Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Elucidation of Genome Wide Selection Signatures for Milk Production and Adaptive Traits in N'dama African Cattle

Kazeem Ajasa

alqasiminternational@yahoo.com

Federal University Gashua

Gargi Ramola

National Dairy Research Institute

Uddhav Paneru

Nepal Agricultural Research Council

Moshood Abiola Popoola

Federal College of Animal Health and Production Technology Ibadan

Anupama Mukherjee

National Dairy Research Institute

Sabyasachi Mukherjee

National Dairy Research Institute

Research Article

Keywords: N'Dama, Genome, ROH, Tajima's D, Candidate gens

Posted Date: August 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10412150/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

ELUCIDATION OF GENOME WIDE SELECTION SIGNATURES FOR MILK PRODUCTION AND ADAPTIVE TRAITS IN N'DAMA AFRICAN CATTLE

Badmus, Kazeem Ajasa*[1,2], Gargi Ramola[2], Uddhav Paneru[3], Popoola Moshood Abiola[4], Anupama Mukherjee[2] and Sabyasachi Mukherjee[2]

[1]Federal University Gashua, Gashua-631101, Yobe State, Nigeria,
[2]ICAR-National Dairy Research Institute, Karna-132001, Haryana, India,
[3]NARC-National Cattle Research Program, Rampur, Chitwan, Nepal,
[4]Federal College of Animal Health and Production Technology, Ibadan, Nigeria

[1]Gashua-631101, Yobe State, Nigeria,
[2]Karna-132001, Haryana, India,
[3]Rampur, Chitwan, Nepal,
[4]Ibadan, Oyo state, Nigeria

E-mail and telephone number of the Corresponding author: alqasiminternational@yahoo.com/ +2349044649931
Corresponding author: Badmus Kazeem Ajasa

Abstract

Background: N'Dama (West African taurine) cattle exemplify African breeds that evolved under distinct ecological and management regimes. The underlying genetic mechanisms behind these potentials are still not well defined. Understanding the genetic basis of milk and adaptive mechanism can guide in developing the effective breeding goals that both conserve the local breeds and improve the commercial cattle. In this study, we evaluated the genome-wide scans with runs of homozygosity (ROH), integrated haplotype scores (iHS) and Tajima's D to unveil selection signatures for milk production and adaptive traits in N'Dama cattle using publicly available high density SNP chips.

Results: ROH islands, iHS and Tajima's D hotspots overlapped milk and adaptation-related genes consistent with trypanotolerance in N'Dama. Manhattan plot and functional enrichment of these regions implicated milk production QTLs and disease resistance in N'Dama. Notably, ROH hotspots contained candidate genes associated with milk protein synthesis (*LEP, NCK1, IL6R, PDGFB, FGF23, CSNK1E, KIT, RPTOR, CSNK1D*), thermal adaptation (*DNAJB2, STRADB, OAZ3, SASH1, SOCS4, STRADA, MAPKAPK3*), adaptive immunity (*FCGR3A, IFIH1, IFI16, PDCD1, CTLA4, FCGR2A, CD1B, TLR6, ANXA1*) while iHS hotpot contained *CDH2, TCN2, ITGA11, ITGA6* for milk production and *SMURF2, TNXB, CD47, THBS1* for adaptation and Tajima's D hot spot revealed *IGF1, STAT5B, LPIN1, GALM* and *NREP, SMAD6, TLR9. TYRO3, SEL1, CAT, TNF, TXN, AP3B1* as candidate genes associated with milk production and adaptive traits respectively. Some genes such as *DCDC2, HECW1, CAPZB, FER, ANAPC4, HECTD2, RFFL, SLIT2* were found to be overlapping between ROH and iHS. Among overlaps between ROH and Tajima'D were found to be *ANXA1, GCG, GCK, SOCS7, ALOX15B* and that between iHS and Tajima's D were *SDCCAG8* and *EPHB2*. No overlap genes were intersected between the three selection signature methods.

Conclusions: The contrasting ROH, iHS and Tajima's D landscapes of N'Dama reflected demographic history in selection and highlighted candidate loci for breeding milk productive and adaptive N'Dama cattle. The overlaps between ROH, iHS and Tajima's D signals provides

robust evidence of recent positive selection, highlighting high-confidence candidate genes likely contributing to adaptation and economically important traits such milk.

Keywords: N'Dama; Genome; ROH; Tajima's D, Candidate gens

BACKGROUND

African indigenous cattle represent a genomic mosaic shaped by historical taurine-indicine admixture, ecological adaptation, and centuries of diverse management practices [1-2,3]. Among these, N'Dama (Bos taurus) cattle are exceptional for their natural trypanotolerance, which enables them to thrive in tsetse-infested West African regions where other breeds succumb to African animal trypanosomiasis [4,5]. This makes them an invaluable genetic resource for improving cattle productivity and resilience in tropical environments. In contrast, Kuri cattle, adapted to semi-aquatic ecosystems, face challenges such as small population sizes and higher inbreeding. Despite their adaptive importance, N'Dama cattle remain underrepresented in genome-wide studies combining runs of homozygosity (ROH), integrated haplotype score (iHS), and Tajima's D, which together provide a comprehensive understanding of selection signatures.

ROH represents long stretches of homozygous SNPs that reveal genome-wide autozygosity and help distinguish between recent and ancient inbreeding [6,7]. The proportion of the genome covered by ROH (FROH) quantifies realized inbreeding, offering insights into population history and genetic diversity [8,9]. ROH islands shared across individuals often mark regions under positive selection, where advantageous alleles have become fixed [10]. In livestock, such regions frequently overlap with QTLs associated with milk production, growth, reproduction, and adaptation [3,13].

In contrast, iHS detects *partial selective sweeps* regions where beneficial alleles are still segregating but surrounded by extended haplotype homozygosity, signalling recent positive selection [10,12]. iHS is particularly useful for identifying genomic regions linked to

adaptation, disease resistance, and production traits [13,14]. Tools such as PLINK, SHAPEIT, and SELSCAN are commonly used for these analyses [15].

Tajima's D, on the other hand, is a site frequency spectrum (SFS)-based neutrality test [16] that compares nucleotide diversity to segregating sites. Negative values suggest recent positive selection or population expansion, while positive values indicate balancing selection or bottlenecks. Unlike ROH and iHS, which depend on haplotype structure and linkage disequilibrium, Tajima's D reflects allele frequency changes across loci, thus detecting older or frequency-based selection signals [17].

Integrating ROH, iHS, and Tajima's D provides a robust, complementary framework for detecting selection signatures across different evolutionary timescales: ROH identifies long-term fixation or autozygosity, iHS captures recent selective sweeps, and Tajima's D highlights ancient or frequency-based selection events [18,19]. Overlapping results from these methods strengthen confidence in true selection signals by reducing false positives due to drift or population structure [20]. In summary, combining these three approaches enhances the detection and interpretation of adaptive, production, and disease-resistance genes in N'Dama cattle. This integrative strategy provides a comprehensive understanding of both historical and ongoing selection processes, essential for the conservation and sustainable genetic improvement of African indigenous cattle.

African cattle breeding can benefit from the systematic and practical application of these tools for a variety of reasons [21]. First, compared to artificial selection for growth and milk in breeding programs, indigenous and composite breeds have experienced different selection regimes and environmental obstacles in harsh rangelands, leading to different genetic imprints. Second, networks of genes linked to metabolism, endocrine regulation, and immune biology mediate the polygenic economic traits under selection in these breeds, including as growth rate, carcass composition, milk output and composition, parasite resistance, and thermotolerance.

Lastly, mapping selection signatures helps prioritize haplotypes to maintain or introgress, offers candidate markers for genomic selection, and helps manage inbreeding by highlighting genomic regions where homozygosity is most concentrated [22,23]. Since Lake Chad is recognized as a source of N'Dama genetics for several tropical countries facing milk production constraints, conserving the genetic diversity of N'Dama cattle in Africa remains a significant challenge. Ensuring the long-term sustainability and productivity of N'Dama breeding programs requires a deeper understanding of the genomic mechanisms underlying autozygosity, haplotype homozygosity and site frequency, which represent an emerging area of research. Accordingly, this study aims to perform genome-wide analyses of autozygosity, haplotype homozygosity and site frequency in N'Dama cattle to identify comprehensive selection signatures associated with milk production and adaptive traits.

MATERIALS AND METHODS

Ethical statement

Genomic analyses were conducted using publicly available sequence data from a cohort of 25 N'Dama cattle. No new animal procedures were performed for this study. All data utilization complied with the relevant ethical guidelines, data-sharing agreements, and standard practices for the use of existing genotype datasets.

Genotype Data and Quality control

Whole-genome SNP genotype data for 25 N'Dama cattle originating from West Africa were retrieved from the WIDDE (Web-Interfaced Next Generation Database for Genetic Diversity Exploration) database [24]. The dataset was generated using the Illumina BovineSNP50 BeadChip, which assays 51998 SNP markers distributed across the bovine autosomes. To

ensure high-quality markers for downstream analyses, standard quality control (QC) filters were applied in PLINK v1.9 [25]. SNPs and individuals were retained if they satisfied the following criteria: Per-SNP call rate ≥ 0.95, Per-individual call rate ≥ 0.95, Minor allele frequency (MAF) ≥ 0.05, Hardy-Weinberg equilibrium (HWE) test p-value > 1×10^{-6}. Autosomal SNPs passing these QC steps were used in all subsequent analyses (Table 1).

Table 1: Quality control/Filtering of the SNPs in N'Dama cattle

Parameters	Total values	Total SNPs filtered	SNPs removed	SNPs retained
Genotype (geno) call rate	>0.95	51998	839	51159
Hardy-weinberg equilibrium (hwe)	0.001	51998	12	51986
Minimum allele frequency (MAF)	0.05	51998	23427	28571
Total (geno/hwe/MAF)	>0.95/0.001/0.05	51998	24278	27720

ROH-Based selections signatures

Heterozygosity and Inbreeding Coefficient (FIS)

Within-breed genetic diversity was quantified by estimating observed heterozygosity (HO) and expected heterozygosity (HE) under Hardy-Weinberg equilibrium using PLINK v1.9. The inbreeding coefficient (FIS) was calculated as:

$FIS=(H_E-H_O)/H_E$

This provided a marker-based estimate of deviation from Hardy-Weinberg equilibrium, reflecting within-population inbreeding.

Detection and Classification of Runs of Homozygosity (ROH)

ROH were identified in PLINK using the --homozyg command set, with parameters optimized for cattle as recommended in previous studies [7,9]. The following thresholds were applied: plink --cow --bfile N'Dama --homozyg --homozyg-snp 50 --homozyg-kb 250 --homozyg-

density 250 --homozyg-gap 1000 --homozyg-window-snp 30 --homozyg-window-het 1 --homozyg-window-threshold 0.05 -out. This is according to window size employed by Illa [26]. ROH were further categorized into five length classes [27] to differentiate between recent and ancient autozygosity: 2-4 Mb, 4-8 Mb, 8-16 Mb, >16 Mb. Short ROH typically represent ancient shared ancestry, whereas long ROH reflect recent parental relatedness or reduced effective population size.

Genomic Inbreeding (FROH) and Identification of ROH Islands

The proportion of the genome covered by ROH was used to estimate realized inbreeding (FROH) following McQuillan method [6]:

FROH = Total length of ROH per individual/Total autosomal genome length covered by SNP. Both per-individual and per-chromosome FROH were computed to capture variation across the genome [28]. Breed-specific ROH islands were defined as genomic regions where SNPs were included in ROH across the top 1-5% of individuals in the dataset, representing loci with the highest frequency of homozygosity. These regions are indicative of selective sweeps or fixation of advantageous alleles.

iHS-Based selecetion signatures

To identify regions under recent positive selection, we employed the integrated haplotype score (iHS), which evaluates the relative extended haplotype homozygosity (EHH) between derived and ancestral alleles at segregating loci [12]. Genotype data were first quality controlled in PLINK v1.9 [27,29], excluding SNPs with call rates below 95%, minor allele frequency (MAF) < 0.05, or significant deviations from Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$). The filtered dataset was then phased into haplotypes using SHAPEIT v2 [30], incorporating breed-specific effective population sizes to enhance phasing accuracy, and subsequently converted to VCF

format for analysis with selscan v2.1 [10]. iHS was calculated by integrating the decay of haplotype homozygosity around each core SNP, comparing ancestral and derived alleles. Resulting iHS scores were standardized within allele frequency bins across the genome to correct for frequency-dependent biases [12], and normalized iHS values exceeding 2 were considered indicative of strong recent positive selection [31].

Tajima'D-Based selection signatures

Tajima's D was calculated to detect deviations from neutral evolution and identify potential regions under selection. Previously filtered SNP genotype data were used for this analysis. The dataset was split by chromosome and converted to Variant Call Format (VCF). Tajima's D statistics were computed using VCFtools v0.1.16 [32], employing a sliding window approach with a 50 kb window size and 25 kb step size to estimate local variations in nucleotide diversity across the genome. The resulting Tajima's D values were aggregated and visualized genome-wide to detect regions exhibiting significant positive or negative deviations, which may indicate balancing or directional selection, respectively [16].

Functional Annotation and Enrichment of Genes and Quantitative Trait loci (QTL) and Overlaps between ROH, iHS and Tajima's D candidate genes

For the QTL annotation and enrichment, ROH islands were identified and then annotated by aligning genomic coordinates to genes in the *Bos taurus* ARS-UCD1.2 reference genome using Ensembl and NCBI genome browsers. To assess their potential economic relevance, overlaps with known quantitative trait loci (QTL) were examined through the Cattle QTL Database (QTLdb_CattleARS_UCD2.gtf) [21], with QTL subsequently grouped into trait categories such as milk production, meat quality, health, adaptation, reproduction, and overall production. Candidate regions of selection were defined as genomic windows harbouring clusters of SNPs

with extreme iHS and Tajima's D values and were annotated to nearby genes using the Bos_taurusARS-UCD2.0.114.gtf reference genome. Functional enrichment analyses were then conducted to explore biological processes associated with selection signatures. Both ROH islands and significant iHS and Tajima's D hotpots regions were subjected to enrichment analysis using the GALLO R software package [33], which facilitated gene and QTL annotation, cross-breed comparisons of functional categories, and QTL enrichment tests via permutation and hypergeometric approaches [21]. Visualization of candidate selection hotspots was performed using Manhattan plots and Venn diagrams generated in GALLO.

