Iranian Society of Animal Science

Journal of Livestock Science and Technologies

Shahid Bahonar University of Kerman

ISSN: 2322-3553 (Print) ISSN: 2322-374X (Online)

Paper type: Original Research

Investigation of selection signatures provides key insights into genetic differences between Holstein and Sarabi dairy cattle breeds

Parisa Biabani[1], Hassan Mehrabani Yeganeh[1], Hossein Moradi Shahrbabak[1], Mahdi Mokhber[2*]*

[1] Department of Animal Science, Faculty of Agricultural Science and Engineering, University College of Agriculture and Natural Resources (UTCAN), University of Tehran, Karaj, Iran
[2] Department of Animal Science, Faculty of Agriculture, Urmia University, Urmia, Iran

[*]Corresponding author,
E-mail address:
hmoradis@ut.ac.ir
m.mokhber@ut.ac.ir

Received: 19 Apr 2025,
Received in revised form: 25 Jun 2025,
Accepted: 12 Jul 2025,
Published online: 13 Jul 2025,

ORCID
Parisa Biabani
0000-0002-1532-6201
Hassan Mehrabani Yeganeh
0000-0001-6820-6759
Hossein Moradi Shahrbabak
0000-0002-6680-7662
Mahdi Mokhber
0000-0002-7184-8426

Abstract Selection increases the frequency of beneficial alleles in subpopulations, leaving genomic signatures associated with genes and QTLs controlling economic traits. Genomic data from 60 Holstein and 72 Sarabi cattle were analyzed to identify these selection signatures. Quality control and data filtering were performed using PLINK_1.9 software. The genetic group was identified using three complementary methods: principal component analysis (PCA) with PLINK 1.9, discriminant analysis of principal components (DAPC) implemented in the adegenet package in R, and Admixture analysis conducted with Admixture version 1.23. Subsequently, F_{ST}, XP-EHH, and Rsb statistics were used to identify selection signatures. The chromosomal positions of the selected regions were aligned with the bovine genome data (ARS-UCD1.2 Bos Taurus) from the Ensembl Biomart database. Genetic analyses revealed the presence of two distinct genetic groups with different origins. In this study, 16 genomic regions were identified using the F_{ST} method, and 18 regions were determined using XP-EHH and Rsb methods, covering approximately 17.5 and 24 Mbp of the bovine genome, respectively. Based on gene ontology analyses, these regions contained coding genes related to key biological processes such as immune response, muscle growth, reproduction, and milk production. Several genes, including *MYO1A*, *STAT6*, and *PRKAA1*, were associated with traits such as carcass quality, fertility, and metabolic processes. The analysis of identified QTLs confirmed the presence of economically important traits, such as growth rate, disease resistance, meat quality, and milk composition. Regions on Bos taurus autosome (BTA) 6 and 10 were identified as key areas for immune-related genes, while milk production traits were observed on regions of BTA 5, 7, and 20. Overall, these findings provide valuable insights into the genetic basis of important economic traits in cattle and can contribute to future breeding programs aimed at improving productivity and disease resistance.

Keywords: dairy cattle, selection sweep, genomic array, population differentiation index

Introduction

Due to the importance of cattle species in meat and milk production and agricultural activities, research into breeding and improving these animals has expanded. The major dairy cattle breeds raised worldwide for milk production include Holstein, Brown Swiss, Ayrshire, Jersey, and Guernsey. Holstein is the dominant breed for milk production among these breeds, as its milk yield has increased due to selective breeding efforts. In recent years, in addition to studies on world-renowned exotic and purebred livestock, numerous studies have also focused on native cattle breeds (Gautier et al., 2010; Kukučková et al., 2017; Maiorano et al., 2018). The increasing knowledge about the structure and genetic diversity of native breeds is

essential for their effective use in sustainable animal husbandry and agriculture, the harsh and underdeveloped breeding conditions, and the genetic preservation of livestock (Groeneveld et al., 2010). Considering that native breeds exhibit high adaptability to their environment and have significantly longer lifespans, the gene pool of unselected native breeds is a valuable genetic resource (Medugorac et al., 2009).

Selection signatures are genomic patterns that arise due to selective forces acting on specific genomic regions (Simianer et al., 2014). They include localized reductions in genetic diversity, deviations from allele frequency distributions, increased linkage disequilibrium, extended haplotype structures, and genetic differentiation between populations (Qanbari and Simianer, 2014). Various tools and methods have been successfully developed and applied to identify selection signatures across different populations at the genomic level (Li et al., 2014).

Numerous studies have been conducted to identify genomic regions under selection in cattle (Stella et al., 2010; Maiorano et al., 2018; Biabani et al., 2022; Salehi et al., 2023), goats and sheep (Moradi et al., 2012; Bertolini et al., 2018; Álvarez et al., 2020; Azizi et al., 2024), and pigs (Zhang et al., 2020). Other studies have focused on horses (Grilz-Seger et al., 2019; Nolte et al., 2019), buffalo (Mokhber et al., 2015; Sun et al., 2020), camels (Bahbahani et al., 2019), and poultry (Almeida et al., 2019).

In this regard, this study aims to utilize genomic data to identify regions of the genome that have been subjected to natural or artificial selection forces over time. These genomic regions may serve as distinguishing markers between the Holstein and Sarabi breeds.

