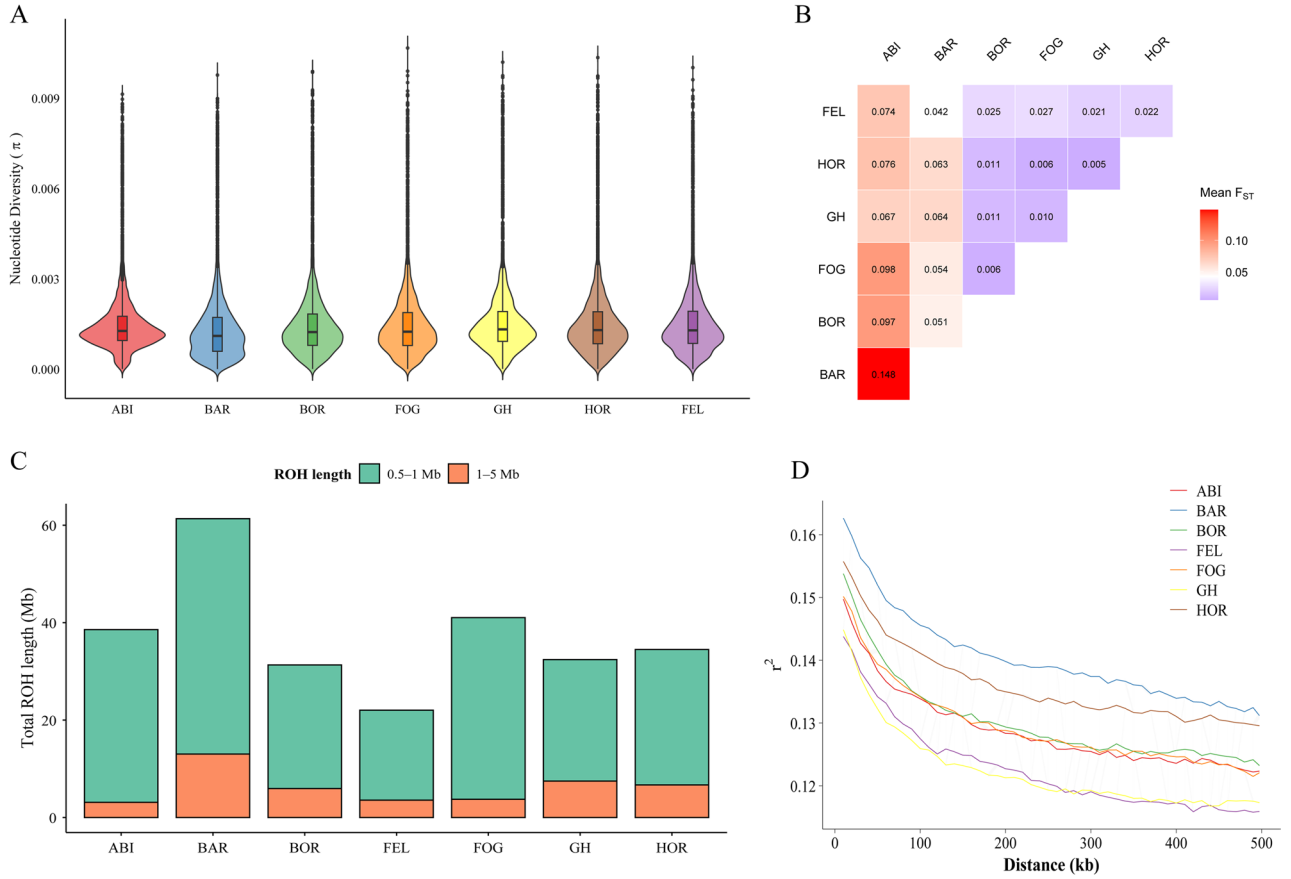

Fig. 3 Population genetic diversity, inbreeding patterns, and LD decay across the Ethiopian indigenous cattle population. **A** Violin plots illustrating nucleotide diversity (π), **B** Pairwise population differentiation based on mean F_{ST} values, **C** Total length of runs of homozygosity (ROH) per individual across populations, categorized into short (0.5–1 Mb) and medium (1–5 Mb) segments, and (**D**) Linkage disequilibrium (LD) decay curves showing the decline of r^2 with increasing physical distance between SNPs