RESULTS

Runs of Homozygosity

Runs of Homozygosity (ROH) analysis in Table 2 identified a total of 544 ROH segments distributed across different length classes in the studied cattle population (Table 2). The majority of ROHs were within the 2-4 Mb class (375; 42.25%), contributing the largest proportion of the genome coverage (41.41%). Short ROHs of 1-2 Mb were comparatively rare (36; 2.73%) and accounted for only 2.67% of genome coverage. Intermediate ROHs of 4-8 Mb (78; 17.52%) and 8-16 Mb (36; 16.93%) contributed substantially to genomic autozygosity, covering 17.17% and 16.59% of the genome, respectively. The longest ROHs (>16 Mb) were fewer in number (19 segments) but made up a considerable proportion of the genome (20.17% coverage). The mean ROH length increased progressively with ROH class, from 1.85 Mb in the shortest class to 26.53 Mb in the longest class. The ROH distribution suggests that the studied cattle population harbours both short and long homozygous segments, indicating signatures of both ancient inbreeding and recent demographic events.

Table 2**:** Summary of ROH metrics per breed (nROH count, cumulative ROH length, mean ROH, total length and genome converage) in N'DAMA cattle

Class	nROH	Percent	Cumm. Percent	Mean Length	SD	Total length	Genome Coverage (%)
1-2Mb	36	2.73	2.73	1855.97	550.05	66815.01	2.67
2-4 Mb	375	42.25	44.98	2760.78	1346.81	1035291.13	41.41
4-8 Mb	78	17.52	62.5	5502.29	2002.19	429178.94	17.17
8-16 Mb	36	16.93	79.43	11519.69	3113.67	414708.72	16.59
>16 Mb	19	20.58	100	26534.01	5567.33	504146.25	20.17
Total							

Individual genomic inbreeding

The genomic autozygosity analysis revealed variable levels of inbreeding (measured as $\%F_{ROH}$) across individuals (Figure 1). Most animals displayed relatively low inbreeding levels (<0.5%), but several individuals exhibited substantially higher $\%F_{ROH}$, with peaks reaching up to about 2.5%. This indicates the presence of both highly outbred and moderately inbred individuals within the population. The observed variation in $\%F_{ROH}$ suggests heterogeneity in inbreeding levels within the population.

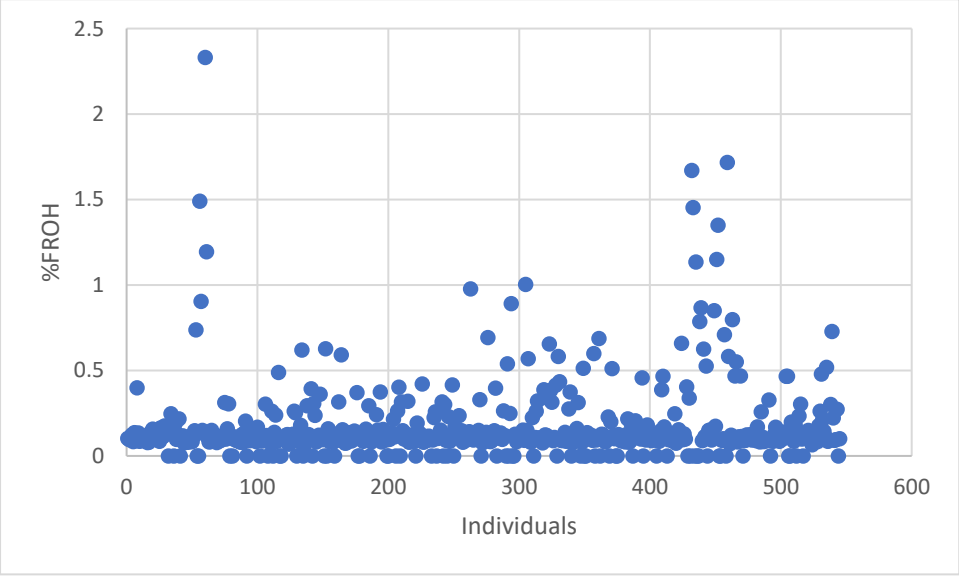


Figure 1. Scatter plot of N'Dama F_{ROH} per individual

Autosomal genomic inbreeding

The chromosomal distribution of runs of homozygosity (%F_{ROH}) in N'Dama cattle revealed variation across the genome (Figure 2). Peaks of homozygosity were particularly evident on chromosomes 3, 6, 16, 17, 19, 20 and 22, where %F_{ROH} exceeded 1-2% in some regions. In contrast, several chromosomes displayed relatively low autozygosity, with most values clustering below 0.5%. These findings indicate that autozygosity is not uniformly distributed but rather concentrated in specific chromosomal regions.

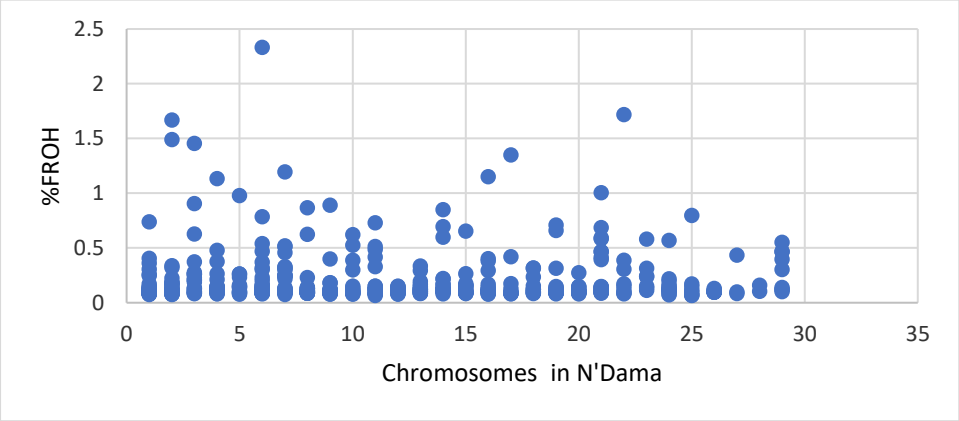


Figure 2: Scatter plot of N'Dama F_{ROH} per chromosomes

Genomic inbreeding coefficient

The mean inbreeding coefficient (FIS) across chromosomes exhibited heterogeneous patterns (Figure 3). Most chromosomes showed values close to zero, indicating Hardy-Weinberg equilibrium or minimal deviations. However, distinct deviations were observed: chromosome 1 displayed a slight excess of homozygosity (positive FIS around 0.02), whereas chromosome 20 exhibited the most pronounced negative FIS (< -0.02), suggesting an excess of heterozygosity. Other chromosomes fluctuated within narrow margins (-0.01 to 0.01). The

chromosomal variation in FIS highlights the non-uniform distribution of inbreeding and genetic diversity across the genome.

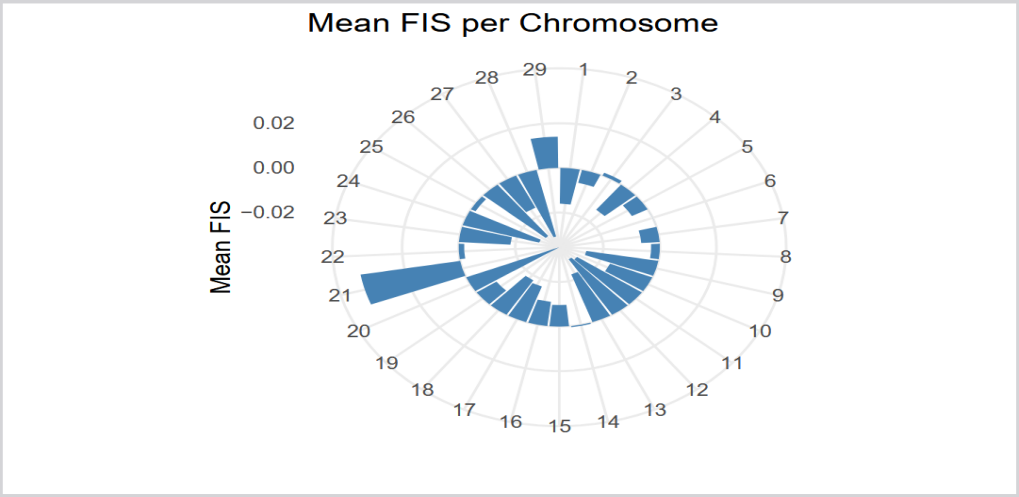


Figure 3: Genomic inbreeding coefficient (FIS) revealing inbreeding in N'DAMA

Genome wide scan with integrated haplotype scores (iHS)

The genome-wide iHS scan for the N'Dama breed revealed several candidate regions under potential positive selection (Figure 4). Most SNPs showed iHS values close to zero, consistent with neutrality, while a subset displayed elevated iHS signals. Using thresholds of iHS ≥ 2 (blue line) and iHS ≥ 3 (red line), we detected multiple loci exceeding these cutoffs across several chromosomes. Notably, strong peaks were observed on chromosomes 5, 6, 13, 14, and 21, suggesting regions potentially associated with adaptation and functional traits in N'Dama cattle.

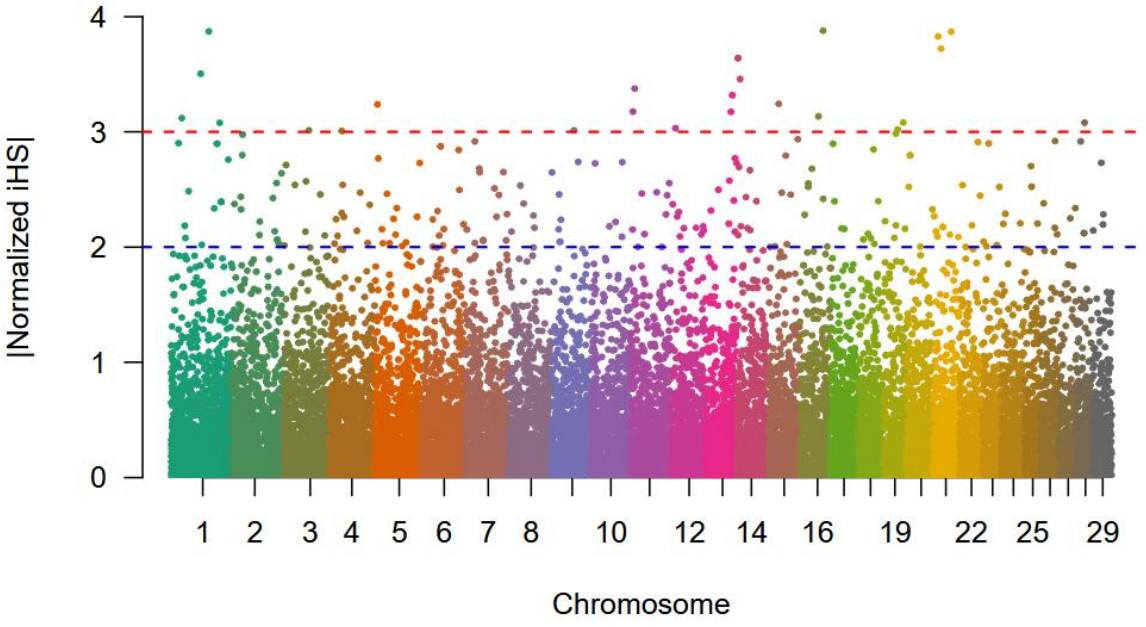


Figure 4: Manhattan plot of iHS showing genome-wide scan for N'Dama using iHS

The genome-wide Tajima's D scan for N'Dama cattle revealed variable patterns of nucleotide diversity across all 29 autosomes (Figure 5). Most genomic regions displayed neutral Tajima's D values around zero, indicating mutation-drift equilibrium. However, several loci showed elevated positive Tajima's D values exceeding the threshold ($D > 2.0$), suggesting regions under balancing selection. These peaks were notably distributed across chromosomes 1, 5, 8, 11, and 17, representing potential selection hotspots associated with adaptive and production traits.