Materials and methods

Breed Selection History and Justification for Sarabi

The Holstein breed (Holstein-Friesian) results from over a century of intensive artificial selection aimed at optimizing dairy production traits. Organized breeding programs since the 1950s, focusing on milk yield traits (e.g., volume and fat percentage), have established it as the dominant dairy breed in global industrial systems (García-Ruiz et al., 2016). In contrast, the Sarabi is an indigenous Iranian cattle breed reared in West Azerbaijan province. This breed has evolved through long-term adaptation to challenging environmental conditions (e.g., cold climate, limited forage resources, and endemic diseases) and natural selection for survival, disease resistance, and production efficiency in traditional farming systems (Dadpasand et al., 2013). The Sarabi breed was selected for this study based on the following rationale: I. Unique Genetic Diversity: As a valuable genetic reservoir with adaptive traits (e.g., resistance to infectious diseases like foot-and-mouth disease and cold stress tolerance) (Mirzaee et al., 2019); II. Divergent Selection Patterns: The absence of industrial breeding interventions allows identification of selection signatures linked to environmental adaptation (contrasting with anthropogenic selection in Holsteins), and III. Conservation Imperative: Genomic data from this breed is critical for designing conservation and sustainable utilization strategies (FAO, 2015).

Genomic data collection

Genomic data from 132 samples, comprising 60 Holstein cows and 72 purebred Sarabi cows, were used in the present study. The Sarabi breed (72 heads purebred Sarabi cows) samples were collected from East-Azerbaijan (north-western part of Iran (37.02° – 38.78° N, 44.81° - 49.52° E), and were sequenced through collaboration with Saina Gostar Alborz Company using Geneseek's 40K arrays. whereas the Holstein genomic data (60 heads related France Holstein population) were obtained from the WIDDE database (http://widde.toulouse.inra.fr/widde/widde/main.do;jsessionid=1DDE4ECBC7809DD4448A9768E60AA7A6?module=cattle. (Sempéré et al., 2015).

Quality control and data filtration

The Plink 1.9 software was used to perform quality control and filtration of individuals and single-nucleotide polymorphism (SNP) markers (Purcell et al., 2007). The quality control criteria were as follows:

First, individuals and SNPs with less than a 95% call rate were excluded from this study. Then, the data were filtered based on a minor allele frequency threshold of less than 1% and markers out of the Hardy-Weinberg equilibrium (P-value $>10e^{-6}$). Considering that the samples used in the current study were sequenced using different platforms, the filtered data were merged, and once common genomic data were obtained, the data were re-filtered to exclude SNP markers with high missing data.

Principal component analysis (PCA), discriminant analysis of principal components (DAPC), and population structure analysis

To elucidate genetic structure and delineate clusters within the studied populations, three complementary multivariate analyses were implemented. Principal Component Analysis (PCA) was performed using PLINK 1.9 (Purcell et al., 2007), applying standard quality control and linkage disequilibrium pruning to identify major axes of variation. Discriminant Analysis of Principal Components (DAPC) was executed in R using the adegenet package (Jombart, 2008; Jombart & Ahmed, 2011), with cross-validation to optimize retained principal components for maximal between-group differentiation. Model-based admixture analysis was conducted using ADMIXTURE (Alexander et al., 2009), employing maximum-likelihood estimation of individual ancestry proportions; the optimal number of ancestral clusters (K) was determined by minimizing cross-validation error. Results from all analyses (PCA

scatterplots, DAPC membership probabilities, ADMIXTURE ancestry proportions) were integrated and visualized using the ggplot2 package (Wickham, 2016) within the R statistical environment (R Core Team, 2023).

Detection of selection signatures

The unbiased fixation index (F_{ST}) estimator (Weir and Cockerham, 1984), cross-population extended haplotype homozygosity (XP-EHH) (Sabeti et al., 2007), and extended haplotype homozygosity between populations (Rsb) statistics (Tang et al., 2007) were calculated to examine the pattern of positive selection between distinct genetic groups (Holstein and Sarabi). All statistics were determined solely on Bos taurus autosomal chromosomes (BTA). with appropriate software packages in the R environment. To better identify selection signals at the genome level, instead of directly considering the numerical value of each SNP, a sliding window approach with a window length of 300 Kbp was employed to average the values of adjacent SNPs (Qanbari et al., 2012). Finally, 0.1% of the studied markers were identified as selection signals. The search for selection signals using XP-EHH and Rsb statistics was performed via the ReHH software. XP-EHH and Rsb statistics are computed via high-frequency and high EHH alleles. The EHH and iHS statistics, after imputation and phasing with Beagle software (Browning and Browning, 2007), were calculated using ReHH software (Gautier and Vitalis, 2012). Eventually, Manhattan plots for each method were drawn using the QQman software package in the R environment (R Core). (Ripley, 2001).

Gene ontology

The selected regions' chromosomal positions aligned with the gene positions were listed for the cattle genome (ARS-UCD1.2 Bos Taurus Genome) in the Ensemble Biomart Tool (http://www.ensemble.org/biomart/martview) (Zhang et al., 2011). Next, an extensive literature review was conducted to comment on the functions of the identified genes. Finally, biological pathways and gene networks were determined by DAVID (Sherman et al., 2022).

Results and discussion

Quality control and data screening

Following various screening stages based on raw data quality control information and animal kinship, a total of 26,492 SNP markers from autosomal BTA of 132 animals (including 60 Holstein and 72 Sarabi heads) passed the quality control process and were selected for final analyses.

Population structure analysis and determination of diversity components

PCA is a widely used method for examining population differentiation and clustering in the genomic studies of domestic animals (Kijas et al., 2013; Uzzaman et al., 2014). This approach estimates and visualizes population structure based on the genetic relationships among the individuals. In other words, the results of PCA reflect the genetic affinity among the studied individuals. Ultimately, animals are positioned close to or distant from each other based on their genetic relatedness. The results of PCA indicated that the studied populations fall into two completely distinct groups (Figure 1), with the first two components explaining 15% of the total variance.