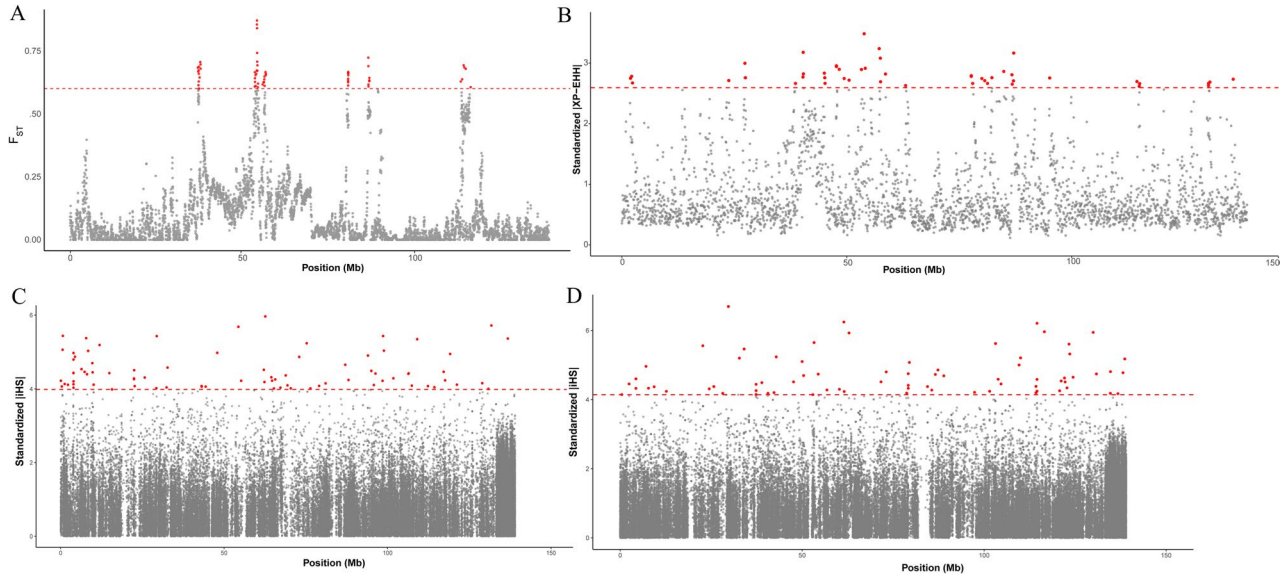

Fig. 4 Manhattan plots of ABI and GH cattle. **A** F_{ST} and (**B**) XP-EHH (**C**) iHS of ABI and (**D**) iHS of GH

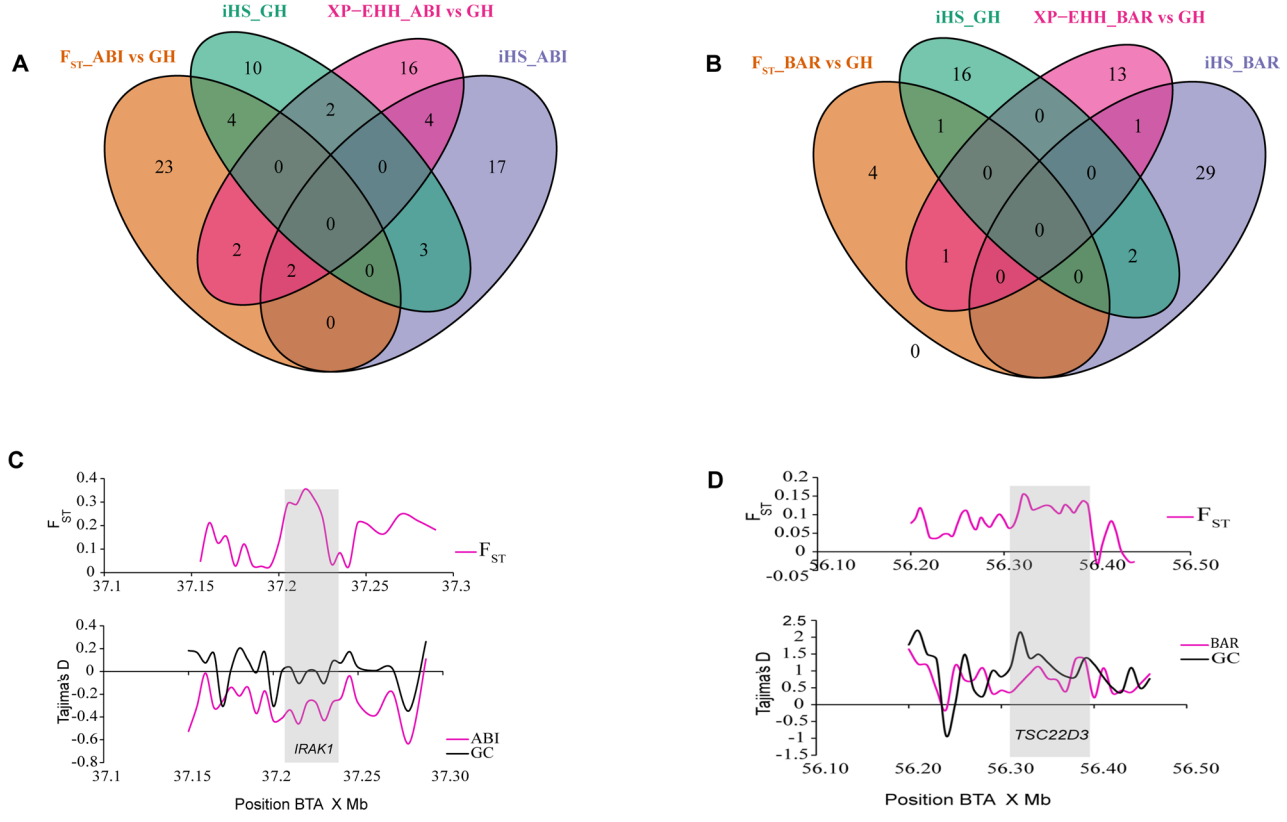

Fig. 5 Overlapping protein-coding genes of three signatures of selection methods and local gene region scans. **A** Venn diagram of genomic regions under selection in ABI vs. GH cattle, identified by F_{ST} , iHS and *XP-EHH* (**B**) Venn diagram of genomic regions under selection in BAR vs. GH cattle, identified by F_{ST} , iHS and *XP-EHH* (**C**) Local scan of selection signatures at the *IRAK1* locus in ABI cattle, detected by F_{ST} and *Tajima's D*, (**D**) Local scan of selection signatures at the *TSC22D3* locus in BAR cattle, detected by F_{ST} and *Tajima's D*