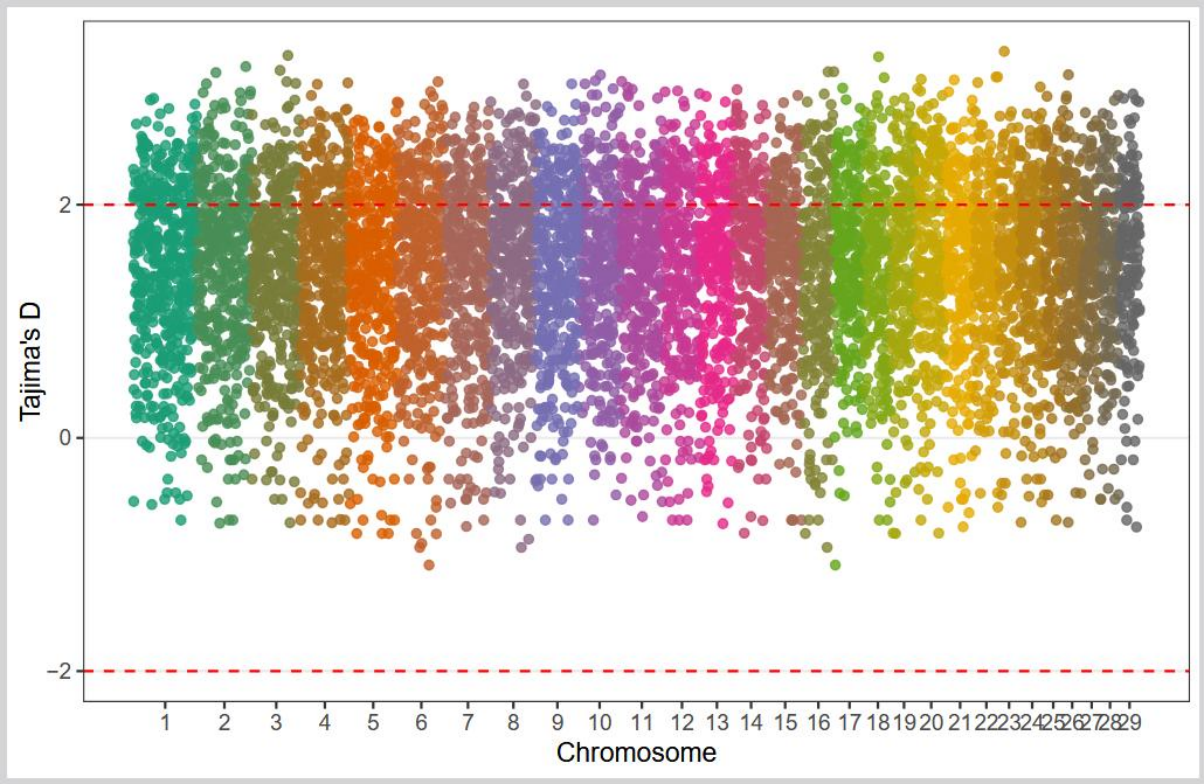


Figure 5: Tajima's D Manhattan plot revealing Genome-wide selection sweep

Gene interaction networks in runs of homozygosity

The protein-protein interaction (PPI) network revealed several highly connected gene clusters (Figure 6). Three major modules were evident. Signal transduction cluster centered on SRC, FGFR1, NRAS, and LYN, with strong connections to genes involved in intracellular signalling (e.g., ITGA4, GAB2, ACTB). This cluster also highlighted SNPs located in bta100 and bta466, suggesting association with chromosomal loci under selection. Futhermore, chromatin and transcriptional regulation cluster was anchored by ZNF, SF1, and H2AC family genes, indicating roles in epigenetic regulation and gene expression modulation. Moreover, metabolic and detoxification cluster were organized around CYP450 family genes (CYP7A1, CYP21A2, CYP17A1), sulfotransferases (SULT1A1, SULT1E1), and flavin-containing monooxygenases (FMO1-FMO5). These are critical for xenobiotic metabolism, hormone regulation, and oxidative stress response.

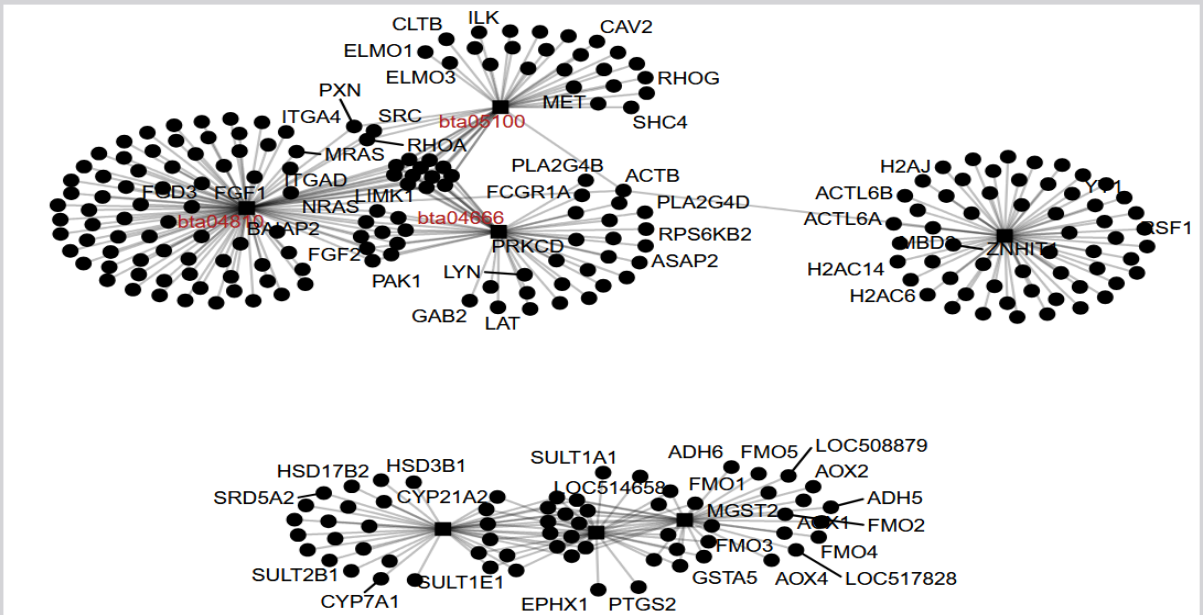


Figure 6: ROH Cytoscape showing Gene interaction network

Genes interaction network in integrated haplotype scores

Pathway enrichment analysis of iHS identified bta04820 (Smooth muscle contraction pathway) as significantly enriched, with several genes connected within the network (Figure 7). Key genes included *ACTG2, MYH1, TPM1, TLN2, ITGA6, ITGA11, COL6A5, COL27A1, THBS1, FBN1,* and *CAPZB*. These genes are functionally linked to cytoskeletal organization, extracellular matrix interactions, and contractile functions. Their strong connectivity highlights coordinated regulation in muscle contraction and structural integrity.

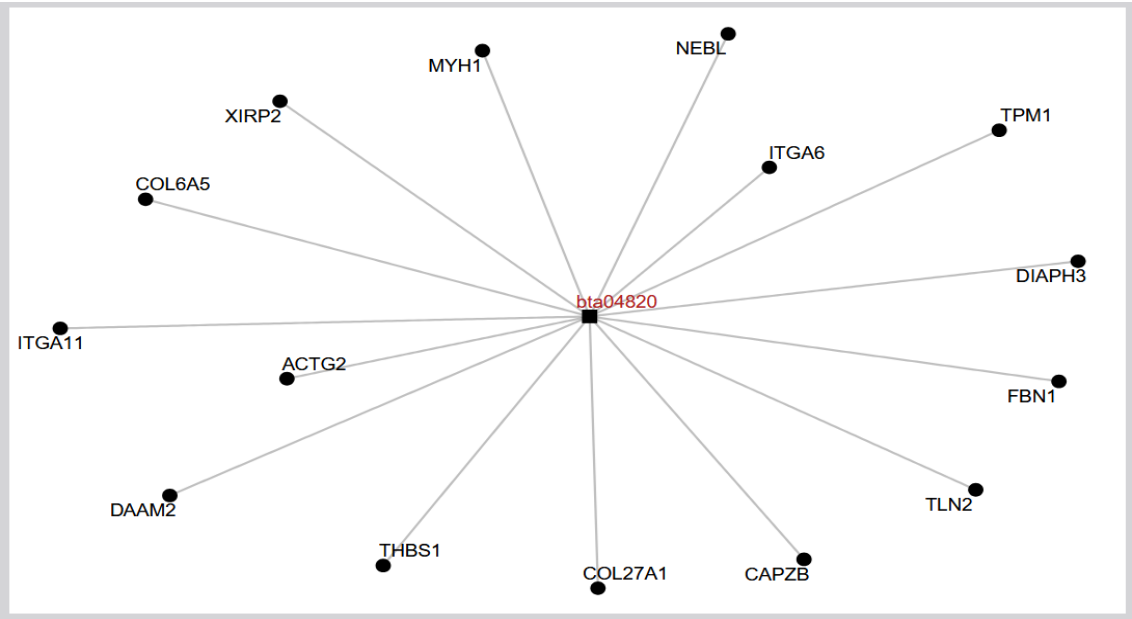


Figure 7: iHS Cytoscape showing Gene interaction network in N'Dama cattle

The gene-gene interaction network for N'Dama (Figure 8) revealed multiple densely connected modules, suggesting extensive molecular interactions among genes under selection. Several hub genes including *ENO3*, *IL1B*, *RB1*, *GSK3A*, and *CCR6* showed high degrees of connectivity, indicating their central roles in the network (Figure 8). These genes form interaction cores linked to both metabolic and immune-related pathways. Peripheral clusters composed of *NUMB*, *HEYL*, *TLE4*, and *DVL2* represent smaller modules associated with cell differentiation and signal transduction.

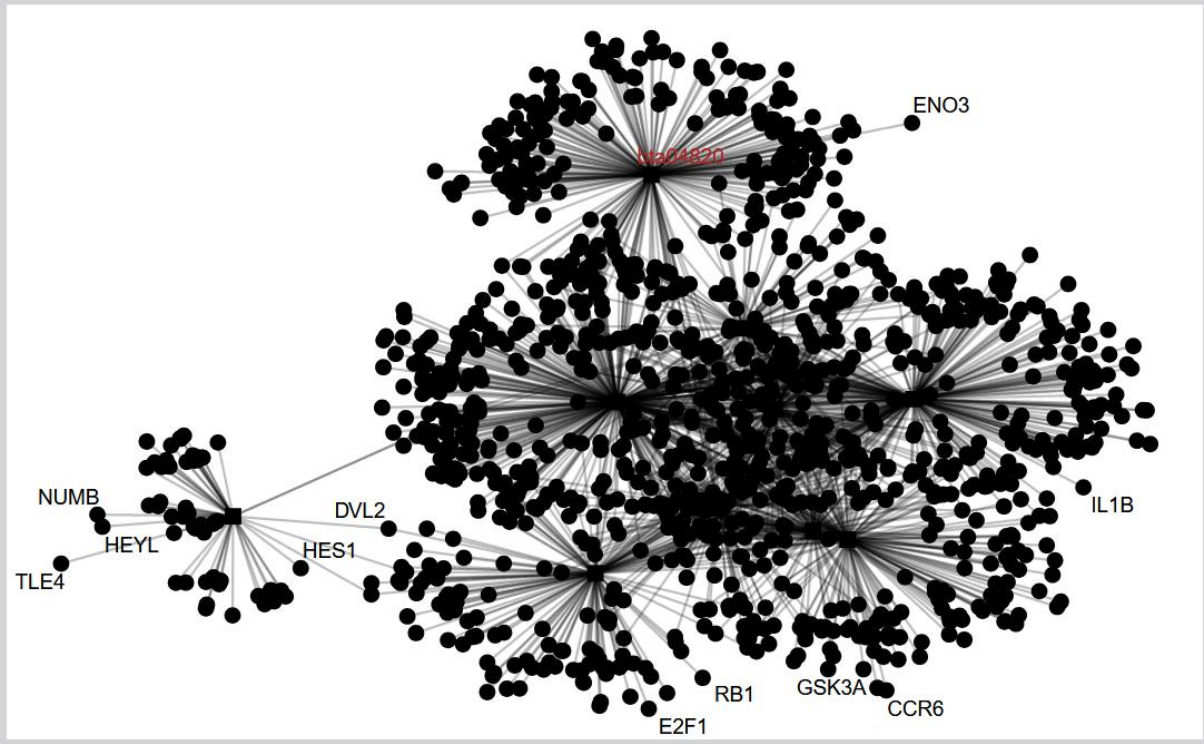


Figure 8: Tajima's D gene interaction network in N'Dama cattle

ROH-Based selection signatures

The runs of homozygosity gene ontology (GO) enrichment analysis of N'Dama candidate genes identified five significantly enriched biological processes and molecular functions (Table 3). Among these, positive regulation of protein metabolic process (GO:0051247) was the most prominent, with genes partitioned into both milk production (*LEP, IL6R, RPTOR, KIT*) and heat stress response (*DNAJB2, MAPKAPK3, SOCS4*). The regulation of immune response (GO:0050776) highlighted several immunity-related genes (*FCGR3A, IFIH1, PDCD1, CTLA4, TLR6*), consistent with the well-documented trypanotolerance and disease resistance in N'Dama cattle. Additionally, endopeptidase activity (GO:0004175) and regulation of actin cytoskeleton organization (GO:0032956) were enriched with calpains (*CAPN1, CAPN2, CAPN3*) and cytoskeletal regulators (*RAC1, ROCK2*), both implicated in stress adaptation and resilience. Therefore, the enriched GO categories suggest that N'Dama cattle balance

production potential (milk-associated pathways) with adaptive advantages (immune and stress resilience), reflecting evolutionary adaptation to tropical environments.

Table 3: Runs of homozygosity gene ontology (GO) enrichment of classification of candidate genes in N'Dama

ID	Description	p.adjust	Genes	Category
GO:0051247	Positive regulation of protein metabolic process	0.020272	*LEP, NCK1, IL6R, PDGFB, FGF23, CSNK1E, KIT, RPTOR, CSNK1D*	Milk Production
GO:0050776	regulation of immune response	0.030173	*FCGR3A, IFIH1, IFI16, PDCD1, CTLA4, FCGR2A, CD1B, TLR6.*	Adaptive / Immunity
GO:0051247	positive regulation of protein metabolic process	0.020272	*ANXA1, DNAJB2, STRADB, OAZ3, SASH1, SOCS4, STRADA, MAPKAPK3*	Heat Stress
GO:0004175	endopeptidase activity	0.00476	*CAPN3,CAPN3, CAPN2CAPN1*	Adaptive/Stress Response
GO:0032956	regulation of actin cytoskeleton organization	0.010612	*CAPZA2, CAPZB, HAX1, RAC1, ROCK2*	Adaptive/Stress Response

iHS-Based selection signatures

GO enrichment analysis revealed significant biological processes associated with both adaptive and milk-related categories. In the adaptive category, enrichment was observed for protein polyubiquitination (GO:0000209; *SMURF2*), indicating a role for ubiquitin-mediated protein

regulation. Processes related to neural development were also prominent, including neurogenesis (GO:0022008; *TNXB*) and generation of neurons (GO:0048699; *TNXB*). Additionally, cell interaction processes such as cell adhesion (GO:0007155; TNXB, CD47, *THBS1*) and cell to cell adhesion (GO:0098609; *CD47*) were significantly represented, suggesting involvement in tissue organization and adaptive structural remodeling. In the milk-related category, strong enrichment was detected for cell morphogenesis (GO:0000902; *CDH2*) and multiple adhesion related processes, including cell adhesion (GO:0007155; *CDH2, TLN2, ITGA11, ITGA6*). Neural development-associated terms such as axonogenesis (GO:0007409), neurogenesis (GO:0022008), neuron differentiation (GO:0030182), generation of neurons (GO:0048699), and neuron projection morphogenesis (GO:0048812) largely driven by *CDH2* were also enriched, reflecting shared molecular pathways between neural development and mammary gland structure and function.