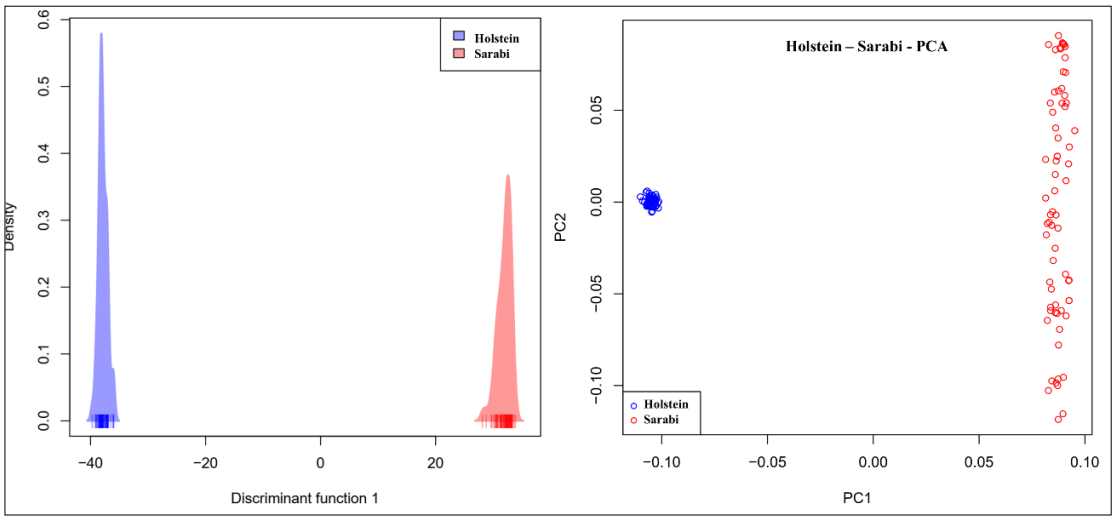

Figure 1. Graphical representation of PCA (right) and DAPC (left) analyses for Holstein and Sarabi cattle breeds based on SNP marker data.

The Admixture analysis results were confirmed the genetic distinction between the two studied groups (Figure 2). Cross-validation was performed under the assumption of 1–3 ancestral populations. The results for the assumption with two genetic groups exhibited the lowest validation error and provided the highest clarity in distinguishing between the populations (Figure 2).

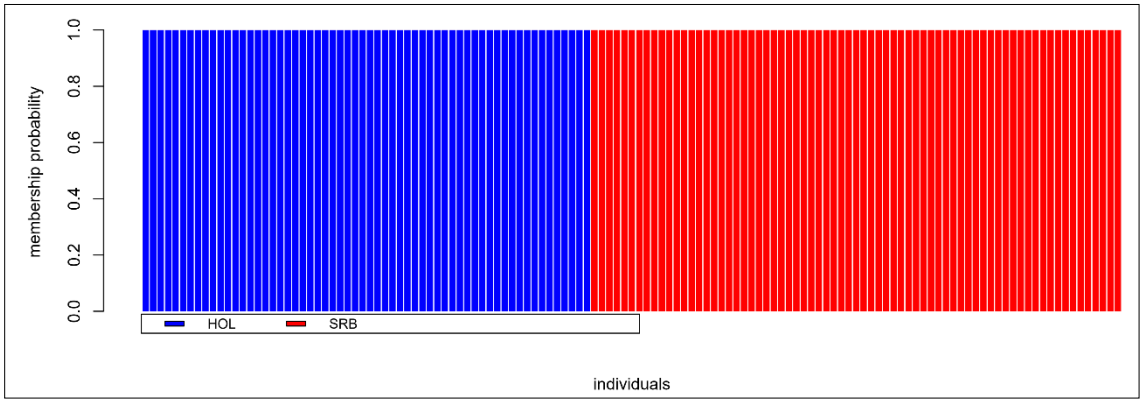

Figure 2. Graph of population admixture analysis results with K=2 for Holstein (blue. and Sarabi (red. cattle breeds using SNP marker data

Detection of selection signatures, Manhattan plots and gene ontology results

The Manhattan plots (Figure 3) depict genome-wide F_{ST}, XP-EHH, and Rsb values for Sarabi vs. Holstein cattle, with a threshold line highlighting selection signatures. Gene ontology analysis of 16 genomic regions (~17.5 Mbp) identified via FST, iHS, and Rsb revealed 9 regions containing 60 protein-coding genes, localized to BTA 2, 3, 6, 9, 10 (two regions), 16, 20, and 26 (Table 1). Notably, the BTA 10 region (22–23 Mbp) aligns with a reported selection signature in buffaloes (Mokhber et al., 2018).

Of the 18 regions (~24 Mbp) detected as selection signatures, 160 protein-coding genes were identified, distributed across BTA 5 (two regions), 6, 7 (two regions), 12 (two regions), 15, 16 (two regions), 18, 20 (six regions), and 25 (Table 1). Extended regions on BTA 7, 18, and 20 (spanning 3–6 Mbp) were attributed to broad marker peaks, cumulatively covering ~24 Mbp.