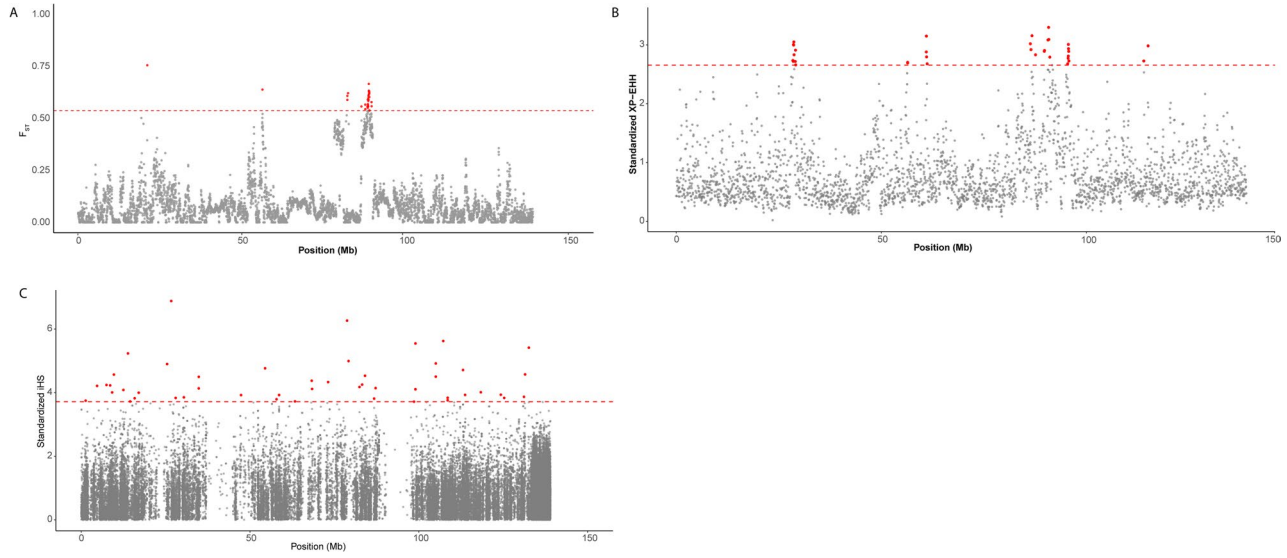

Fig. 6 Manhattan plots of BAR and GH cattle. (A) F_{ST} BAR vs. GC, (B) *XP-EHH* BAR vs. GH, and (C) iHS of BAR

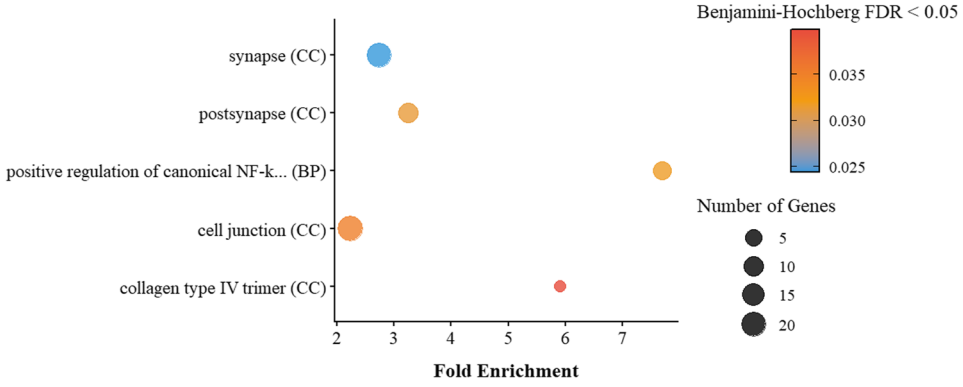

Fig. 7 Dot-plot illustrating the top functionally enriched terms identified by gene ontology (GO) and pathway analysis, filtered by Benjamini-Hochberg false discovery rate (FDR) < 0.05. The x-axis represents fold enrichment (a measure of how overrepresented a term is relative to expectation), and the y-axis lists enrichment terms, with their GO category (cellular component [CC] or biological process [BP]) indicated in parentheses

candidate regions, (20 protein-coding genes) with FST and 32 regions, (15 protein-coding genes) with XP-EHH, (Additional File 1: Table S6-S7). Within-population selections scan uncovered 46 iHS signals annotated to 18 protein-coding genes in BAR, (Additional File 1: Table S8). Notably, candidate genes such as *HDAC6*, *RBM3*, *SUV39H1*, and *TSC22D3* were positioned near the strongest SNPs among the three selection scan methods, (Fig. 5B). The F_{ST} and Tajima's D statistics revealed differentiation between BAR and GH cattle at the *TSC22D3* gene region, (Fig. 5C). The Manhattan plots of the three complementary selection signatures are presented in Figs. 6 A, B, and C.

Functional annotation of the candidate genes

Functional enrichment analysis, corrected for multiple testing (FDR < 0.05), identified five significant terms (Fig. 7). Prominent enrichments included "*synapse*," "*postsynapse*," and "*cell junction*," indicating robust involvement in cellular communication and tissue structural integrity. Additionally, the "*collagen type IV trimer*" was enriched, suggesting roles in extracellular matrix organization. In the signaling pathways, "*positive regulation of canonical NF-κB signal transduction*" was significantly overrepresented, implicating activation of immune and inflammatory processes. Details are provided in Additional file 2: Table S9.