Table 4: iHS Gene ontology (GO) enrichment in N'Dama

ID	Description	p.adjust	Genes	Category
GO:0000209	protein polyubiquitination	0.025255	*SMURF2*	Adaptive
GO:0000902	cell morphogenesis	0.00075	*CDH2*	Milk
GO:0007155	cell adhesion	0.001988	*TNXB, CD47, THBS1*	Adaptive
GO:0022008	Neurogenesis	0.00051	*TNXB*	Adaptive
GO:0022008	Neurogenesis	0.00051	*CDH2*	Milk
GO:0030182	neuron differentiation	0.014796	*CDH2*	Milk
GO:0048699	generation of neurons	0.00075	*TNXB*	Adaptive
GO:0098609	cell-cell adhesion	0.00075	*CD47*	Adaptive
GO:0098609	cell-cell adhesion	0.00075	*CDH2, TLN2, ITGA11, ITGA6*	Milk

Tajima's D-Based selection signatures

The Gene Ontology (GO) enrichment analysis of genes under Tajima's D selection in N'Dama cattle identified multiple biological processes associated with milk production and adaptive functions (Table 5). Among the milk production-related processes, positive regulation of molecular function (GO:0044093; P = 0.002541) and cell surface receptor protein tyrosine kinase signalling pathway (GO:0007169; P = 0.04505) were significantly enriched and represented by *IGF1* and *STAT5B*, respectively-both of which are known regulators of mammary gland development and lactation performance. Additional significant milk-related terms included lipid biosynthetic process (GO:0008610; P = 0.04505; *LPIN1*) and small molecule catabolic process (GO:0044282; P = 0.04505; *GALM*). Adaptive processes were also prominent, with positive regulation of molecular function (GO:0044093; *TLR9, TYRO3*), cell surface receptor protein tyrosine kinase signalling pathway (GO:0007169; *NREP, SMAD6*), and regulation of molecular function (GO:0065009; TLR9) showing strong enrichment. Other adaptive terms included response to nitrogen compound (GO:1901698; SEL1), lipid biosynthetic process (GO:0008610; *CAT*), DNA metabolic process (GO:0006259; *TNF*), and protein transport (GO:0015031; *TXN, AP3B1*).

Table 5: Tajima's D GO genes in N'Dama cattle

GO_Term	Description	p.adjust	Gene	Category
GO:0044093	Positive regulation of molecular function	0.002541	*IGF1*	Milk Production
GO:0007169	Cell surface receptor protein tyrosine kinase signalling pathway	0.04505	*STAT5B*	Milk Production
GO:0008610	Lipid biosynthetic process	0.04505	*LPIN1*	Milk Production
GO:0044282	small molecule catabolic process	0.04505	*GALM*	Milk Production
GO:0007169	Cell surface receptor protein tyrosine kinase signalling pathway	0.04505	*NREP, SMAD6*	Adaptive
GO:0044093	Positive regulation of molecular function	0.002541	*TLR9, TYRO3*	Adaptive
GO:1901698	Response to nitrogen compound	0.000302	*SEL1*	Dual
GO:0065009	Regulation of molecular function	0.005435	*TLR9*	Adaptive
GO:0008610	Lipid biosynthetic process	0.04505	*CAT*	Adaptive
GO:0006259	DNA metabolic process	0.04505	*TNF*	Adaptive
GO:0015031	Protein transport	0.04505	*TXN, AP3B1*	Adaptive

Shared candidate between iHS and ROH

Functional enrichment analysis of overlapped (Shared) candidate genes under selection between iHS and ROH (Table 6 and Figure 9) identified several significant Gene Ontology (GO) terms related to neurogenesis (GO:0022008, P = 0.00051), cell morphogenesis (GO:0000902, P = 0.00075), cell adhesion (GO:0007155, P = 0.0019), and protein polyubiquitination (GO:0000209, P = 0.025). Notably, two genes, *HECW1* and *SLIT2*, were identified across both iHS and ROH signals, indicating robust evidence of selection. *SLIT2* was annotated with roles in both milk production and adaptive responses, while *HECW1* was mainly linked to adaptive functions. Other genes, including *DCDC2*, *ANAPC4*, *HECTD2*,

RFFL, *FER*, and *CAPZB*, were enriched in cellular organization and protein modification pathways, although their direct roles in cattle adaptation or milk traits remain unclear.

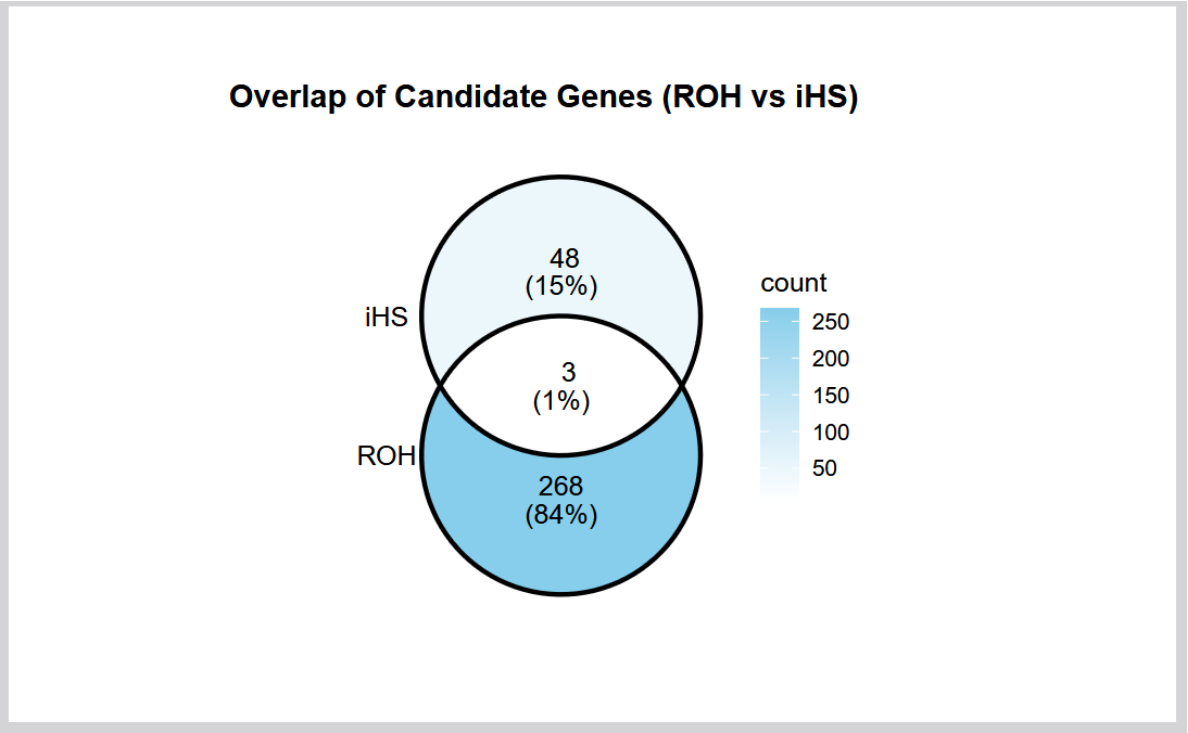


Figure 9: Shared candidate genes between Runs of Homozygosity and Integrated Haplotype Scores

Table 6: Gene ontology (GO) overlapped (shared) genes between iHS and ROH in N'dama cattle

ID	Description	p.adjust	Genes	Category
GO:0022008	Neurogenesis	0.00051	*HECW1*	Adaptive
GO:0022008	Neurogenesis	0.00051	*SLIT2*	Milk and adaptive
GO:0022008	Neurogenesis	0.00051	*DCDC2*	Adaptive
GO:0000902	Cell morphogenesis	0.00075	*CAPZB*	Milk/Adaptive
GO:0007155	Cell adhesion	0.001988	*FER*	Milk and Adaptive
GO:0000209	Protein polyubiquitination	0.025255	*ANAPC4*	Milk
GO:0000209	Protein polyubiquitination	0.025255	*HECTD2*	Adaptive
GO:0000209	Protein polyubiquitination	0.025255	*RFFL*	Adaptive

Shared candidate genes between iHS and Tajima's D

Two candidate genes, SDCCAG8 and EPHB2 were identified in overlapping regions between iHS and Tajima's *D* selection scans in N'Dama cattle (Table 7 and Figure 10). *SDCCAG8*, associated with cell cycle, ciliary function, and stress response (GO:0007049; $P = 0.045$), indicates selection for adaptive mechanisms related to environmental and physiological stress tolerance. *EPHB2*, linked to enzyme-linked receptor protein signalling and mammary gland morphogenesis (GO:0007167; $P = 0.024$), suggests a role in mammary gland development and milk production.

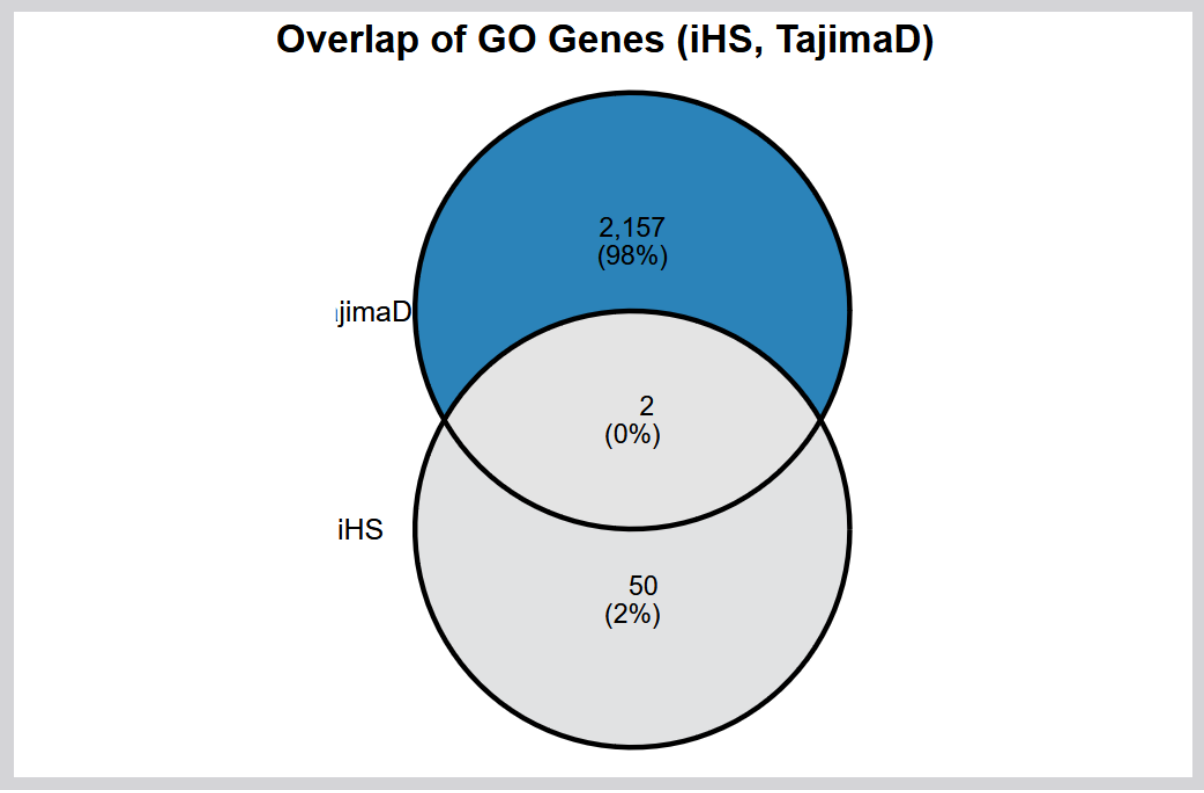


Figure 10: Shared candidate genes between Integrated Haplotype Scores and TajimaD

Table 7: iHS - Tajima's D GO overlapped (Shared) genes in N'DAMA cattle

GO terms	Description	P.adjusted	Genes	Category
GO:0007049	cell cycle, Ciliary function & stress response	0.0450498	*SDCCAG8*	Adaptive
GO:0007167	enzyme-linked receptor protein signalling pathway, Mammary gland morphogenesis	0.023557	*EPHB2*	Milk production

Shared candidate genes between ROH and Tajima's D

The ROH-Tajima's D GO overlap in N'Dama cattle (Table 8 and Figure 10) revealed enrichment of key biological processes primarily related to lipid metabolism and secretion, signifying their strong influence on milk production and adaptive physiology. The most significantly enriched GO term was lipid biosynthetic process (GO:0008610; $P = 4.88\times10^{-106}$), involving *ANXA1*, *SOCS7*, *ALOX15B*, *SPHK2*, *PGAP3*, *DEGS1*, and *ACSM2B*. These genes play crucial roles in fatty acid synthesis, phospholipid remodeling, and sphingolipid signalling, essential for maintaining milk fat composition and membrane fluidity in the mammary gland. Similarly, genes such as *DPM3*, *HMGCS2*, and *FDPS* were enriched in phospholipid and glycerophospholipid biosynthetic processes (GO:0008654; $P = 3.12\times10^{-40}$ and GO:0046474; $P = 2.08\times10^{-35}$), indicating active lipid membrane synthesis required for milk secretion. The sphingolipid biosynthetic process (GO:0030148; $P = 1.52\times10^{-23}$) included *CERS2*, *CERS6*, *STK11*, *GCK*, and *LRP5*, linking lipid metabolism with cellular energy balance and stress response. Additionally, genes like *ANXA1*, *GCG*, *TRH*, and *SSTR5* were enriched under peptide and hormone secretion pathways (GO:0002790, GO:0030072; $P = 1.13\times10^{-8}$), highlighting hormonal control in milk release and mammary function. Adaptive pathways included carbohydrate homeostasis (GO:0033500; $P = 1.76\times10^{-8}$) and regulation of cellular response to

stress (GO:0080135; $P = 0.034$), with genes such as *STK11*, *CYP7A1*, *RNFT1*, and *OTUB1* indicating metabolic and stress adaptation in tropical conditions.