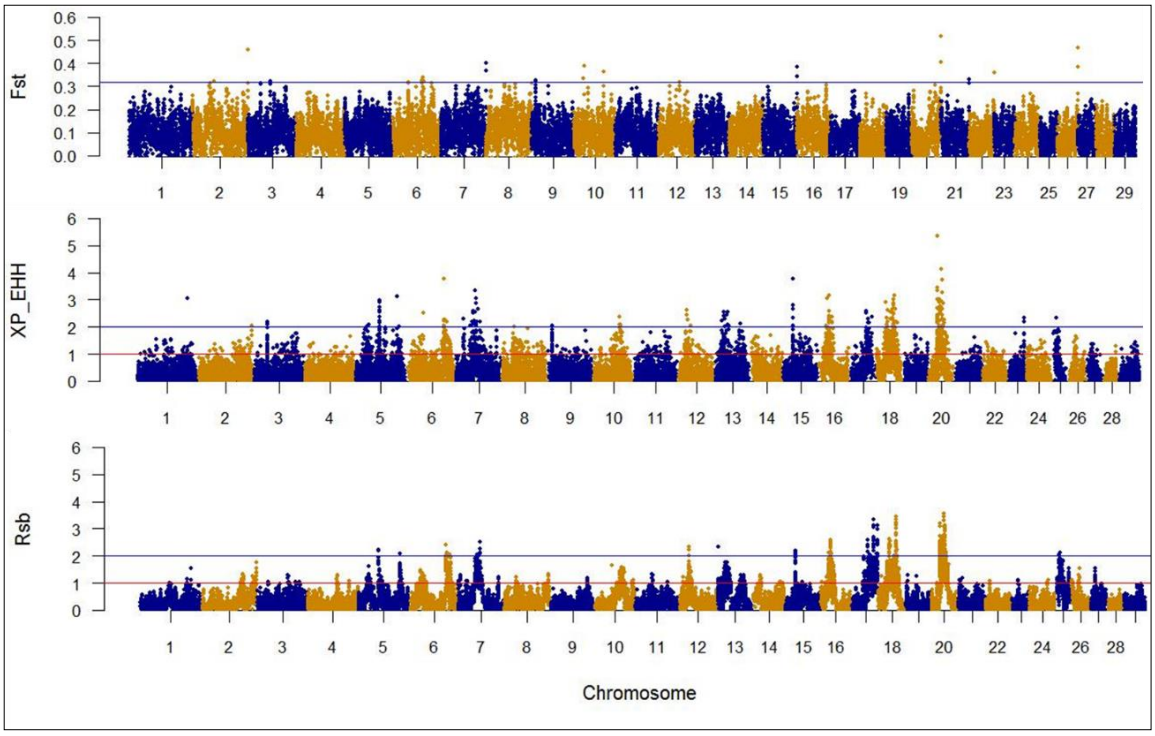

Figure 3. Manhattan plot of F_{ST}, XP-EHH, and Rsb values comparing Sarabi and Holstein cattle breeds. The blue line represents the threshold for identifying genomic regions under selection

Several genomic regions identified in this study, including BTA5 (55.4–57 Mbp), BTA6 (71.5–72.5 Mbp), BTA7 (44.5–46 Mbp), and BTA2 (32.5–35 Mbp), align with selection signatures reported in prior research (Yurchenko et al., 2018; Moravčikova et al., 2019). Notably, the BTA20 region has been recurrently highlighted across studies (Bovine HapMap Consortium et al., 2009; Qanbari et al., 2010; Stella et al., 2010), underscoring its significance in cattle genetic improvement.

The BTA5 region (55.4–57 Mbp) emerged as a critical locus enriched with functional genes such as *PRIM1*, *PTGES3*, *BAZ2A*, *GLS2*, *APOF*, *MIP*, and *TIMELESS*, previously linked to carcass and meat quality traits (Yurchenko et al., 2018). These findings corroborate earlier reports associating this locus with economically vital phenotypes in cattle (Peters et al., 2012; Saatchi et al., 2014).

Key genes associated with meat quality traits include *MYO1A* (muscle development), *NACA* (myoblast regulation, lipid metabolism), *R3HDM2*, *TAC3*, and *STAT6* (carcass traits, immune responses). These genes, previously linked to carcass weight, backfat thickness, and skeletal muscle formation in cattle (Cui et al., 2012; Leal-Gutierrez et al., 2018; Wang et al., 2022), underscore their role in economically vital phenotypes. *STAT6*, identified as a selection signature on BTA 5, influences traits such as shear force, daily weight gain, and immune pathway activation (Rincon et al., 2009; Signer-Hasler et al., 2017).

Table 1. List of identified genes related to selected regions between Holstein and Iranian Sarabi cattle's using F_{ST}, XP-EHH and Rsb methods