Discussion

The X chromosome plays a unique role in mammalian evolution. To fill the existing gap on X-linked analysis in Ethiopian cattle, we explored the population structure and the genetic basis of local adaptation. Population structure analyses revealed a clear divergence between the ABI and BAR populations, indicating limited recent gene flow. In contrast, the remaining groups formed a

tight cluster, reflecting substantial shared ancestry and historical connectivity. This pattern conforms to previous autosomal studies [6, 13]. Thus, suggests a shared demographic history for most populations and unique adaptive histories, ultimately influencing current Ethiopian cattle genetic structure. Moreover, our analysis revealed several X-linked selection signals shaped by long-term evolutionary pressures, highlighting the potential contribution of the X chromosome to key adaptive processes, including heat tolerance, high-altitude adaptations, and robust immune function in Ethiopian cattle.

Adaptive selection signals in response to environmental stresses

High-altitude adaptations

High-altitude environments impose unique physiological challenges, particularly chronic hypoxia, which exerts intense selective pressures on genomic regions associated with oxygen transport, energy metabolism, and cellular stress responses. Previous analyses based on autosomal SNP-derived copy number variations [9] and selection signature studies [8] revealed potential genetic adaptations in Ethiopian high-altitude cattle, including the GH cattle population. In line with these findings, the current X chromosome–based analysis identified additional candidate genes, including *RBM3*, *RPS4X*, and *TSC22D3*, which may contribute to physiological processes potentially advantageous under high-altitude conditions through their roles in cellular stress regulation, neuroprotection, and translational control. Notably, *RBM3* is a highly conserved, cold, and stress-inducible RNA-binding protein that plays critical roles in cell survival, proliferation, and homeostasis under adverse conditions [39]. It is rapidly upregulated by cold exposure [40, 41] and also responds to hypoxia independently of *HIF-1* signaling [42, 43]. In neural stem cells and pulmonary endothelial cells, *RBM3* alleviates hypoxia-induced cell cycle arrest and apoptosis, maintaining cellular proliferation and viability [43, 44]. Under low-temperature and hypoxic environments, *RBM3* mechanistically stabilizes mRNAs, modulates protein translation, and interacts with stress-response pathways to confer cytoprotection [40, 45]. Likewise, *RPS4X*, an X-linked ribosomal protein, forms part of the 40 S ribosomal subunit and plays a vital role in maintaining translational capacity under cellular stress. Recent RNA-seq analyses in human pulmonary microvascular endothelial cells revealed that females express higher levels of *RPS4X* than males under both normoxic and hypoxic conditions, suggesting a role in sex-linked adaptation to low-oxygen environments [46]. Although direct studies of *RPS4X* under hypoxic conditions are limited, a highly similar peptide, *RPS4XL*, encoded by *lnc-Rps4l*, has been shown to interact with ribosomal protein S6 and to inhibit the proliferation of pulmonary artery smooth muscle cells under hypoxic stress [47]. These observations support the hypothesis that *RPS4X* may contribute to cellular resilience under hypoxia, making it a candidate gene for high-altitude adaptation. The *TSC22D3* gene, also known as glucocorticoid-induced leucine zipper (*GILZ*), encodes a transcription factor involved in cellular stress responses, immune regulation, and cell survival. In mammalian models, *TSC22D3* is induced under hypoxic conditions via *HIF1α*-dependent pathways, facilitating survival in low-oxygen environments by modulating transcriptional networks, including *SREBP1* signaling [48]. Furthermore, *TSC22D3* regulates inflammatory responses by inhibiting *NF-κB/NLRP3* signaling and macrophage polarization toward the M1 pro-inflammatory state, suggesting a role in maintaining tissue homeostasis under stress [49]. These stress-response and immunomodulatory functions indicated that the gene may provide a molecular mechanism, enabling livestock to cope with the stressful conditions of mountainous regions.

Hot environments adaptation

In hot environments, where heat stress imposes intense selective pressures, adaptive alleles on the X chromosome may play a role in shaping population resilience. Among the candidate genes, *EDA*, *SUV39H1*, and *HDAC6* emerged as promising candidates due to their plausible biological roles in thermal stress adaptation, relevant to ABI and BAR cattle. Ectodysplasin A (*EDA*) is a key regulator of ectodermal development, controlling the formation of sweat glands, hair follicles, and other skin-derived structures. These ectodermal components contribute to thermoregulation, sensory perception, and protection against environmental stressors, thereby playing an essential role in maintaining physiological homeostasis [50, 51]. Deficiencies in the *EDA–EDAR* signaling pathway lead to hypohidrotic ectodermal dysplasia, a condition marked by a reduced number of sweat glands, sparse hair growth, and impaired evaporative cooling, resulting in heat intolerance [52–54]. These anatomical and physiological links highlight the gene's crucial role in thermoregulation by improving heat dissipation. *HDAC8* encodes a class I histone deacetylase whose enzymatic activity is directly regulated by a redox-sensitive disulfide switch between cysteines C102 and C153, thereby linking oxidative stress to gene regulatory capacity [55]. Heat stress is known to elevate reactive oxygen species (ROS) and disrupt redox homeostasis, so allelic variation in *HDAC8* may underpin enhanced resilience by maintaining chromatin state and transcriptional control under thermal challenge.

Furthermore, *HDAC8* has been shown to increase under general stress conditions (e.g., UV) and to modulate protective gene networks, supporting its role in

cellular stress adaptation [56]. *SUV39H1* encodes a histone H3 lysine 9 methyltransferase that establishes the heterochromatic mark *H3K9me3* and is critical for chromatin compaction and transcriptional silencing [57]. Notably, heat-stress experiments in human cells revealed an immediate enrichment of *H3K9me3* at alpha-satellite repeats following exposure to elevated temperature, suggesting that heterochromatin dynamics are responsive to thermal stress [58]. Thus, selection on *SUV39H1* in heat-adapted cattle may reflect enhanced capacity to reorganize heterochromatin and stabilize genome architecture under prolonged heat exposure, which helps to maintain cellular integrity and transcriptional homeostasis.

