

RESEARCH

Open Access



Genomic analysis reveals population structure and selection signatures in plateau dairy cattle

Shangzhen Huang^{1†}, Longgang Ma^{1†}, Bin Li², Jinhuan Dou^{3*}, Qing Xu^{4*} and Yachun Wang^{1*}

Abstract

Background To solve the problem of an insufficient supply of dairy products in Tibet, work has been carried out to improve native dairy cattle and introduce purebred dairy cattle from low-altitude areas. The harsh environment of the plateau not only severely limits the production performance of high-yielding dairy cattle, such as Holstein and Jersey cattle, but also challenges their survival. The population structure and plateau adaptation mechanism of plateau dairy cattle are rarely reported. In this study, key genes and pathways affecting plateau purebred and crossbred dairy cattle were explored using genetic chip information.

Results The results showed that the genetic diversity of the Tibet dairy cattle population was higher than that of the native cattle and plains dairy cattle. Purebred Holstein and Jersey cattle in Tibet were genetically closer to dairy cattle in the plains, and crossbred dairy cattle were admixed with more Tibet cattle and Apaijiza cattle. Based on the fixation index (F_{ST}), integrated haplotype score (iHS), and cross-population extended haplotype homozygosity (XP-EHH) approaches, 60 and 40 genes were identified in plateau Holstein and Jersey cattle, respectively. A total of 78 and 70 genes were identified in crossbred cattle compared to Holstein and Tibet cattle respectively. These genes are related to cardiac health and development, neuronal development and function, angiogenesis and hematopoietic, pigmentation, growth and development, and immune response.

Conclusions Our results provide a glimpse into diverse selection signatures in plateau dairy cattle, which can be used to enhance our understanding of the genomic basis of plateau adaptation in dairy cattle. These results support further research on breeding strategies such as marker-assisted selection and gene editing in plateau dairy cattle populations.

Keywords Holstein cattle, Crossbred cattle, Tibet native cattle, Plateau adaptation, Population structure, Selection signature

[†]Shangzhen Huang and Longgang Ma contributed equally to this work.

*Correspondence:

Jinhuan Dou

doujinhuan@bua.edu.cn

Qing Xu

qingxu@bjtu.edu.cn

Yachun Wang

wangyachun@cau.edu.cn

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



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

The Qinghai–Tibet Plateau is located in western China and has an average elevation of more than 4,000 m above sea level. This high-altitude area is characterized by extreme environmental conditions, including low oxygen concentrations, alpine cold temperatures and high intensities of ultraviolet (UV) radiation, which restrict the survival of domestic animals [1]. Milk is indispensable for the Tibetan people, and due to the insufficient lactation performance of yaks and local cattle, a considerable portion of dairy products need to be transferred from the mainland [2]. The Tibet Autonomous Region has been utilizing Holstein cows for local cattle improvement since the early 1980s, and the limited stock of crossbred cows is still unable to meet the large demand for dairy products in Tibet [3, 4]. A large number of good-breed dairy cattle have been introduced to Tibet since 2014, the mortality of Holstein is 26.5% and the mortality of Jersey is 8%. They have not demonstrated long-term adaptability to the plateau environment in terms of either disease resistance or production performance [5], to the extent that the original breeding objectives and appointment requirements could not be met. Therefore, it is urgent to make it clear that how plateau dairy cattle adapt to plateau environment and improve their performance through genetic breeding.

Plateau adaptation is considered a complex trait that is influenced by multiple genes [6, 7]. A series of genes with functions related to energy metabolism, oxidative response, stress response, hypoxic response, temperature acclimatization, cardiovascular system and body morphology were considered as candidate genes in studies on selection signatures for high-altitude acclimatization [8–10]. Several studies have used genomic data from different breeds of cattle and yaks in different regions to analyze genetic variation resulting from selection at high altitudes, to explore the mechanisms of plateau adaptation and to search for genetic markers [10–14]. Since plateau dairy cattle are still undergoing plateau habituation and such groups are relatively rare, the analysis of their population structure and selection signatures will differ from previous studies.

In this study, the SNP chip data was used to analyze the genetic diversity and population structure of the populations in two large-scale dairy farms in Lhasa, to determine the genetic status of dairy cattle populations and to provide a research basis for genetic improvement and breeding planning for plateau dairy cattle. Selection signatures analysis was also used to detect genomic differences between purebred dairy cattle at different altitudes, as well as genomic differences between hybrid cattle and possible parent breeds. The aim was to reveal the genes that were selected after several generations of

introduction of dairy cattle to the plateau, as well as the genes carried in the crossbred cows that differed from those of the parents. The present results can be useful for future plateau adaptation studies and genetic improvement of dairy cattle in Tibet.

Materials and methods

Animal information and genomic data

The plateau dairy cattle came from Gaba Ecological Farm (Tibet, China) and Zhizhao Industrial Park's high-standard dairy farm (Tibet, China). Two plains dairy cattle populations used for the comparison were from Shounong Animal Husbandry Jingyindao Farm (Beijing, China) and Hebei Shounong Modern Agricultural Science and Technology Co. (Hebei, China). Three plateau native cattle populations used for comparison came from herdsman's homes in different parts of Tibet. About 2 mL of blood was collected from the tail-heel vein of each cattle and Blood Cards were submitted to Neogen Bio-Scientific Technology (Shanghai) Co., Ltd. (Shanghai, China) for processing. This study did not involve the euthanasia or sacrifice of animals.