Chr	Start (Kbp)	End (Kbp)	Methods	Detected Genes
1	52.17	52.83	F_{ST}	*ZEB2 - GTDC1 -ENSBTAG00000052105*
3	55.35	56.03	F_{ST}	*U6 -ENSBTAG00000051499 -ENSBTAG00000049274*
5	56.08	57.16	XP_EHH و Rsb	*R3HDM2 - STAC3 - NDUFA4L2 - SNORA62 - STAT6 - NEMP1 - MYO1A - TAC3 - ZBTB39 - RDH16 - SDR9C7 - PRIM1 – NACA - PTGES3 - ATP5F1B - bta-mir-677 - bta-mir-677 - BAZ2A - GLS2 – MIP – TIMELESS – APON – APOF -STAT2 - PAN2 - CNPY2 - bta-mir-12054 – CS - ANKRD52 - RNF41 - SMARCC2*
5	96.77	97.70	XP_EHH و Rsb	*GSG1 - HEBP1 - GPRC5D - GPR19 - DUSP16 - 5S_rRNA - MANSC1 - LRP6*
6	71.60	75.20	F_{ST}	*CRACD – PAICS - SRP72 - ARL9 – THEGL - RESTB - POLR2B - Y_RNA - ENSBTAG00000052745 - ENSBTAG00000042899*
6	88.48	90.64	XP_EHH و Rsb	*ALB – AFP – AFM - 7SK - CXCL8 - CXCL5 - CXCL2 - CXCL3 - GRO1 - MTHFD2L – EPGN – EREG –AREG -PARM1 - THAP6 - USO1*
7	44.40	44.78	XP_EHH و Rsb	*SOWAHA – UQCRQ - LEAP2 - HSPA4 - 5S_rRNA*
7	45.77	46.46	XP_EHH و Rsb	*TCF7 - bta-mir-2285di - UBE2B - JADE2 - SEC24A – CAMLG - DDX46 - C7H5orf24 - 5S_rRNA - TXNDC15 - PCBD2 - CATSPER3*
9	10.63	11.70	XP_EHH و Rsb	*OGFRL1 - RIMS1 - ENSBTAG00000048046*
10	22.75	23.27	F_{ST}	*ENSBTAG00000052267- ENSBTAG00000054833 - ENSBTAG00000050913 - ENSBTAG00000047690 - ENSBTAG00000054417 - ENSBTAG00000050391 - ENSBTAG00000050781 - ENSBTAG00000054091 - ENSBTAG00000046819 - ENSBTAG00000038544 - ENSBTAG00000045863 - ENSBTAG00000054130 - ENSBTAG00000049260 - ENSBTAG00000052371 - ENSBTAG00000051127 - ENSBTAG00000051187 - ENSBTAG00000018947 - ENSBTAG00000030792*
10	73.94	74.26	F_{ST}	*SYT16 - ENSBTAG00000030792*
12	22.80	23.52	XP_EHH و Rsb	*LHFPL6 - PROSER1 - STOML3 -U6*
12	23.97	24.17	XP_EHH و Rsb	*TRPC4*
15	23.71	24.78	XP_EHH و Rsb	*NCAM1- TTC12 - CLDN25- HTR3B - U6 - HTR3A- ZBTB16- U8*
16	25.26	25.66	XP_EHH و Rsb	*snoRNA*
	26.10	26.36	XP_EHH و Rsb	*MIA3 - BROX - FAM177B - DISP1*
16	72.97	73.71	F_{ST}	*U5 - IRF6 - C16H1ORF74 - LAMB3*
18	40.42	43.55	XP_EHH و Rsb	*CCNE1 - URI1 - ZNF536 - bta-mir-2899 - SNORA70 - ZNF507 - DPY19L3 - PDCD5 - RGS9BP - NUDT19 - TDRD12 - FAAP24 - GPATCH1*
20	21.80	22.60	XP_EHH و Rsb	*ACTBL2 - bta-mir-2285f2 - MIER3 - U1 - ENSBTAG00000042330 - ENSBTAG00000026505 - ENSBTAG00000053164 - ENSBTAG00000051954*
20	32.67	33.73	XP_EHH و Rsb	*OXCT1 - PLCXD3 - C6 - MROH2B - RPL37 - PRKAA1 - TTC33*
20	34.85	34.89	XP_EHH و Rsb	*U2 _ rRNA*
20	35.00	35.95	XP_EHH و Rsb	*DAB2 - C9 -FYB1 _ RICTOR _ LIFR + ENSBTAG00000054984*
20	36.45	36.64	XP_EHH و Rsb	*GDNF + 2 snoRNA*
20	36.97	38.28	XP_EHH و Rsb	*NUP155 - CPLANE1 - RANBP3L - NADK2 - LMBRD2 - UGT3A2 - CAPSL*
20	71.39	71.95	F_{ST}	*ENSBTAG00000006971 - ENSBTAG00000054687 - ENSBTAG00000052247 - ENSBTAG00000004629 - ENSBTAG00000045056 - ENSBTAG00000047780 - ENSBTAG00000055240 - ENSBTAG00000048135 - ENSBTAG00000047700 - ENSBTAG00000047632*
25	10.68	11.65	XP_EHH و Rsb	*SNX29 - SHISA9*
26	51.24	51.98	F_{ST}	*JAKMIP3 - DPYSL4 -LRRC27 - PWWP2B - ENSBTAG00000050527 - ENSBTAG00000050923 - ENSBTAG00000052910*

The genomic region spanning 88.48–90.64 Mbp on BTA 6 harbors immune-related genes (*CXCL5*, *CXCL8*, *EPGN*, *EREG*), which regulate cytokine activity, inflammation, and mammary gland development. *ALB* in this locus is critical for fatty acid metabolism, apoptosis regulation, and reproductive processes. Additionally, *APOF*, *APON*, and *ATP5F1B* were linked to lipid biosynthesis, cholesterol metabolism, and lung development (Gutiérrez-Gil et al., 2015; Bertolini et al., 2018). These findings highlight genomic hotspots governing meat quality, immune function, and metabolic efficiency in cattle

Genomic regions on BTA 7 harbored *TCF7* (T-cell receptor recombination, *Wnt* signaling), *UBE2B*

(apoptosis regulation), *JADE2* (neurogenesis), and *TXNDC15* (smoothened signaling). On BTA 9, the *RIMS1* (synapse formation) and *OGFRL1* (opioid receptor activity) genes were identified. BTA 10 (22.75–23.27 Mbp) and BTA 12 (23.96–24.17 Mbp) contained immune-related loci (e.g., *ENSBTAG00000052267*, *TRPC4*) linked to pathogen resistance and oligodendrocyte differentiation.

BTA 15 featured *TTC12* (sperm axoneme assembly), *HTR3B* (serotonin signaling), and *ZBTB16* (skeletal/immune cell differentiation). On BTA 16, *IRF6* (epithelial/organ development) and *LAMBP3* (tissue morphogenesis) were highlighted. These regions underscore critical roles in immunity, neural development, and tissue differentiation.