Immune response

In hot and humid environments, cattle are continuously exposed to a high microbial load and environmental stressors that challenge immune homeostasis. The innate immune system provides the first line of defense, with *IRAK1* serving as a pivotal mediator through Toll-like receptor (*TLR*) and interleukin-1 receptor (*IL-1*R) signaling pathways [59]. Variations in *IRAK1* may influence the efficiency of pro-inflammatory signaling, including cytokines such as *TNF-α* and *IL-6*, which are essential for rapid pathogen clearance [60]. Functional studies indicate that *IRAK1* is required by dendritic cells (DCs) and macrophages to be fully activated, highlighting its role in early immune responses. In murine models, *IRAK1* deficiency in conventional DCs impairs *IL-12* production during Toxoplasma gondii infection [61], while in human DCs, *IRAK1* knockdown reduces maturation markers, pro-inflammatory cytokines, and T-cell proliferation [62]. These findings suggest that selection on *IRAK1* in ABI cattle may enhance innate immune responses under high microbial and environmental stress. Likewise, *BDA20* has emerged as a candidate linked to host defense against ectoparasites, particularly ticks. Ticks, such as *Rhipicephalus microplus*, are a significant source of stress for cattle in tropical and subtropical regions, where high temperatures and humidity favor the proliferation of these parasites. Variation in *BDA20* expression has been observed between resistant and susceptible animals, suggesting a role in modulating the immune response to tick infestations [63–65]. Enhanced expression of *BDA20* may support the activation of local inflammatory responses or promote cellular mechanisms that reduce parasite attachment and survival. Thus, *BDA20* represents a functional link between immunity and environmental adaptation in ABI cattle. Similarly, *IL1RAPL1* and *IL1RAPL2* are accessory proteins of the *IL-1* receptor family that can modulate *TLR/IL-1R*–driven innate signaling, a pathway central to early host defense against infection [66]. Activation of the *IL-1/TLR* signaling pathway triggers downstream *NF-κB* and *MAPK* cascades, driving the production of pro-inflammatory cytokines that are critical for rapid pathogen control in barrier tissues [67]. Although *IL1RAPL1* and *IL1RAPL2* were first studied in the nervous system, emerging functional data indicate they can influence local immune responses, supporting a role beyond neuronal biology [68]. In hot, humid low-altitude environments where pathogen pressure and heat stress combine, genetic variation near these loci could plausibly adjust innate responses, contributing to resilience; however, direct confirmation in cattle will require functional immunology follow-ups, and adaptive interpretations remain inferential, as pathogen load and microbial exposure were not directly quantified in the studied environments.

Study limitations

While this study provides novel insights into X chromosome–linked adaptation in Ethiopian cattle, several limitations should be acknowledged. The analyses were restricted to X-linked variants derived from female animals and did not include direct comparisons with autosomal variation. Functional validation of candidate genes was not feasible due to the absence of biological samples; therefore, associations between candidate loci and adaptive phenotypes remain inferential. Historical demographic processes, sex-biased gene flow, and variation in local recombination rate may influence haplotype-based statistics. In addition, linked purifying selection (background selection) may contribute to elevated haplotype-based signals. Although recombination-adjusted or simulation-based null models were not implemented, we reduced the risk of false positives by applying stringent empirical thresholds and requiring concordant signals across complementary methods and populations. Finally, environmental adaptation was inferred from altitude and ecological classification rather than from quantitative environmental variables such as temperature gradients, oxygen availability, or pathogen burden. Consequently, the detected signals likely reflect composite and multifactorial selective pressures rather than direct associations with single measured environmental factors. Future studies integrating fine-scale environmental data, recombination-aware modeling, and functional validation will further refine the interpretation of these candidate adaptive regions.

Conclusions

This study provides the first comprehensive insight into X-linked selection signatures in Ethiopian cattle, highlighting their pivotal role in local adaptation and population divergence. Integrating complementary selection signature detections revealed genomic regions associated with resistance to hypoxia, heat, and pathogen pressure. Notably, *EDA*, *IRAK1*, *BDA20*, and *TSC22D3* are among

the candidate genes involved in coordinating evolutionary responses that enable cattle to thrive under challenging environmental conditions. Together, these findings broaden our understanding of the adaptive landscape of Ethiopian cattle and emphasize the evolutionary importance of the X chromosome in shaping resilience to environmental challenges. The insights gained here offer a valuable genomic foundation for designing breeding strategies that enhance both productivity and adaptation to extreme environmental conditions in the face of climate variability.

Abbreviations

BP	Biological process
CC	Cellular component
CV	Cross-validation
DAVID	Database for Annotation, Visualization and Integrated Discovery
EVA	European Variation Archive
F_{ST}	Fixation index
GO	Gene ontology
His	Integrated Haplotype Score
KEGG	Kyoto Encyclopedia of Genes and Genomes
LD	Linkage disequilibrium
masl	Meter above sea level
NJ	Neighbor-joining
PCA	Principal component analysis
QC	Quality control
ROH	Runs of homozygosity
SNP	Single-nucleotide polymorphism
VCF	Variant calling format
XP-EHH	Cross-Population Extended Haplotype homozygosity
Π	Nucleotide diversity

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12817-5.

Supplementary Material 1: Additional File 1: Figure S1. Cross-validation error of admixture analysis K= 2 to 7.

Supplementary Material 2.