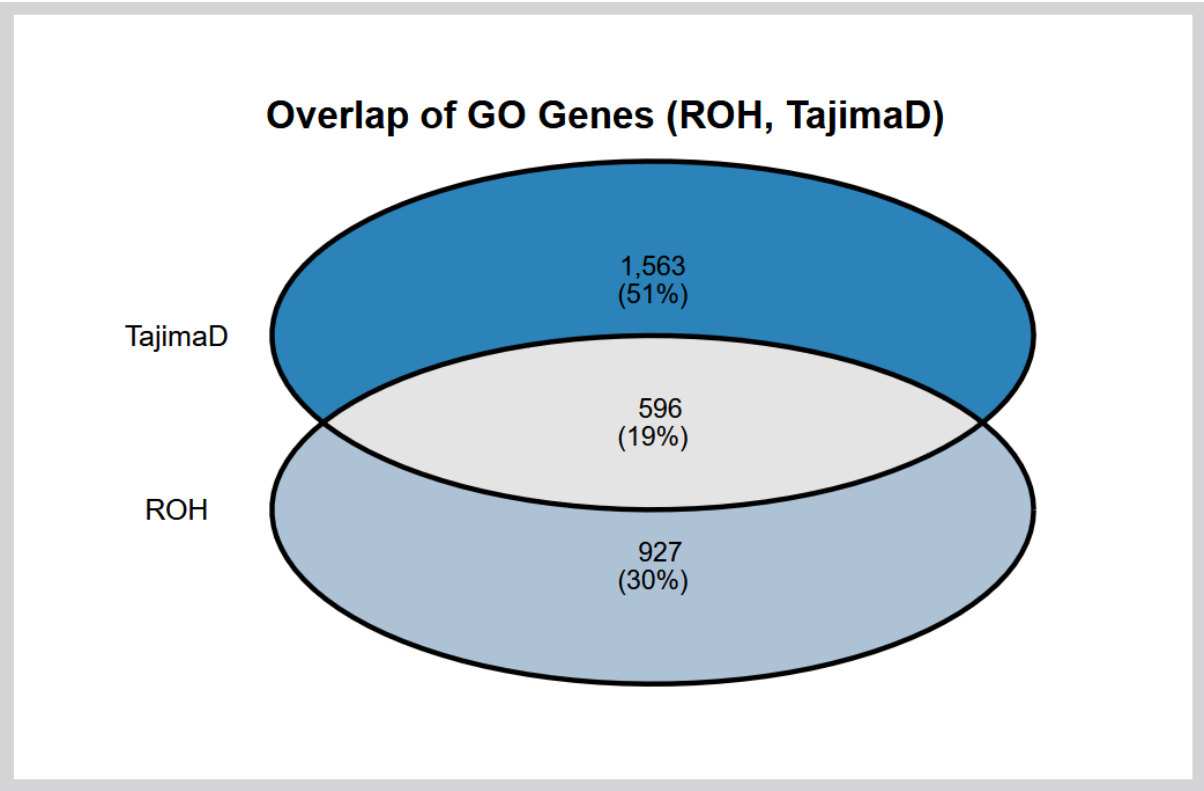


Figure 11: Shared candidate genes between Runs of Homozygosity and TajimaD

Table 8: ROH- Tajima's D GO overlapped (Shared) genes in N'DAMA cattle

	Description	p.adjust	Genes	Category
GO:0008610	lipid biosynthetic process	4.88E-106	*ANXA1, SOCS7, ALOX15B, SPHK2, PGAP3, DEGS1, ACSM2B*	Milk Production
GO:0008654	phospholipid biosynthetic process	3.12E-40	*DPM3, HMGCS2, FDPS, HYCC1, PIP5K1A, INPP*	Milk Production
GO:0046474	glycerophospholipid biosynthetic process	2.08E-35	*PLD2, PTDSS1/BPNT2PIGC, PIGV, LPCAT3, HDHD5*	Milk Production
GO:0046467	membrane lipid biosynthetic process	1.69E-34	*CERS2, GAL3ST2, PIGV, ORMDL1, PLA2G6, AGK*	Milk Production
GO:0006644	phospholipid metabolic process	3.91E-31	*FDPS, HYCC1, CERS2, DPM3, GAL3ST2*	Milk Production
GO:0006650	glycerophospholipid metabolic process	1.50E-27	*CHKA, IP6K2, CISH, SOCS6, INPP5K, INPPL1*	Milk Production
GO:0046486	glycerolipid metabolic process	2.54E-26	*DPM3, HYCC1, SOCS7*	Milk Production
GO:0030148	sphingolipid biosynthetic process	1.52E-23	*CERS2, CERS6, STK11, GCK, LRP5*	Milk Production
GO:0002790	peptide secretion	1.13E-08	*ANXA1, GCG, SSTR5*	Milk Production
GO:0030072	peptide hormone secretion	1.13E-08	*NNAT, HNF1B, TRH, LRP5, SSTR5*	Milk Production
GO:0033500	carbohydrate homeostasis	1.76E-08	*STK11, CYP7A1, ASPSCR*	Adaptation
GO:0006688	glycosphingolipid biosynthetic process	7.02E-08	*UGT, RHBDF1, IGF1R, HNF1B, TRH, LRP5, SSTR5*	Milk Production
GO:0080135	regulation of cellular response to stress	0.034041	*RTEL1, RNFT1, USP1, OTUB1, SHLD1*	Adaptation
GO:0016042	lipid catabolic process	0.038406	*ABHD2, CYP7A1, APOE, APOC2, ACOX2*	Milk Production

Identified milk production and adaptive traits QTLs in ROH

ROH-based QTL enrichment in N'Dama cattle identified numerous loci associated with production, reproduction, exterior, meat and carcass, and milk traits across multiple chromosomes (Figure 9). Exterior traits showed significant enrichment on CHR11 (189 QTLs, adj. P = 1.13E-23) and CHR17 (93 QTLs, adj. P = 2.82E-26). Meat and carcass traits were enriched on several chromosomes, with prominent signals on CHR14 (345 QTLs, adj. P = 4.96E-20) and CHR15 (142 QTLs, adj. P = 2.00E-42), as well as CHR3, CHR16, and CHR17. Milk-related traits exhibited strong enrichment on CHR7 (272 QTLs, adj. P = 2.62E-34), CHR13 (555 QTLs, adj. P = 1.01E-58), and CHR19 (657 QTLs, adj. P = 1.14E-27), consistent with previous reports in Bos taurus breeds involving key milk genes on CHR6, CHR13, CHR14 (DGAT1), and CHR20 (GHR)(S1).

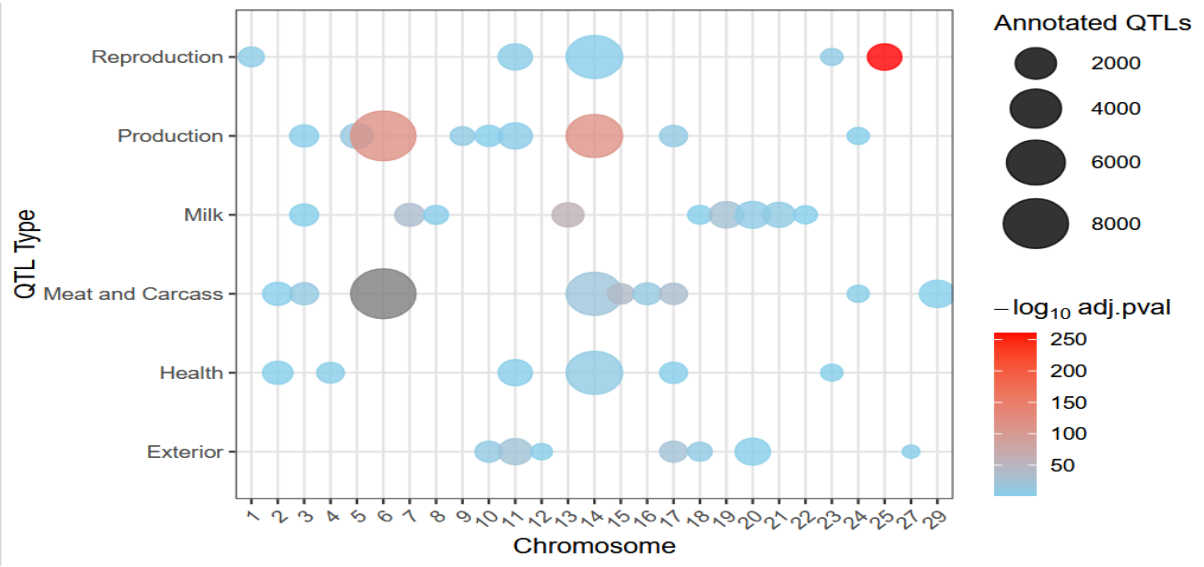


Figure 9: QTL enrichment across chromosomes in ROH

Identified milk production and adaptive traits QTLs in iHS

QTL enrichment analysis for iHS identified significant associations with both milk production traits and health-related traits (Figure 10). For milk traits, strong signals were observed on chromosomes 6, 14, and 25, particularly linked to milk fat yield, casein composition, and fatty acid content. Notably, the most significant associations were for the C18 index (Chr14, adj.P = 5.37E−07) and fat yield (Chr23, adj.P = 7.19E-05; Chr25, adj.P = 0.000354). Casein-related QTLs (*αS2-casein and its phosphorylated form, κ-casein*) were enriched mainly on Chr6 and Chr13, aligning with known milk protein loci. Mineral content traits (zinc, potassium) and fatty acids (linoleic, linolenic, oleic, stearic acids) also showed significant enrichment, emphasizing the nutritional aspects of milk. In the health category, a significant enrichment was detected for bovine tuberculosis susceptibility (Chr20, adj.P = 0.0037) (S2).

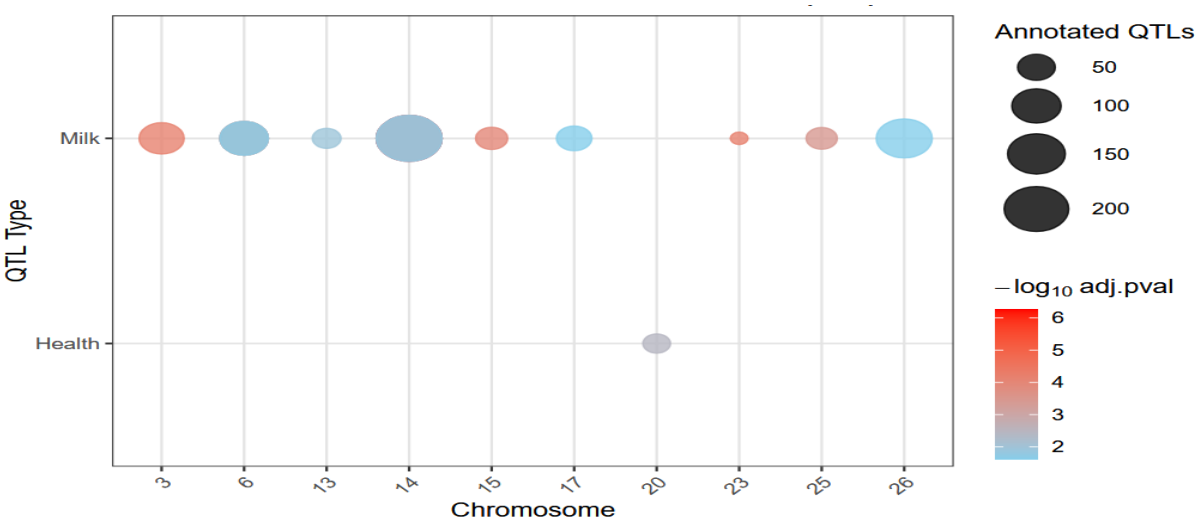


Figure 10: QTL enrichment across chromosomes in iHS

Identified milk production and adaptive traits QTLs in Tajima's D

The analysis of Quantitative Trait Loci (QTLs) under Tajima's D selection in N'Dama cattle identified several significant regions associated with milk production and adaptive traits (Figure 11). For milk-related QTLs, the most enriched signals were detected on chromosomes 5, 6, 16, and 21 with highly significant adjusted *p*-values (e.g., CHR21, $P = 9.21\times10^{-13}$; CHR5, $P = 4.16\times10^{-6}$). These chromosomes harbor QTLs previously associated with milk yield, fat, and protein content, as well as mammary gland development. In contrast, adaptive QTLs were concentrated on chromosomes 4, 10, 19, 20, 21, and 23, with CHR4 ($P = 1.64\times10^{-13}$) and CHR21 ($P = 7.62\times10^{-12}$) showing the strongest signals. These regions correspond to genes involved in immune response, stress tolerance, and morphological adaptation, supporting N'Dama's resilience in tropical environments (S3).