The BTA 18 region harbors *CCNE1* (*Wnt* signaling, cell division), *URI1* (growth regulation, viral response), *PDCD5* (apoptosis regulation), and *RGS9BP* (visual perception). *ZNF536*/*ZNF507* are linked to structural traits, longevity, and calving difficulty in Holsteins (Cole et al., 2011), with QTLs associated with fertility, milk production, and calving ease. On BTA 20, *PRKAA1* (metabolic regulation, lipid biosynthesis), *GDNF* (neural development), *C6*/*FYB1* (immune response), and *LIFR* (inflammatory signaling) were identified. *PRKAA1* and *PLCXD3* further influence lipid metabolism, while *RICTOR* and *CPLANE1* regulate muscle/kidney development. Notably, the *GHR* gene in this region (Yurchenko et al., 2018) underscores its role in growth and milk production.

Genes such as *JADE2*, *RIMS1*, *TRPC4*, and *PRKAA1* are enriched in neural differentiation pathways. Immune-related loci (e.g., *CXCL8*, *CXCL5*, *TCF7*) cluster on BTA 6 and 10, highlighting their significance in pathogen resistance. *PRKAA1*, *CCNE1*, and *TCF7* modulate milk production via *Wnt* signaling, while *APOF*, *APON*, and *PLCXD3* regulate lipid metabolism. These pathways align with QTLs for milk yield/composition on BTA 5, 7, 18, and 20, emphasizing Wnt signaling as a central driver of lactation traits.

QTLs related to divergent selected regions

This study identified 18 genomic regions (spanning ~24 Mbp) via XP-EHH/rSB methods and 16 regions (~17.5 Mbp) via FST as selection signatures harboring quantitative trait loci (QTLs) associated with critical functional traits in cattle. These QTLs were linked to economic and health-related traits, including herd longevity (BTA 6, 7, 12, 16, 21), resistance to bovine tuberculosis (BTA 15), respiratory disease susceptibility, and immune cell counts. Growth and meat quality traits encompassed birth weight, body weight gain (BTA 2, 26), muscle composition (e.g., carnosine, taurine), fat deposition (subcutaneous, intramuscular), and meat tenderness (BTA 16). Feed efficiency traits included residual feed intake (BTA 2, 21), while reproductive traits involved fertility rates (BTA 2, 3, 6, 15, 18), insemination success (BTA 6, 18, 22, 26), and age at puberty.

Milk production QTLs were associated with fat/protein yield (BTA 5, 7, 12, 20), κ-casein percentage (BTA 6, 7, 15, 18), and fatty acid profiles (e.g., palmitic, linoleic acids). Physiological and structural traits included dystocia/stillbirth (BTA 16) and body conformation (udder depth, hip height; BTA 12, 25). These findings underscore the genomic architecture underlying economically vital traits, providing a foundation for targeted breeding strategies to enhance productivity, disease resilience, and structural soundness in cattle populations (Tables S1, S2).

The findings of this study demonstrated a strong association between the selected genomic regions and important traits, such as milk production and composition, fertility, growth, meat quality, and susceptibility to respiratory diseases. These results highlight the significance of these genomic regions in improving economically essential traits in dairy cattle, particularly in enhancing milk production and disease resistance.

Conclusion

This study employed F_{ST}, XP-EHH, and Rsb statistical methods to identify genomic differences between Holstein and Sarabi cattle in Iran. A total of 16 genomic regions were detected using FST, while 18 regions (spanning ~17.5 and 24 Mbp, respectively) were identified via XP-EHH and Rsb. Gene ontology analysis revealed these regions harbor genes linked to critical biological functions, including immune response, muscle growth, reproduction, and milk production. Key genes such as *MYO1A*, *STAT6*, and *PRKAA1* were associated with traits like carcass quality, fertility, and metabolic processes. Immunity-related genes were concentrated on chromosomes 6 and 10, while milk production traits correlated with regions on chromosomes 5, 7, and 20. These findings offer valuable insights for optimizing cattle breeding programs to enhance productivity and disease resistance.

Acknowledgment

The authors appreciate the anonymous reviewers for their valuable and constructive comments.

Conflicts of Interest

The authors declare that they have no competing interests.

References

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19, 1655-1664.

Almeida, O.A.C., Moreira, G.C.M., Rezende, F.M., Boschiero, C., de Oliveira Peixoto, J., Ibelli, A.M.G., Ledur, M.C., de Novais, F.J., Coutinho, L.L., 2019. Identification of selection signatures involved in

performance traits in a paternal broiler line. *BMC Genomics* 20, 1-20.

Álvarez, I., Fernández, I., Traoré, A., Pérez-Pardal, L., Menéndez-Arias, N.A., Goyache, F., 2020. Genomic scan of selective sweeps in Djallonké (West African Dwarf. sheep shed light on adaptation to harsh environments. *Scientific Reports* 10(1), 2824: 1-13.

Azizi, P., Moradi Shahrbabak, M., Moradi Shahrbabak, H., Mokhber, M., 2024. Identification of genomic regions related to litter size with divergent selection between Iranian indigenous sheep breeds and Romanov high reproductive sheep breed. *Iranian Journal of Animal Science*. doi: 10.22059/ijas.2024.369473.653982.

Bahbahani, H., Musa, H.H., Wragg, D., Shuiep, E.S., Almathen, F., Hanotte, O., 2019. Genome diversity and signatures of selection for production and performance traits in dromedary camels. *Frontiers in Genetics* 10, 893 (1-14).

Bertolini, F., Servin, B., Talenti, A., Rochat, E., Kim, E.S., Oget, C., Palhière, I., Crisà, A., Catillo, G., Steri, R., 2018. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genetics Selection Evolution* 50, 1-24.