Acknowledgements

We thank the institutions involved in generating and releasing the whole-genome sequencing and VCF data used in our previous publications. We also acknowledge South China Agricultural University for providing the computational resources that supported this study.

Authors' contributions

W.A. and G.M.T. conceptualized and designed the research; W.A. performed data analysis with input from G.M.T.; W.A. interpreted the results; G.M.T., X.W., R.N., T.S.T., E.B.R., E.N., M.C., P.Y., and Z.Z. contributed to the interpretation of the results and critically revised the manuscript; W.A. drafted the manuscript. All authors read and approved the final version.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 32022078). Computational resources were provided by South China Agricultural University.

Data availability

The VCF data analyzed in this study are publicly accessible at the EVA under Project PRJEB75238.

Declarations

Ethics approval and consent to participate

The study utilized publicly available datasets previously generated by the authors. All original experiments were conducted under appropriate ethical approvals, as cited in the original publications.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]State Key Laboratory of Swine and Poultry Breeding Industry, Guangdong Provincial Key Lab of Agro-Animal Genomics and Molecular Breeding, College of Animal Science, South China Agricultural University, Guangzhou 510642, China
[2]School of Veterinary Medicine and Biosciences (SVMB), Scotland's Rural College (SRUC), Peter Wilson Building, The King's Buildings, West Mains Road, Edinburgh EH9 3JG, UK
[3]Institute of Biotechnology, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia
[4]Key Laboratory of Animal Genetics and Breeding on Tibetan Plateau, Ministry of Agriculture and Rural Affairs, Lanzhou 30050, China
[5]Key Laboratory of Yak Breeding Engineering, Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China
[6]American University of the Middle East, Egaila 54200, Kuwait
[7]Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala 75007, Sweden
[8]Natural Resources Institute Finland (Luke), Jokioinen 31600, Finland
[9]Institute of Western Agriculture, the Chinese Academy of Agricultural Sciences, Changji 831100, China

Received: 2 December 2025 / Accepted: 30 March 2026
Published online: 10 April 2026

References

1. Mwai O, Hanotte O, Kwon YJ, Cho S. African Indigenous Cattle: Unique Genetic Resources in a Rapidly Changing World. Asian-Australas J Anim Sci. 2015;28(7):911–21.
2. Kim J, Hanotte O, Mwai OA, Dessie T, Bashir S, Diallo B, Agaba M, Kim K, Kwak W, Sung S, et al. The genome landscape of indigenous African cattle. Genome Biol. 2017;18(1):34.
3. Gebrehiwot NZ, Strucken EM, Aliloo H, Marshall K, Gibson JP. The patterns of admixture, divergence, and ancestry of African cattle populations determined from genome-wide SNP data. BMC Genomics. 2020;21(1):869.
4. Friedrich J, Bailey RI, Talenti A, Chaudhry U, Ali Q, Obishakin EF, Ezeasor C, Powell J, Hanotte O, Tijjani A, et al. Mapping restricted introgression across the genomes of admixed indigenous African cattle breeds. Genet Sel Evol. 2023;55(1):91.
5. (CSA) CSA. The Federal Democratic Republic of Ethiopia Central Statistical Agency Agricultural Sample Survey Report on the Area, Production of Major Crops and Livestock and Their Products. In.; 2021/22.
6. Edea Z, Dadi H, Kim SW, Dessie T, Lee T, Kim H, Kim JJ, Kim KS. Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. Front Genet. 2013;4:35.
7. Li MH, Zerabruk M, Vangen O, Olsaker I, Kantanen J. Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. Heredity (Edinb). 2007;98(4):214–21.
8. Terefe E, Belay G, Han J, Hanotte O, Tijjani A. Genomic adaptation of Ethiopian indigenous cattle to high altitude. Front Genet. 2022;13:960234.
9. Ayalew W, Xiaoyun W, Tarekegn GM, Tessema TS, Chu M, Liang C, Naboulsi R, Van Damme R, Bongcam-Rudloff E, Ping Y. Whole-genome sequencing of copy number variation analysis in Ethiopian cattle reveals adaptations to diverse environments. BMC Genomics. 2024;25(1):1088.
10. Ayalew W, Wu X, Tarekegn GM, Sisay Tessema T, Naboulsi R, Van Damme R, Bongcam-Rudloff E, Edea Z, Enquahone S, Yan P. Whole-Genome