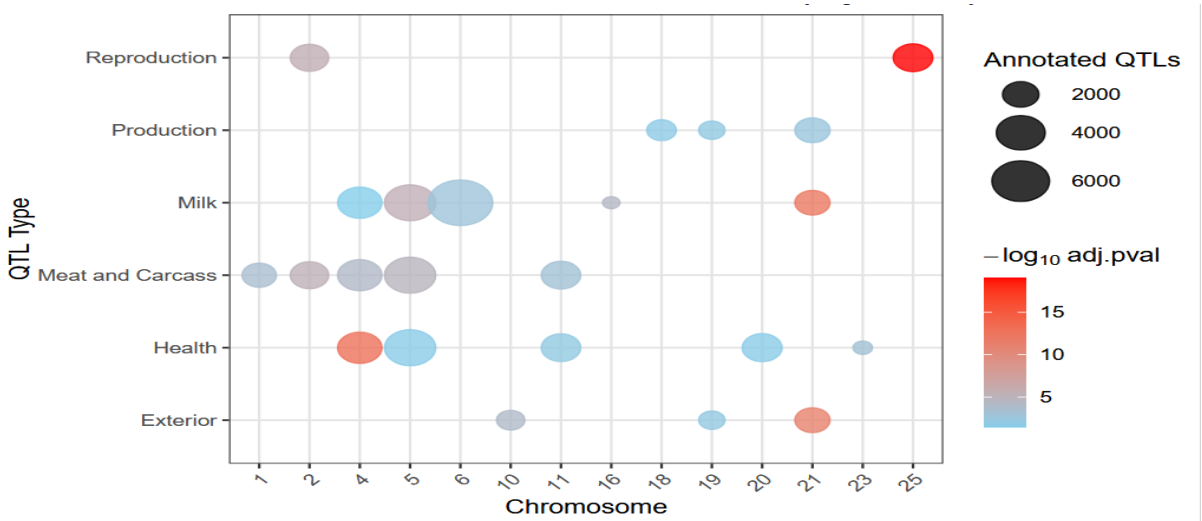


Figure 11: QTL enrichment across chromosomes in Tajima D

DISCUSSION

Runs of homozygosity

The predominance of short ROHs (2-4 Mb) implies the persistence of homozygous segments originating from more ancient common ancestors, where recombination over generations has fragmented longer ROHs into smaller tracts [7,9]. Conversely, the presence of long ROHs (>16 Mb) reflects more recent inbreeding or population bottlenecks, as longer uninterrupted homozygous stretches are less affected by recombination and typically arise from recent parental relatedness [15,34]. The considerable proportion of intermediate ROHs (4 - 16 Mb) further suggests a mixed demographic history, where both historical and more recent events shaped genomic autozygosity. These findings are consistent with reports in other indigenous cattle populations, where moderate to long ROHs indicate the influence of selection for adaptive traits under small effective population sizes [10,35]. The relatively high contribution of longer ROHs (>8 Mb, around 37% of coverage) suggests that recent inbreeding cannot be ruled out, possibly due to restricted breeding practices or small herd sizes, which is typical for indigenous and locally adapted breeds. Therefore, the ROH profile highlights a dual pattern of ancient shared ancestry and recent inbreeding, underscoring the need for management strategies to conserve genetic diversity while maintaining the adaptive potential of the population.

Individual genomic inbreeding

The majority of individuals displayed low autozygosity, implying a broad genetic base and limited recent inbreeding. However, the presence of individuals with higher $\%F_{ROH}$ values points to segments of the population with increased genomic homozygosity, likely resulting from recent common ancestry, bottlenecks, or small effective population sizes [7]. The relatively low mean $\%F_{ROH}$ across the population is consistent with the traditional breeding and management practices in indigenous cattle, where natural mating and lack of intensive

selection often maintain genetic diversity [36]. Nonetheless, the outlier individuals with high F_{ROH} may represent closed breeding within herds or geographically isolated subpopulations, raising potential concerns for long-term genetic health and adaptability [37]. Overall, these findings highlight the need for careful monitoring of genomic autozygosity, particularly in conservation and breeding programs, to prevent the accumulation of deleterious alleles and loss of genetic variability.

Autosomal genomic inbreeding

The presence of chromosomal hotspots of ROH suggests the influence of selection, demographic history, or linkage disequilibrium patterns. Chromosomes 3 and 6, where pronounced peaks were detected, have been previously associated with immune response genes, trypanotolerance loci, and milk production QTLs in N'Dama and other African taurine cattle [3,38]. Elevated homozygosity in these regions may reflect selection for adaptive traits, such as disease resistance and environmental tolerance, which are critical for the survival of N'Dama in tsetse-infested and humid tropical environments. Conversely, the low $\%F_{ROH}$ across many chromosomes suggests that the overall genomic diversity is preserved, consistent with the relatively low levels of inbreeding reported for N'Dama populations [39]. The heterogeneity of $\%F_{ROH}$ across chromosomes highlights the importance of region-specific selective pressures rather than genome-wide fixation.

Manhattan genome wide scan with iHS and Tajima's D

These findings indicate ongoing or incomplete selective sweeps, consistent with the utility of iHS for identifying recent selection [10,12]. The significant iHS signals identified in N'Dama cattle likely reflect historical and recent selection related to adaptation and production traits.

N’Dama is renowned for its trypanotolerance, disease resistance, and adaptation to harsh tropical environments [4,40]. Selection signals on chromosome 21 may be linked to disease resistance loci, consistent with previous reports of immune-related genes in African taurine cattle [13]. Likewise, strong signals on chromosome 6 are noteworthy, as this region harbors QTLs associated with milk production, casein genes, and immune functions in Bovine QTLdb [38]. Furthermore, signals on chromosome 14 coincide with known QTLs for milk fat and fatty acid composition [41,42], suggesting that although N’Dama is primarily adapted for survival traits, there is also evidence of selection on production-related loci. The presence of multiple peaks across the genome highlights the polygenic nature of adaptation in indigenous breeds [43]. To sum it all, the iHS results demonstrate the interplay between natural and artificial selection in shaping the N’Dama genome. The combination of health related and production-related signatures emphasizes the dual importance of environmental adaptation and economic traits in breed evolution.

The presence of positive Tajima’s D peaks in N’Dama cattle reflects signatures of balancing selection, a pattern typical of populations exposed to diverse environmental pressures, such as parasites, heat stress, and nutritional fluctuations [3,39]. The detected regions likely harbor functional variants that contribute to disease resistance and environmental adaptability, consistent with N’Dama’s well-documented trypanotolerance and heat tolerance. Similar studies have shown that balancing selection maintains advantageous alleles at intermediate frequencies, promoting immune system diversity and metabolic resilience [44]. In N’Dama, this adaptive pattern aligns with the enrichment of genes such as TLR9, TNF, and CAT, identified in the GO analysis, which are involved in innate immunity, oxidative stress regulation, and metabolic adaptation. The combined evidence underscores the role of balancing

selection in shaping N'Dama's genetic architecture for environmental fitness and sustainable milk production.

Gene interaction networks in ROH, iHS and Tajima's D

In ROH, the network highlights the functional diversity of candidate genes detected in selection regions such as the *SRC,FGFR,NRAS* signalling axis suggests roles in cell proliferation, immune signalling, and growth regulation, consistent with reports of selective pressures in cattle linked to disease resistance and productivity traits [39]. Meanwhile, the *ZNF, SF1* and *H2A* cluster implicates chromatin remodeling and transcriptional control, which may underlie adaptive gene expression plasticity in response to environmental challenges. In addition to this, the *CYP450*, *SULT* and *FMO* cluster reflects metabolic adaptation, especially in detoxification pathways, which is crucial for cattle exposed to diverse diets, pathogens, and environmental stressors [45]. Together, these findings indicate that genomic regions under selection are enriched for genes in signalling, immunity, transcriptional regulation, and metabolism, reflecting the combined pressures of adaptation, productivity, and resilience in cattle populations.

Meanwhile, in iHS, **t**he enrichment of smooth muscle contraction pathway (bta04820) is biologically relevant to cattle physiology. Genes such as *MYH1* (myosin heavy chain 1), *TPM1* (tropomyosin 1), and *ACTG2* (actin gamma 2) are directly involved in muscle fibre contraction and sarcomere organization [46,47]. Structural proteins like *COL6A5, COL27A1,* and *FBN1* contribute to extracellular matrix remodeling, which is essential for tissue strength and resilience [48]. Importantly, several integrin-related genes (*ITGA6, ITGA11*) and cytoskeleton regulators (*TLN2, DIAPH3*) play roles in mechanotransduction and muscle adaptation [49]. These pathways may influence not only muscle physiology but also traits linked to growth,

meat quality, and overall resilience in cattle. In African taurine breeds such as N'Dama and Kuri, enrichment of smooth muscle-related pathways could reflect adaptation to endurance, energy efficiency, and disease tolerance, as smooth muscle contraction is critical for vascular function, thermoregulation, and immune responses [13,40]. This suggests that selection in these cattle may favor structural and metabolic efficiency, which contributes to their survival in harsh tropical environments.

Moreover, in Tajima's D, the observed high connectivity of *ENO3* (enolase 3) reflects its key function in energy metabolism, particularly in glycolytic processes that sustain tissue activity during stress or lactation [39]. The clustering of *IL1B* (*interleukin-1 beta*) underscores strong immune and inflammatory signalling, consistent with adaptive responses to environmental and pathogenic pressures [39]. The presence of *GSK3A* (glycogen synthase kinase 3 alpha) and *RB1* (*retinoblastoma gene*) within the network center suggests selection on cell-cycle regulation and metabolic control, both vital for tissue growth and repair [50]. *CCR6* (*C-C* chemokine receptor 6), another highly connected node, mediates immune cell trafficking, supporting adaptation to infection challenges. Peripheral nodes such as *NUMB*, *HEYL*, *TLE4*, and *DVL2* are linked to the Notch and Wnt signalling pathways, which regulate cell fate determination and developmental plasticity [51]. The integration of these functional clusters highlights the interplay between metabolic efficiency, immune competence, and developmental regulation, supporting evidence of polygenic adaptation in N'Dama cattle.

Selection signatures for milk production and adaptive traits

The enrichment of candidate ROH genes in N'Dama cattle reveals a dual pattern of selection for production and adaptation traits. Genes involved in positive regulation of protein metabolism, including *LEP*, *IL6R*, *PDGFB*, *RPTOR*, and *CSNK1D/E*, highlight pathways regulating energy balance, lactation efficiency, mammary gland function, and mTOR signalling

associated with milk protein synthesis [52,53]. Concurrently, enrichment of immune response genes such as *FCGR2A/FCGR3A*, *PDCD1*, *CTLA4*, and *TLR6* underscores roles in antibody-dependent responses, immune checkpoint regulation, and pathogen resistance, supporting N'Dama's well-documented trypanotolerance [54]. Genes related to heat and stress tolerance, including *DNAJB2*, *SOCS4*, *MAPKAPK3*, and calpains (*CAPN1-3*), reflect molecular chaperone, proteolytic, and signalling mechanisms that buffer cellular stress [55,56]. Additionally, enrichment in cytoskeletal regulation through genes such as *RAC1* and *ROCK2* suggests roles in stress adaptation, apoptosis resistance, and thermotolerance [57]. Collectively, these findings demonstrate that N'Dama cattle maintain genetic variants that support both milk production traits and adaptive resilience, reflecting the selective pressures of tropic.

In iHS, the enrichment of cell adhesion and morphogenesis pathways highlights their dual role in adaptation and milk production. Adhesion molecules such as *CDH2* (N-cadherin) and *ITGA6/ITGA11* regulate cell-cell interactions, mammary gland development, and tissue remodeling, processes essential for milk synthesis and secretion [58]. The repeated detection of *CDH2* across multiple neuron- and adhesion-related terms suggests it plays a pleiotropic role linking neurodevelopmental signalling with lactation physiology. Adaptive signals were particularly strong for *TNXB*, a gene implicated in extracellular matrix stability and stress response, which could contribute to resilience under environmental challenges [59]. Similarly, *CD47* and *THBS1* are known for their roles in immune modulation and cellular adhesion, supporting their potential contribution to adaptive immune defense. Overall, the findings suggest that adhesion and neurogenesis-related genes (*CDH2, TNXB, ITGA6*) may function at the interface of production and adaptation, reflecting selective pressures for both milk yield and environmental resilience.

Moreover, The Tajima's D-based GO enrichment in N'Dama cattle highlights selective pressures on both milk production and adaptation-related genes, reflecting the dual-purpose nature and environmental resilience of this trypanotolerant West African breed. Genes such as IGF1 (*Insulin-like Growth Factor 1*) and STAT5B are crucial regulators of mammary gland development, cell proliferation, and milk synthesis, with established roles in enhancing milk yield and composition. Both *IGF1* and *STAT5B* had been identified in multiple breeds [21] as growth factors. Similarly, *LPIN1*, involved in lipid biosynthetic process, influences milk fat synthesis through its role in triglyceride and phospholipid metabolism. The identification of *GALM*, related to *small molecule catabolic process*, indicates metabolic optimization for energy utilization during lactation. The adaptive gene set reflects immune regulation and stress response mechanisms essential for survival in humid and tsetse-infested environments. *TLR9* and *TYRO3* are pattern recognition receptors participating in innate immune signalling and pathogen defense, while *TNF* (*tumor necrosis factor*) modulates inflammatory responses under environmental and parasitic stress. *TLR9* has been reported to involve in innate immune system which are ubiquitous across climatic regions [60,61]. *CAT* (*Catalase*) contributes to lipid biosynthetic and antioxidant defense, protecting cells from oxidative damage. *NREP* and *SMAD6*, part of the tyrosine kinase signalling pathway, regulate cell differentiation and tissue repair, possibly contributing to thermotolerance and disease resilience. Thes findings align with that of Merve [21] and Dzomba [62] who reported genes such as *SMAD2, ESR2, DMC1* which were implicated in oocyte production, growth and seasonal reproduction. The enrichment of dual-function genes such as *SEL1* (*response to nitrogen compound*) further demonstrates an adaptive metabolic plasticity, enabling N'Dama cattle to maintain production efficiency under nutritional and environmental constraints. Overall, these findings emphasize the integrated selection for milk productivity and environmental adaptation in N'Dama cattle consistent with

its long history of natural and artificial selection for trypanotolerance and sustainable production in sub-Saharan Africa.