Biabani, P., Yeganeh, H.M., Shahrbabak, H.M., Mokhber, M., 2022. Detection of genetic differences between Holstein and Iranian north-west indigenous hybrid cattles using genomic data. *Research on Animal Production* 13(37): 175-186.

Browning, S.R., Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics* 81, 1084-1097.

Chen, N., Cai, Y., Chen, Q., Li, R., Wang, K., Huang, Y., Hu, S., Huang, S., Zhang, H., Zheng, Z., 2018. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nature Cmmunications* 9, 2337: 1-13.

Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., Lawlor, T.J., Crooker, B.A., Van Tassell, C.P., Yang, J., Wang, S., Matukumalli, L.K., 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. *BMC Genomics* 12, 1-17.

Consortium, B.H., Gibbs, R.A., Taylor, J.F., Van Tassell, C.P., Barendse, W., Eversole, K.A., Gill, C.A., Green, R.D., Hamernik, D.L., Kappes, S.M., 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324, 528-32.

Cui, H.-X., Liu, R.-R., Zhao, G.-P., Zheng, M.-Q., Chen, J.-L., Wen, J., 2012. Identification of differentially expressed genes and pathways for intramuscular fat deposition in pectoralis major tissues of fast-and slow-growing chickens. *BMC Genomics* 13, 1-12.

Dadpasand, M., Zamani, P., Mirhoseini, S.Z., 2013. Genetic diversity among Sarabi, Holstein, and Simmental cattle breeds using microsatellite markers. *Iranian Journal of Applied Animal Science* 3(3), 559-564 .

FAO. 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture. *Food and Agriculture Organization of the United Nations*. http://www.fao.org/3/a-i4787e/index.html

García-Ruiz, A., Cole, J.B., VanRaden, P.M., Wiggans, G.R., Ruiz-López, F.J., Van Tassell, C.P., 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proceedings of the National Academy of Sciences* 113(28): E3995-E4004. https://doi.org/10.1073/pnas.1519061113

Gautier, M., Laloë D., Moazami-Goudarzi, K., 2010. Insights into the genetic history of French cattle from dense SNP data on 47 worldwide breeds. *PloS One* 5, e13038.

Gautier, M., Vitalis, R., 2012. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* 28, 1176-7.

Grilz-Seger, G., Neuditschko, M., Ricard, A., Velie, B., Lindgren, G., Mesarič, M., Cotman, M., Horna, M., Dobretsberger, M., Brem, G., 2019. Genome-wide homozygosity patterns and evidence for selection in a set of European and near eastern horse breeds. *Genes* 10, 491, 1-23.

Groeneveld, L., Lenstra, J., Eding, H., Toro, M., Scherf, B., Pilling, D., Negrini, R., Finlay, E., Jianlin, H., Groeneveld, E., 2010. Genetic diversity in farm animals–a review. *Animal Genetics* 41, 6-31.

Gutiérrez-Gil, B., Arranz, J.J., Wiener, P., 2015. An interpretive review of selective sweep studies in Bos taurus cattle populations: identification of unique and shared selection signals across breeds. *Frontiers in Genetics* 6(167), 1-20.

Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521.

Jombart, T., 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11), 1403-1405.

Kijas, J.W., Ortiz, J.S., McCulloch, R., James, A., Brice, B., Swain, B., Tosser-Klopp, G., Consortium, I.G.G., 2013. Genetic diversity and investigation of polledness in divergent goat populations using 52 088 SNPs. *Animal Genetics* 44, 325-35.

Kukučková, V., Moravčíková, N., Ferenčaković, M., Simčič, M., Mészáros, G., Sölkner, J., Trakovická, A., Kadlečík, O., Curik, I., Kasarda, R., 2017. Genomic

characterization of Pinzgau cattle: genetic conservation and breeding perspectives. *Conservation Genetics* 18, 893-910.

Leal-Gutierrez, J.D., Rezende, F.M., Elzo, M.A., Johnson, D., Penagaricano, F., Mateescu, R.G., 2018. Structural equation modeling and whole-genome scans uncover chromosome regions and enriched pathways for carcass and meat quality in beef. *Frontiers in Genetics* 9, 532.

Li, C., Sun, D., Zhang, S., Wang, S., Wu, X., Zhang, Q., Liu, L., Li, Y., Qiao L., 2014. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. *PloS One* 9, e96186.

Maiorano, A.M., Lourenco, D.L., Tsuruta, S., Ospina, A.M.T., Stafuzza, N.B., Masuda, Y., Filho, A.E.V., Cyrillo, J.N.d.S.G., Curi, R.A., Silva, J.A.I.d.V., 2018. Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PloS One* 13, e0200694.

Medugorac, I., Medugorac, A., Russ I., Veit-Kensch, C.E., Taberlet, P., Luntz, B., Mix, H.M., Förster, M., 2009. Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds with high effective population size. *Molecular Ecology* 18, 3394-410.

Mirzaee, M., Ghavi Hossein-Zadeh, N., Shadparvar, A.A., Shahrbabak, H.M., 2019. Genetic diversity in Sarabi cattle: An ancestral Iranian breed. *Journal of Livestock Science and Technologies* 7(1), 59-66. https://doi.org/10.22103/jlst.2019.12322.1231

Mokhber, M., Moradi Shahrbabak. M., Sadeghi, M., Moradi Shahrbabak, H., Williams, J., 2015. Genome-wide survey of signature of positive selection in Khuzestani and Mazandrani buffalo breeds. *Iranian Journal of Animal Science* 46, 119-31.

Mokhber, M., Moradi-Shahrbabak, M., Sadeghi, M., Moradi-Shahrbabak, H., Stella, A., Nicolzzi, E., Rahmaninia, J., Williams, J.L., 2018. A genome-wide scan for signatures of selection in Azeri and Khuzestani buffalo breeds. *BMC Genomics* 19, 1-9.