Resequencing Reveals Selection Signatures of Abigar Cattle for Local Adaptation. *Animals (Basel)* 2023, 13(20).
11. Bahbahani H, Afana A, Wragg D. Genomic signatures of adaptive introgression and environmental adaptation in the Sheko cattle of southwest Ethiopia. PLoS ONE. 2018;13(8):e0202479.
12. Mekonnen YA, Gultas M, Effa K, Hanotte O, Schmitt AO. Identification of Candidate Signature Genes and Key Regulators Associated With Trypanotolerance in the Sheko Breed. Front Genet. 2019;10:1095.
13. Terefe E, Belay G, Tijjani A, Han J, Hanotte O. Whole Genome Resequencing Reveals Genetic Diversity and Selection Signatures of Ethiopian Indigenous Cattle Adapted to Local Environments. Diversity 2023, 15(4): 540.
14. Lahn BT, Page DC. Four evolutionary strata on the human X chromosome. Science. 1999;286(5441):964–7.
15. Rajawat D, Panigrahi M, Nayak SS, Bhushan B, Mishra BP, Dutt T. Dissecting the genomic regions of selection on the X chromosome in different cattle breeds. 3 Biotech. 2024;14(2):50.
16. Harris M, Garud NR. Enrichment of Hard Sweeps on the X Chromosome in Drosophila melanogaster. Mol Biol Evol 2023, 40(1): msac268.
17. Ellegren H. The different levels of genetic diversity in sex chromosomes and autosomes. Trends Genet. 2009;25(6):278–84.
18. Liu XX, Pan JF, Zhao QJ, He XH, Pu YB, Han JL, Ma YH, Jiang L. Detecting selection signatures on the X chromosome of the Chinese Debao pony. J Anim Breed Genet. 2018;135(1):84–92.
19. Heyer E, Segurel L. Looking for signatures of sex-specific demography and local adaptation on the X chromosome. Genome Biol. 2010;11(1):203.
20. Wu H, Liu YH, Wang GD, Yang CT, Otecko NO, Liu F, Wu SF, Wang L, Yu L, Zhang YP. Identifying molecular signatures of hypoxia adaptation from sex chromosomes: A case for Tibetan Mastiff based on analyses of X chromosome. Sci Rep. 2016;6:35004.
21. Ma Y, Zhang H, Zhang Q, Ding X. Identification of selection footprints on the X chromosome in pig. PLoS ONE. 2014;9(4):e94911.
22. Zhu C, Fan H, Yuan Z, Hu S, Zhang L, Wei C, Zhang Q, Zhao F, Du L. Detection of Selection Signatures on the X Chromosome in Three Sheep Breeds. Int J Mol Sci. 2015;16(9):20360–74.
23. Ayalew W, Xiaoyun W, Tarekegn GM, Naboulsi R, Sisay Tessema T, Van Damme R, Bongcam-Rudloff E, Chu M, Liang C, Edea Z, et al. Whole genome sequences of 70 indigenous Ethiopian cattle. Sci Data. 2024;11(1):584.
24. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
25. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655–64.
26. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-; 2016.
27. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35(6):1547–9.
28. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.
29. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
30. Xu L, Zhao G, Yang L, Zhu B, Chen Y, Zhang L, Gao X, Gao H, Liu GE, Li J. Genomic Patterns of Homozygosity in Chinese Local Cattle. Sci Rep. 2019;9(1):16977.
31. Weir BS, Cockerham CC. Estimating F-Statistics for the Analysis of Population Structure. Evolution. 1984;38(6):1358–70.
32. Browning BL, Tian X, Zhou Y, Browning SR. Fast two-stage phasing of large-scale sequence data. Am J Hum Genet. 2021;108(10):1880–90.
33. Szpiech ZA, Hernandez RD. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol. 2014;31(10):2824–7.
34. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biol. 2006;4(3):e72.
35. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007;449(7164):913–8.
36. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, Rowan TN, Low WY, Zimin A, Couldrey C et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 2020, 9(3): giaa021.
37. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2.
38. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.
39. Zhou RB, Lu XL, Zhang CY, Yin DC. RNA binding motif protein 3: a potential biomarker in cancer and therapeutic target in neuroprotection. Oncotarget. 2017;8(13):22235–50.
40. Hu Y, Liu Y, Quan X, Fan W, Xu B, Li S. RBM3 is an outstanding cold shock protein with multiple physiological functions beyond hypothermia. J Cell Physiol. 2022;237(10):3788–802.
41. Bastide A, Peretti D, Knight JR, Grosso S, Spriggs RV, Pichon X, Sbarrato T, Roobol A, Roobol J, Vito D, et al. RTN3 Is a Novel Cold-Induced Protein and Mediates Neuroprotective Effects of RBM3. Curr Biol. 2017;27(5):638–50.
42. Wellmann S, Buhrer C, Moderegger E, Zelmer A, Kirschner R, Koehne P, Fujita J, Seeger K. Oxygen-regulated expression of the RNA-binding proteins RBM3 and CIRP by a HIF-1-independent mechanism. J Cell Sci. 2004;117(Pt 9):1785–94.
43. Yan J, Goerne T, Zelmer A, Guzman R, Kapfhammer JP, Wellmann S, Zhu X. The RNA-Binding Protein RBM3 Promotes Neural Stem Cell (NSC) Proliferation Under Hypoxia. Front Cell Dev Biol. 2019;7:288.
44. Feng J, Pan W, Yang X, Long F, Zhou J, Liao Y, Wang M. RBM3 Increases Cell Survival but Disrupts Tight Junction of Microvascular Endothelial Cells in Acute Lung Injury. J Surg Res. 2021;261:226–35.
45. Zhu X, Buhrer C, Wellmann S. Cold-inducible proteins CIRP and RBM3, a unique couple with activities far beyond the cold. Cell Mol Life Sci. 2016;73(20):3839–59.
46. Kostyunina DS, McLoughlin P. RNA-Seq Analysis Revealed Sex Differences in Pulmonary Microvascular Endothelial Cells in the Response to Hypoxia. FASEB J 2022, 36(S1) :R2951.
47. Li Y, Zhang J, Sun H, Chen Y, Li W, Yu X, Zhao X, Zhang L, Yang J, Xin W, et al. lnc-Rps4l-encoded peptide RPS4XL regulates RPS6 phosphorylation and inhibits the proliferation of PASMCs caused by hypoxia. Mol Ther. 2021;32(8):2799.
48. Koizume S, Takahashi T, Yoshihara M, Nakamura Y, Ruf W, Takenaka K, Miyagi E, Miyagi Y. Cholesterol Starvation and Hypoxia Activate the FVII Gene via the SREBP1-GILZ Pathway in Ovarian Cancer Cells to Produce Procoagulant Microvesicles. Thromb Haemost. 2019;119(7):1058–71.
49. Li Y, Huang H, Zhu Z, Chen S, Liang Y, Shu L. TSC22D3 as an immune-related prognostic biomarker for acute myeloid leukemia. iScience. 2023;26(8):107451.
50. Cui CY, Yin M, Sima J, Childress V, Michel M, Piao Y, Schlessinger D. Involvement of Wnt, Eda and Shh at defined stages of sweat gland development. Development. 2014;141(19):3752–60.
51. Dermitzakis I, Chatzi D, Kyriakoudi SA, Evangelidis N, Vakirlis E, Meditskou S, Theotokis P, Manthou ME. Skin Development and Disease: A Molecular Perspective. Curr Issues Mol Biol. 2024;46(8):8239–67.
52. Kowalczyk-Quintas C, Schneider P. Ectodysplasin A (EDA) - EDA receptor signalling and its pharmacological modulation. Cytokine Growth Factor Rev. 2014;25(2):195–203.
53. Cui CY, Schlessinger D. EDA signaling and skin appendage development. Cell Cycle. 2006;5(21):2477–83.
54. Krull F, Bleyer M, Schafer J, Brenig B. A missense mutation in the highly conserved TNF-like domain of Ectodysplasin A is the candidate causative variant for X-linked hypohidrotic ectodermal dysplasia in Limousin cattle: Clinical, histological, and molecular analyses. PLoS ONE. 2024;19(1):e0291411.
55. Jansch N, Meyners C, Muth M, Kopranovic A, Witt O, Oehme I, Meyer-Almes FJ. The enzyme activity of histone deacetylase 8 is modulated by a redox-switch. Redox Biol. 2019;20:60–7.
56. Emmons MF, Faiao-Flores F, Sharma R, Thapa R, Messina JL, Becker JC, Schadendorf D, Seto E, Sondak VK, Koomen JM, et al. HDAC8 Regulates a Stress Response Pathway in Melanoma to Mediate Escape from BRAF Inhibitor Therapy. Cancer Res. 2019;79(11):2947–61.
57. Kudithipudi S, Schuhmacher MK, Kebede AF, Jeltsch A. The SUV39H1 Protein Lysine Methyltransferase Methylates Chromatin Proteins Involved in Heterochromatin Formation and VDJ Recombination. ACS Chem Biol. 2017;12(4):958–68.
58. Feliciello I, Sermek A, Pezer Z, Matulic M, Ugarkovic D. Heat Stress Affects H3K9me3 Level at Human Alpha Satellite DNA Repeats. Genes (Basel) 2020, 11(6): 663.