Shared candidate genes between iHS and ROH

The detection of genes such as *SLIT2* and *HECW1* in both iHS and ROH analyses highlights their potential importance in cattle adaptation. *SLIT2* is known to regulate axon guidance and neurogenesis, but in livestock it has also been linked to mammary gland development and milk-related pathways. Enhanced alveoli development and increased milk production have noticed in the SLIT2/SLIT3 knockout mammary gland epithelia [63]. Its dual annotation in our dataset (milk and adaptive) suggests that selection on this gene may underlie both production and environmental resilience. Similarly, *HECW1*, an HECT E3 ubiquitin-protein ligase, plays a role in protein homeostasis and cellular stress responses, supporting its classification as an adaptive candidate [64]. The enrichment of processes such as protein polyubiquitination and cellular organization suggests that protein quality control and structural regulation are important targets of selection in these cattle breeds. Previous studies have reported that ubiquitination-related genes contribute to cellular adaptation under environmental stress [65,66]. Genes with unclear roles (*ANAPC4, HECTD2, RFFL, DCDC2, FER, CAPZB*) may represent background selection signatures or have species specific functional relevance not yet fully characterized in cattle. Nevertheless, their enrichment in core cellular processes points to potential indirect contributions to adaptation. Overall, these findings suggest that selection in Kuri, N'Dama, and Bororo cattle is acting on a combination of neurogenesis, protein regulation, and cell organization pathways, with *SLIT2* and *HECW1* emerging as strong candidates underlying both adaptive and milk production traits.

Shared candidate genes between iHS and Tajima's D

These genes reflect concurrent selection pressures on adaptation and milk production traits in N'Dama cattle, consistent with their dual importance in tropical environments. The *SDCCAG8* which is functionally linked to NEK8 gene, has been found to be associated with cell cycle and stress response, has been implicated in cellular organization and ciliogenesis [67]. Its involvement in stress adaptation suggests a potential role in environmental resilience and disease tolerance, key adaptive features of N'Dama cattle [4]. Conversely, *EPHB2* functions in cell signalling and mammary gland morphogenesis [68], linking it to mammary epithelial development and milk production. Selection signals at this locus indicate evolutionary pressure on milk-related pathways, consistent with the breed's dual-purpose (milk and adaptive) nature. The overlapping signals between iHS and Tajima's D thus highlight genomic regions subject to both recent and historical selection, underscoring the adaptive-productive balance that defines N'Dama cattle genetics.

Shared candidate genes between ROH and Tajima's D

The overlapping GO signatures between ROH and Tajima's D analyses suggest long-term selection and adaptive balancing of genes governing lipid metabolism and milk-related functions in N'Dama cattle. The dominance of lipid biosynthetic and metabolic processes underscores the evolutionary importance of efficient lipid utilization for milk synthesis, even under environmental stress. Genes such as *ANXA1*, *SPHK2*, and *CERS2* are integral to phospholipid signalling, sphingolipid metabolism, and fatty acid turnover, processes that are critical for maintaining mammary gland integrity and milk fat yield [69,70] and have also been associated with regulation of cardiovascular diseases, inflammatory injury and diabetes [71,72]. Similarly, *PLA2G6* and *ACSM2B* contribute to membrane lipid remodeling and energy

balance, consistent with adaptation to nutritional fluctuations typical of tropical environments and as well as maintaining neurodegeneration [73,74]. The enrichment of *GCG*, *TRH*, and *SSTR5* within peptide and hormone secretion pathways points to a tight endocrine regulation of lactation, reinforcing the coordination between lipid metabolism and hormonal signalling [75]. The involvement of *STK11* and *CYP7A1* in glucose and bile acid homeostasis further supports adaptive metabolic efficiency, which is crucial for N'Dama's resilience to thermal and nutritional stress [76]. Therefore, the shared GO terms highlight a functional convergence between selection signals and homozygosity regions, reflecting the dual optimization of milk productivity and environmental adaptation in N'Dama cattle.

Identified milk production and adaptive traits QTLs in ROH, iHS and Tajima's D

In ROH, production traits showed the highest overall enrichment, particularly on CHR6 and CHR14. Reproduction traits were dominated by a major hotspot on CHR25 and this has been reported to be aligning with known associations for fertility related traits such as calving ease, gestation length, and ovarian function [77]. The enrichment patterns reflect the polygenic architecture of adaptive and productive traits in N'Dama cattle, a trypanotolerant African taurine breed renowned for disease resistance and environmental adaptability [4,44]. Our results revealed high QTLs on CHR14 and CHR15 in this study. Strong signals on CHR14 and CHR15 are consistent with regions harboring genes affecting growth, muscle development, and carcass traits [78.79], while enrichment on CHR6 and CHR14 overlaps known selection sweeps for growth and feed efficiency, an observation that agreed with that of Flori et al. [80]. Generally, these results highlight CHR25 as a key reproductive adaptation hotspot and CHR6, CHR13, and CHR14 [81,82] as central hubs integrating milk, growth, and meat traits, supporting the dual adaptive–productive potential of N'Dama cattle.

However, in iHS, the enrichment of QTLs for milk composition and yield in this finding confirms the major role of chromosomes 6, 14, and 25 in dairy cattle production traits. CHR6 harbors the casein gene cluster, explaining its association with casein percentages and milk minerals [69,83]. Similarly, CHR14 is well-established as a hotspot for fatty acid metabolism genes (notably *DGAT1*), which drives variation in milk fat yield and composition [83,84]. The strong signal for the C18 index on CHR14 supports the importance of this region in shaping milk fat quality. The enrichment for minerals (zinc, potassium) and fatty acids (linoleic, linolenic, oleic, stearic) further highlights the genetic basis of nutritional quality in milk [70,83,85]. These findings are highly relevant for improving functional milk properties in breeding programs. On the adaptive side, the identification of QTLs for bovine tuberculosis susceptibility (CHR20) suggests a possible overlap between production and resilience traits, indicating selection pressures balancing milk yield and health [38,86]. This supports the integration of disease resistance loci into genomic selection strategies to enhance both productivity and animal health.

Meanwhile, in Tajima's D, the identification of significant QTLs across multiple chromosomes highlights the polygenic nature of selection in N'Dama cattle, which combines both milk production efficiency and environmental adaptation. Chromosomes 5 and 6, showing strong milk-related QTL signals, are consistent with previous reports linking these regions to milk yield and composition traits. This report aligns with that of Lopdell [83] Bionaz and Loor [69] who emphasized the association of CHR6 with milk casein percentage and minerals. Genes such as LPIN1 and STAT5B, located within these QTLs, regulate lipid biosynthesis and lactation signalling, emphasizing their functional importance for milk productivity under tropical stress conditions. The adaptive QTLs observed on chromosomes 4, 10, 20, and 23

overlap with regions previously associated with immune defense, thermoregulation, and parasite resistance [39,44,81]. Notably, CHR4 and CHR21 harbored both health and milk-related QTLs, suggesting possible pleiotropic effects where genes contribute to both productivity and adaptation which is a hallmark of genomic balancing selection in African indigenous breeds. Overall, these findings demonstrate that N'Dama cattle have undergone strong natural and artificial selection targeting dual-purpose traits, reinforcing their economic and ecological significance in tropical livestock systems.

CONCLUSION

Chromosome-level analyses revealed breed-specific homozygosity peaks in N'Dama cattle, with ROH length distributions indicating both recent inbreeding and ancient demographic events. Overlapping ROH islands and iHS hotspots highlighted genes linked to milk production, thermal tolerance, and immune adaptation, reflecting the breed's trypanotolerant nature. Functional enrichment pointed to key QTLs for milk yield, composition, and disease resistance. ROH regions contained genes for milk protein synthesis (e.g., *LEP, IL6R, KIT, RPTOR*) and thermal adaptation (*DNAJB2, MAPKAPK3*), while iHS hotspots captured recent selection on milk (*CDH2, ITGA6*) and adaptive (*TNXB, THBS1*) genes. Tajima's D identified frequency-based selection on both milk (*IGF1, STAT5B*) and adaptive (*TLR9, TNF*) genes. Shared signals between ROH-iHS (*DCDC2, SLIT2*), ROH-Tajima's D (*ANXA1, SOCS7*), and iHS-Tajima's D (*SDCCAG8, EPHB2*) reinforce complementary detection of selection across different timescales. Although no single gene overlapped among all three methods, their convergence on related functional pathways underscores the combined power of ROH, iHS, and Tajima's D in revealing genomic regions underlying milk production and adaptive traits in N'Dama cattle.

Supplementary information

The supplementary data used in this study are attached.

Acknowledgements

We gratefully acknowledge research support grant provided by UNESCO-DBT-TWAS to carryout this study and the enabling environment and teaching aid provided by the Animal Genetics and Breeding Division, ICAR-National Dairy Research Institute, India.

Authors' contributions

Badmus Kazeem Ajasa and Anupama Mukherjee conceived and designed the study, and Badmus Kazeem Ajasa conducted the research, data analysis and wrote the first draft. Popoola Moshood edited the first draft. Udhav Paneru revised the manuscript. All authors read and approved the final manuscript.

Funding

Enabling research environment and study grant was provided by UNESCO-DBT-TWAS-2023 with UTR number-ICIC52025090900536591; date-09092025 and is acknowledged.

Availability of data and materials

The genotyping data of African cattle breeds used for conducting this research were already deposited into the WIDDE public data repository. http://widde.toulouse.inra.fr/widde/widde/main.do?module=cattle

Ethics approval and consent to participate

Animal Care Committee approval was not obtained for this study as all the analyses were performed using pre-existing databases.

Consent for publication

Not applicable.

Competing interests

The authors declared no competing interests for this study.

Authors' details

[1]Federal University Gashua, Yobe State, 631101, Nigeria, [2]ICAR-National Dairy Research Institute, 132001, Haryana, India, [3]NARC-National Cattle Research Programm, Rampur, Chitwan, Nepal, [4]Federal College of Animal Health and Production Technology, Ibadan, Nigeria

REFERENCES

1. Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JEO. African pastoralism: genetic imprints of origins and migrations. Science 2002;296:336-339.

2. Hanotte OTD, Kemp S. Time to tap Africa's livestock genomes. Science (Washington) 2010;328:1640-1641.

3. Taye M, Kim J, Yoon SH, Lee W, Hanotte O, Dessie T, et al. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. BMC Genet. 2017, 18, 11. [CrossRef]

4. Hanotte O, Ronin Y, Agaba M, Nilsson P, Gelhaus A, Horstmann, R. et al. Mapping of quantitative trait loci controlling trypanotolerance in a cross of tolerant West African N'Dama and susceptible East African Boran cattle. PNAS, 2003;100(13), 7443-7448.

5. Murray M, Trail JC, D'Ieteren GD. Trypanotolerance in cattle and prospects for the control of trypanosomiasis by selective breeding. Rev Sci Tech. 1990;9(2):369-86. doi: 10.20506/rst.9.2.506. PMID: 2132686.

6. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al.. Runs of homozygosity in European populations. *American Journal of Human Genetics*, 2008;83(3), 359-372.

7. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. *BMC Genetics*, 2012;13:70. https://doi.org/10.1186/1471-2156-13-70

8. Keller MC, Visscher PM, Goddard ME. Quantification of inbreeding due to distant ancestors and its detection using dense SNP data. *Genetics*, 2011;189(1), 237-249.

9. Ceballos FC, Joshi PK, Clark DW, RamsayM, Wilson JF. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 2018;19(4), 220-234.

10. Szpiech ZA, Hernandez RD. selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. Mol. Biol. Evol. 2014;31(10), 2824-2827.

11. Kim ES, Cole JB, Huson H, Wiggans GR, Van Tassell CP, et al. Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. PLOS ONE 2013;8(11): e80813. https://doi.org/10.1371/journal.pone.0080813

12. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biology, 2006;4(3), e72. https://doi.org/10.1371/journal.pbio.0040072

13. Bahbahani H, Clifford H, Wragg D, Mbole-Kariuki MN, Van Tassell C, Sonstegard T, wt al. Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. Scientific Reports, 2015;5, 11729. https://doi.org/10.1038/srep11729

14. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. A meta-assembly of selection signatures to reveal candidate genes for adaptation and production traits in sheep. Scientific Reports. 2016; 6, 27613. https://doi.org/10.1038/srep27613

15. Peripolli E, Munari DP, Silva MVGB, Lima ALF, Irgang R, Baldi F. Runs of homozygosity: current knowledge and applications in livestock. Anim. Genet. 2017;48(3), 255-271. https://doi.org/10.1111/age.12526

16. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 1989;123(3), 585-595. https://doi.org/10.1093/genetics/123.3.585

17. Qanbari S. On the Extent of Linkage Disequilibrium in the Genome of Farm Animals. Front. Genet. 2020;10:1304. doi: 10.3389/fgene.2019.01304

18. Biswas S, Akey JM. Genomic insights into positive selection. Trends in Genetics, 2006;22(8), 437-446. https://doi.org/10.1016/j.tig.2006.06.005

19. Oleksyk TK, Smith MW, O'Brien SJ. Genome-wide scans for footprints of natural selection. Phil. Trans. R. Soc. B: Biological Sciences, 2010;365(1537), 185-205. https://doi.org/10.1098/rstb.2009.0219

20. Ruiz-Larrañaga O. et al. Genomic selection signatures in sheep from the Western Pyrenees. Genet. Selection Evol. 2018;50, 9.

21. Mervan, B, Şükrü D, Mesut K, Bülent B, Yalçın Y. Comparative Detection of Selection Signatures in Indigenous and Crossbred Turkish Sheep Breeds, Research Square, PREPRINT (Version 1) available at Research Square. https://doi.org/10.21203/rs.3.rs-7364156/v1;2025.