Moradi, M.H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K.G., McEwan, J.C., 2012. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics* 13, 1-15.

Moravčikova, N., Trakovicka, A., Kadlečik, O., Kasarda, R., 2019. Genomic signatures of selection in cattle through variation of allele frequencies and linkage disequilibrium. *Journal of Central European Agriculture* 20, 576-80.

Nolte, W., Thaller, G., Kuehn, C., 2019. Selection signatures in four German warmblood horse breeds: Tracing breeding history in the modern sport horse. *PloS One* 14, e0215913.

Peters, S., Kizilkaya, K., Garrick, D., Fernando, R., Reecy, J., Weaber, R., Silver, G., Thomas, M., 2012. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *Journal of Animal Science* 90, 3398-3409.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 38, 904-909.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81, 559-575.

Qanbari, S., Pimentel, E., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A.R., Simianer, H., 2010. A genome-wide scan for signatures of recent selection in Holstein cattle. *Animal Genetics* 41, 377-389.

Qanbari, S., Simianer, H., 2014. Mapping signatures of positive selection in the genome of livestock. *Livestock science* 166, 133-143.

Qanbari, S., Strom, T.M., Haberer, G., Weigend, S., Gheyas, A.A., Turner, F., Burt, D.W., Preisinger, R., Gianola, D., Simianer, H., 2012. A high resolution genome-wide scan for significant selective sweeps: an application to pooled sequence data in laying chickens. *PloS One* 7, e49525.

Rincon, G., Farber, E., Farber, C., Nkrumah, J., Medrano, J., 2009. Polymorphisms in the STAT6 gene and their association with carcass traits in feedlot cattle. *Animal Genetics* 40, 878-882.

Ripley, B.D., 2001. The R project in statistical computing. MSOR connections. *The Newsletter of the LTSN Maths, Stats & OR Network* 1, 23-25.

Saatchi, M., Beever, J.E., Decker, J.E., Faulkner, D.B., Freetly, H.C., Hansen, S.L., Yampara-Iquise, H., Johnson, K.A., Kachman, S.D., Kerley M.S., 2014. QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. *BMC Genomics* 15, 1-14.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X.H., Byrne, E.H., McCarroll, S.A., Gaudet, R., et al., 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449(7164), U912-U913.

Salehi, R., Javanmard, A., Mokhber, M., Alijani, S., 2023. Genomic Selection Signatures in Two French and Swedish Holstein Cattle Breeds Provide Evidence for Several Potential Candidate Genes Linked to Economic Traits. *Iranian Journal of Applied Animal Science* 13, 677-684.

Sempéré, G., Moazami-Goudarzi, K., Eggen, A., Laloë, D., Gautier, M., Flori, L., 2015. WIDDE: a Web-Interfaced next generation database for genetic diversity exploration, with a first application in cattle. *BMC Genomics* 16, 1-8.

Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., Chang, W., 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research* 50, 216-221.

Signer-Hasler, H., Burren, A., Neuditschko, M., Frischknecht, M., Garrick, D., Stricker, C., Gredler, B., Bapst, B., Flury, C., 2017. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genetics Selection Evolution* 49, 1-13.

Simianer, H., Ma, Y., Qanbari, S., 2014. Statistical problems in livestock population genomics. In: *Proceedings of the 10th World Congress on Genetics Applied to Livestock Production*, pp. 17-22.

Smad, G., 2020. GeneCards—The human gene database https://www. genecards. org/cgi-bin/carddisp. pl? gene= SMAD6&keywords= smad6. Accessed.

Stella, A., Ajmone-Marsan, P., Lazzari, B., Boettcher, P., 2010. Identification of selection signatures in cattle breeds selected for dairy production. *Genetics* 185, 1451-61.

Sun, T., Huang, G., Wang, Z., Teng, S.-h., Cao, Y., Sun, J., Hanif, Q., Chen, N., Lei, C., Liao Y., 2020. Selection signatures of Fuzhong Buffalo based on whole-genome sequences. *BMC Genomics* 21, 1-10.

Tang, K., Thornton, K.R., Stoneking, M., 2007. A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology* 5, 1587–1602.

Uzzaman, M.R., Edea, Z., Bhuiyan, M.S.A., Walker, J., Bhuiyan, A., Kim K.S., 2014. Genome-wide single nucleotide polymorphism analyses reveal genetic diversity and structure of wild and domestic cattle in Bangladesh. *Asian-Australasian journal of Animal Sciences* 27(10), 1381-1386.

Wang, X., Ran, X., Niu, X., Huang, S., Li, S., Wang, J., 2022. Whole-genome sequence analysis reveals selection signatures for important economic traits in Xiang pigs. *Scientific Reports* 12, 11823, 1-17.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38(6), 1358-1370.

Yurchenko, A.A., Daetwyler, H.D., Yudin, N., Schnabel, R.D., Vander Jagt, C.J., Soloshenko, V., Lhasaranov, B., Popov, R., Taylor, J.F., Larkin, D.M., 2018. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Scientific Reports* 8 (Article No.12984), 1-16.

Zhang, J., Haider, S., Baran, J., Cros, A., Guberman, J.M., Hsu, J., Liang, Y., Yao, L., Kasprzyk, A., 2011. BioMart: a data federation framework for large collaborative projects. *Database* 2011, bar038.

Zhang, W., Yang, M., Zhou, M., Wang, Y., Wu, X., Zhang, X., Ding, Y., Zhao, G., Yin, Z., Wang, C., 2020. Identification of signatures of selection by whole-genome resequencing of a Chinese native pig. *Frontiers in Genetics* 11, 566255: 1-11.