59. Tichý L, Novák K, Kyselová J, Přibáňová M, Calta J, Vostrý L. Diversity of the bovine genes IRAK1 and IRAK4 in the Toll-like receptor signaling pathway. Czech J Anim Sci. 2023;68(8):323–32.
60. Taraktsoglou M, Szalabska U, Magee DA, Browne JA, Sweeney T, Gormley E, MacHugh DE. Transcriptional profiling of immune genes in bovine monocyte-derived macrophages exposed to bacterial antigens. Vet Immunol Immunopathol. 2011;140(1–2):130–9.
61. Pereira M, Ramalho T, Andrade WA, Durso DF, Souza MC, Fitzgerald KA, Golenbock DT, Silverman N, Gazzinelli RT. The IRAK1/IRF5 axis initiates IL-12 response by dendritic cells and control of Toxoplasma gondii infection. Cell Rep. 2024;43(2):113795.
62. Gao Y, Xu X, Feng J, Ma Y, Zheng D, Meng Y, Shan F. Effects of interleukin-1 receptor-associated kinase 1 RNA interference in dendritic cells on inflammatory cytokine release and T-cell proliferation. Mol Med Rep. 2016;14(6):5685–92.
63. Wang YH, Reverter A, Kemp D, McWilliam SM, Ingham A, Davis CK, Moore RJ, Lehnert SA. Gene expression profiling of Hereford Shorthorn cattle Following challenge with Boophilus microplus tick larvae. Aust J Exp Agric. 2007;47:1397–407.
64. Marufu MCMJ, Nel CL, Man'ombe E, Dube B, Mapholi N, Jonsson NN, Muchadeyi F. K. D: Differential Expression of Tick-Resistance Related Genes Following Artificial Infestation with R. Microplus and R. Decoloratus Ticks. In: *9th Veterinary, Paraveterinary & SASVEPM Congress: 2017*; 2017.
65. Marima JK, Nel CL, Marufu MC, Jonsson NN, Dube B, Dzama K. A genetic and immunological comparison of tick-resistance in beef cattle following artificial infestation with Rhipicephalus ticks. Exp Appl Acarol. 2020;80(4):569–90.
66. Valnegri P, Montrasio C, Brambilla D, Ko J, Passafaro M, Sala C. The X-linked intellectual disability protein IL1RAPL1 regulates excitatory synapse formation by binding PTPdelta and RhoGAP2. Hum Mol Genet. 2011;20(24):4797–809.
67. Guo Q, Jin Y, Chen X, Ye X, Shen X, Lin M, Zeng C, Zhou T, Zhang J. NF-kappaB in biology and targeted therapy: new insights and translational implications. Signal Transduct Target Ther. 2024;9(1):53.
68. Han Y, Mora J, Huard A, da Silva P, Wiechmann S, Putyrski M, Schuster C, Elwakeel E, Lang G, Scholz A, et al. IL-38 Ameliorates Skin Inflammation and Limits IL-17 Production from gammadelta T Cells. Cell Rep. 2019;27(3):835–46. e835.

Publisher's Note