22. Bayraktar M, Shoshin O. Association between CAST and MSTN gene polymorphisms with growth traits in Awassi sheep. Kuwait J. Science, 2022;49(2). https://doi.org/10.48129/kjs.10955

23. Ceccobelli, S. et al. A comprehensive analysis of the genetic diversity and environmental adaptability in worldwide Merino and Merino-derived sheep breeds. Genet. Selection Evol. 2023;55, 24

24. Sempéré G, Moazami-Goudarzi K, Eggen A, Laloë D, Gautier M, Flori L. WIDDE: a Web-Interfaced next generation database for genetic diversity exploration, with a focus on bovine species. BMC Genomics, 2015;16, 940. https://doi.org/10.1186/s12864-015-2141-0

25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D.et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;3:559-75. doi: 10.1086/519795.

26. Illa SK, Mukherjee S, Nath S, Mukherjee A. Genome-Wide Scanning for Signatures of Selection Revealed the Putative Genomic Regions and Candidate Genes Controlling Milk Composition and Coat Color Traits in Sahiwal Cattle. Front. Genet. 2021;12:699422. doi: 10.3389/fgene.2021.699422

27. Vyas J, Kumar I, Chitra A *et al.* Longitudinal genomic and pedigree inbreeding metrics reveal inbreeding depression in milk quality traits of Sahiwal cattle maintained in organised herd. *Nucleus* 2025;9 https://doi.org/10.1007/s13237-025-00611-9

28. Vyas J, Kumar I, Chitra A. *et al.* Variation in genetic diversity evaluated under different breeding schemes in Sahiwal herd. Trop Anim Health Prod 2025;**57**, 318. https://doi.org/10.1007/s11250-025-04564-8.

29. Fleiss JL. Statistical methods for rates and proportions, 2nd ed. Wiley, New York; 1981.

30. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nature Methods, 2012;9(2), 179-181. https://doi.org/10.1038/nmeth.1785

31. Ben-Jemaa S, Senczuk G, Ciani E, Ciampolini R, Catillo G, Boussaha M, et al. Genome-Wide Analysis Reveals Selection Signatures Involved in Meat Traits and Local Adaptation in Semi-Feral Maremmana Cattle. Front. Genet. 2021;12:675569. doi: 10.3389/fgene.2021.675569

32. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics, 2011;27(15):e2156-2158. https://doi.org/10.1093/bioinformatics/btr330

33. Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *GigaScience*, 2020;9(12): giaa149.

34. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. *PLoS One*, 2010;5: e13996.

35. Mastrangelo S, Ciani E, Marsan PA, et al. Conservation status and historical relatedness of Italian cattle breeds. Genet. Sel. Evol. 2016;48: 35.

36. Ouédraogo D, Ouédraogo-Koné S, Yougbaré B, Soudré A, Zoma-Traoré B, Mészáros G, Khayatzadeh N, Traoré A, Sanou M, Mwai OA, Wurzinger M, Burger PA, Sölkner J. Population structure, inbreeding and admixture in local cattle populations managed by community-based breeding programs in Burkina Faso. J Anim Breed Genet. 2021;138(3):379-388. doi: 10.1111/jbg.12529.

37. Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science, 2014;166, 26-34. https://doi.org/10.1016/j.livsci.2014.05.034

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

39. Bahbahani H, Salim B, Almathen F, Al Enezi F, Mwacharo J, Hanotte O. Signatures of positive selection in African cattle reveal adaptive diversity to environmental challenges. Gen. Biol. Evol. 2018;10(4), 1142-1158. https://doi.org/10.1093/gbe/evy062

40. Flori L, Thevenon S, Dayo GK, Senou M, Sylla S, Berthier D, et al. Adaptive admixture in the West African bovine trypanotolerant population. Mol. Ecol. 2014;23(13), 3241-3257.

41. Stella A, Ajmone-Marsan P, Lazzari B, Boettcher P. Identification of selection signatures in cattle breeds selected for dairy production. *Genetics*, 2010; 185(4), 1451-1461.

42. Raven LA, Cocks BG, Hayes BJ. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics, 2014;15, 62.

43. Edea Z, Kim KS. Genomic signatures of selection in African cattle revealed by whole genome SNP analysis. BMC Genetics, 2014;15, 80.

44. Tajima F. Simple methods for testing molecular clock hypothesis. Genetics 1993;135, 599– 607

45. Bickhart DM, Liu GE. The challenges and importance of structural variation detection in livestock. Front. Genet. 2014;5, 37. https://doi.org/10.3389/fgene.2014.00037

46. Weber A, Pennise CR, Babcock GG, Fowler VM. Tropomodulin caps the pointed ends of actin filaments. J Cell Biol. 1994;127(6 Pt 1):1627-35. doi: 10.1083/jcb.127.6.1627. PMID: 7798317; PMCID: PMC2120308.

47. Otey CA, Carpen O. Alpha-actinin revisited: a fresh look at an old player. Cell Motil Cytoskeleton. 2004;58(2):104-11. doi: 10.1002/cm.20007. PMID: 15083532.

48. Kielty CM, Sherratt, MJ, Shuttleworth CA. Elastic fibres.J. Cell Sci. 2002;115(14), 2817-2828.

49. Legate KR, Wickström SA, Fässler R. Genetic and cell biological analysis of integrin outside-in signalling. Genes Dev. 2009;15;23(4):397-418. doi: 10.1101/gad.1758709. PMID: 19240129.

50. Lichtinghagen, Ralf & Huber, René. (2023). GSK3 as a Master Regulator of Cellular Processes. IJMS. 24. 15503. 10.3390/ijms242115503.

51. Katoh M. WNT signalling in stem cell biology and regenerative medicine. Curr Drug Targets. 2008;9(7):565-70. doi: 10.2174/138945008784911750. PMID: 18673242.

52. Block SS, Butler WR, Ehrhardt RA, Bell AW, Van Amburgh ME, Boisclair YR. Decreased concentration of plasma leptin in periparturient dairy cows is caused by negative energy balance. J Endocrinol. 2001;171(2):339-48. doi: 10.1677/joe.0.1710339.

53. Cai Z, Dusza M, Guldbrandtsen B, Lund MS, Sahana G. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. Genet. Sel. Evol. 2020;52:19. doi: 10.1186/s12711-020-00538-6

54. Noyes H, Brass A, Obara I, Anderson S, Archibald AL, Bradley DG, et al. Genetic and expression analysis of cattle identifies candidate genes in pathways responding to trypanosome infection. Proc Natl Acad Sci *USA*, 2011;108(22), 9304-9309. https://doi.org/10.1073/pnas.1013486108

55. Goll DE, Thompson VF, Li H, Wei W, Cong J. The calpain system. Physiological Reviews, 2003;83(3), 731-801. 10.1152/physrev.00029.2002

56. Kampinga HH, Craig EA. The HSP70 chaperone machinery: J. Proteins as drivers of functional specificity. Nat Rev Mol Cell Biol. 11(8):579-92. doi: 10.1038/nrm2941. Erratum in: Nat Rev Mol Cell Biol. 2010;11(10):750.

57. Ilnitskaya AS, Litovka NI, Rubtsova SN, Zhitnyak IY, Gloushankova NA Actin Cytoskeleton Remodeling Accompanied by Redistribution of Adhesion Proteins Drives Migration of Cells in Different EMT States. Cells, 2024;13, 780. https://doi.org/ 10.3390/cells13090780

58. Burgoyne RD, Duncan JS, Mather IH. Adhesion molecules in the mammary gland: Roles in development and disease. J. Mammary Gland Biol. and Neop., 2020;25(2), 145-160. https://doi.org/10.1007/s10911-020-09467-2

59. Bruschi M, et al. The extracellular matrix and autoimmunity: the case of the tenascin family of glycoproteins. *Front. Immun.* 2013;4, 39. https://doi.org/10.3389/fimmu.2013.00039

60. Wanjala G. et al. Genetic diversity and adaptability of native sheep breeds from different climatic zones. Sci. Rep. 2025;15, 14143

61. Patiabadi, Z. et al. Whole-genome scan for selection signature associated with temperature adaptation in Iranian sheep breeds. PLOS ONE. 2024;19, e0309023

62. Dzomba, E. et al. Selection signature analysis and genome-wide divergence of South African Merino breeds from their founders. Front. Genet. 2023;13, 932272.

63. Cazares O, Chen M, Menendez J, Molinuevo R, Thomas G, Cervantes J, Yee M, Cadell M, Durham M, Zhu Y, Strietzel C, Bubolz JW, Hinck L. SLIT Loss or Sequestration Increases Mammary Alveologenesis and Lactogenesis. MicroPubl Biol.2024;6;2024:10.17912/micropub.biology.001264.doi:10.17912/micropub.biology.001264.

64. Glickman MH, Ciechanover A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiological Reviews, 2002;82(2), 373-428. https://doi.org/10.1152/physrev.00027.2001

65. Pickart CM. Mechanisms underlying ubiquitination. Annu Rev Biochem.2001;70:503-33. doi: 10.1146/annurev.biochem.70.1.503. PMID: 11395416.

66. Komander D, Rape M. The ubiquitin code. Annu Rev Biochem. 2012;81:203-29. doi: 10.1146/annurev-biochem-060310-170328.

67. Otto EA, Trapp ML, Schultheiss UT, Helou J, Quarmby LM, Hildebrandt F. NEK8 mutations affect ciliary and centrosomal localization and may cause nephronophthisis. J Am Soc Nephrol. 2008;19(3):587-92. doi: 10.1681/ASN.2007040490.

68. Vaught D, Brantley-Sieders DM, Chen J. Eph receptors in breast cancer: roles in tumor promotion and tumor suppression. Breast Cancer Res. 2008;10(6):217. doi: 10.1186/bcr2207.

69. Bionaz M, Loor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics. 2008;31:9:366. doi: 10.1186/1471-2164-9-366.

70. Bernard L, Leroux C, Chilliard Y. Expression and nutritional regulation of lipogenic genes in the ruminant mammary gland. Adv. Exp. Med. Biol. 2008;606, 67-108.

71. Purvis GSD, Solito E, Thiemermann C.. Annexin-A1: Therapeutic Potential in Microvascular Disease. Front Immunol. 2019;30;10:938. doi: 10.3389/fimmu.2019.00938.

72. Qi Y, Wang W, Song Z, Aji G, Liu XT, and Xia P. Role of Sphingosine Kinase in Type 2 Diabetes Mellitus. Front. Endocrinol. 2021;11:627076. doi: 10.3389/fendo.2020.627076

73. Bauman DE, Griinari JM. Nutritional regulation of milk fat synthesis. Annu Rev Nutr. 2003;23:203-27. doi: 10.1146/annurev.nutr.23.011702.073408.

74. Gregory A, Kurian MA, Maher ER, et al. PLA2G6- Associated Neurodegeneration. In: Adam MP, Bick S, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2025;2017 Bookshelf URL: https://www.ncbi.nlm.nih.gov/books

75. Wall E, Coffey MP, Brotherstone S. The relationship between fertility, health and milk production in dairy cattle. Animal, 2015;9(1), 176-187.

76. Rizzolo D, Buckley K, Kong B, Zhan L, Shen J, Stofan M, Brinker A, Goedken M, Buckley B, Guo GL. Bile Acid Homeostasis in a cyp7a1 and cyp27a1 Double Knockout Mouse Model. Hepatology. 2019;70(1):389-402. doi: 10.1002/hep.30612.

77. Frischknecht M, Bapst B, Seefried FR, Signer-Hasler H, Garrick D, Stricker C; Intergenomics Consortium; Fries R, Russ I, Sölkner J, Bieber A, Strillacci MG, Gredler-Grandl B, Flury C. Genome-wide association studies of fertility and calving traits in Brown Swiss cattle using imputed whole-genome sequences. BMC Genomics. 2017;18(1):910. doi: 10.1186/s12864-017-4308-z.

78. Casas E, Shackelford S, Keele J. Koohmaraie M. Smith T, Stone R. Detection of quantitative trait loci for growth and carcass composition in cattle. J. Ani. Sc. 2003;81. 2976-83. 10.2527/2003.81122976x.

79. Porto-Neto LR, Kijas JW, Reverter A. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. Genet. Sel. Evol' 2013;45: 17.

80. Flori L, Fritz S, Jaffrézic F. et al. The genome response to artificial selection: a case study in dairy cattle. PLoS One, 2009;4(8): e6595.

81. Kühn C, Thaller G, Winter A, Bininda-Emonds OR, Kaupe B, Erhardt G, Bennewitz J, Schwerin M, Fries R. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics. 2004;167(4):1873-81. doi: 10.1534/genetics.103.022749.

82. Karim L, Takeda H, Lin L, et al.. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nature Genetics, 2011;43: 405-413.

83. Lopdell TJ. Using QTL to Identify Genes and Pathways Underlying the Regulation and Production of Milk Components in Cattle. Animals (Basel). 2023;2;13(5):911. doi: 10.3390/ani13050911.

84. Schennink A, Stoop WM, Visker MH, Heck JM, Bovenhuis H, van der Poel JJ, et al.,DGAT1 underlies large genetic variation in milk-fat composition of dairy cows. Anim Genet. 2007;38(5):467-73. doi: 10.1111

85. Matsumoto H, Sasaki K, Bessho T, Kobayashi E, Abe T, Sasazaki S. et al. The SNPs in the ACACA gene are effective on fatty acid composition in Holstein milk. Mol. Biol. Rep. 2012;39:8637–8644.

86. Bermingham ML, Brotherstone S, Berry DP. Evidence for genetic variance in resistance to tuberculosis in Great Britain and Irish Holstein-Friesian populations. BMC Proc 2011;**5** (Suppl 4), S15. https://doi.org/10.1186/1753-6561-5-S4-S15

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.xls