DNA from each blood card sample was extracted using the Sbeadex Livestock Nucleic Acid Livestock Extraction Kit (LGC Biosearch Technologies, Petaluma, CA). DNA samples were quantified using NanoDrop™ 8000 (Thermo Scientific, Waltham, MA) and genotyped according to the Illumina Infinium XT protocol (Illumina, San Diego, CA). SNP markers were scored on the Bovine Geneseek Genomic Profiler –100 K Beadchip or Bovine Geneseek Genomic Profiler –150 K Beadchip (Neogen Inc, Lincoln, NE) at Neogen. The beadchips were scanned using the Illumina iScan (Illumina) and the raw data was processed and exported using Genome Studio 2.0 software according to the Genome Studio Framework User Guide (Illumina) based on selection criteria of at least 90% animal call rate.

Because accurate pedigrees were not available and crossbred dairy cattle were indistinguishable from Holsteins in coat color, we obtained the bloodline composition calculated from chip data at Neogen. The percentage of DNA contributed to the animal by each of 5 evaluated breeds: Holstein, Jersey, Brown Swiss, Ayrshire, and Guernsey was estimated and termed as Breed base representation (BBR) [15]. BBR values represent the breed composition of an individual that is more accurate and much easier to interpret than breed-check markers. A BBR value of 94 to 99% is typically set to 100%, which representing animals with 100% purebred ancestry. Therefore, the plateau dairy cattle were categorized to 5 subpopulations according to the BBR values: 119 crossbred dairy cattle with 40%~74% Holstein bloodlines (GH40), 92 dairy cattle with 75%~84% Holstein

bloodlines (GH75), 92 dairy cattle with 85%~93% Holstein bloodlines (GH85), 129 Holstein cattle with 94%~100% Holstein bloodlines (GH100) and 28 Jersey cattle with 76%~100% Jersey bloodlines (GJ). The genomic data used in this study were shown in Table 1.

The liftover tool [16] was used to update the reference genome version from Bos_taurus_UMD_3.1 to ARS-UCD1.2 for the 150 K chip. Data from chips of the same density were merged and subjected to quality control, including (1) excluding individuals with a genotyping detection rate less than 0.9; (2) excluding SNPs with a genotyping detection rate less than 0.9; (3) excluding SNPs with a minor allele frequency (MAF) less than 0.01; (4) excluding sites severely deviating from Hardy–Weinberg equilibrium ($P < 10^{-6}$); and (5) retaining SNPs located on autosomes and the X chromosome. The Beagle (v5.2) [17] tool was used to impute missing genotypes in the quality-controlled data. After retaining common sites between the two versions of the chips, a total of 752 cattle with 62,422 SNPs were obtained for subsequent analysis.

Genetic diversity and linkage disequilibrium

The average MAF and observed heterozygosity rate (HET) were estimated using PLINK (v1.90) [18]. The parameters for detecting Runs of homozygosity (ROH) were set as follows: for each 70 kb segment of ROH, there must be one SNP; the interval between consecutive SNPs in the same ROH must be less than 500 kb; one heterozygous locus is allowed; only ROH with a length exceeding 30 SNPs and greater than 500 kb are detected. The ratio of ROH length to chromosome length was taken as inbreeding coefficient (F_{ROH}). To estimate and compare decay patterns of linkage disequilibrium (LD) at the genome-wide level among populations and

subpopulations, PopLDdecay v.3.40 [19] was used to calculate the square correlation coefficient (r^2) between SNP pairs up to 500 kb. The effective population size 13 generations ago (N_{E13}) was estimated from LD by using the SNeP tool [20].

Population structure

The genetic distance based on identity-by-state (IBS) was calculated using PLINK (v1.90), and multidimensional scaling (MDS) was carried out to show the populations and subpopulations structure. The genetic distance matrix between individuals, computed through VCF2Dis (v1.47) [21], was uploaded to FastME (v2.0) [22] for the construction of a neighbor-joining evolutionary tree, which was ultimately visualized using iTOL (v6) [23]. Based on the LD decay result, PLINK (v1.90) was submitted to use the command “– indep-pairwise 50 10 0.2” to remove linkage sites, followed by the use of ADMIXTURE (v1.3.0) [24] to conduct population structure analysis under the assumption of ancestral numbers $k = 1 \sim 20$.

Selection signature detection

Detection of selection signatures in the plateau dairy cattle genome was based either on the evaluation of allele frequency differences among populations and on properties of haplotypes segregating in the populations. The applied methods included within-population (integrated haplotype score, iHS) and between-population (fixation index, F_{ST} and cross-population extended haplotype homozygosity, XP-EHH) tests. In comparisons between the same breeds at different altitudes, the within population signatures of selection in GH100 and GJ were computed using iHS, and between-population tests were applied to identify potential sweeps that occurred in

Table 1 Sources of experimental animal and genetic data

Population	Region	Breed	Number	Chip type
Plateau dairy cattle				
GH100	Lhasa, Tibet	Holstein cattle (94~100% Holstein bloodlines)	129	150 K SNP chip
GH85	Lhasa, Tibet	cows with 85%~93% Holstein bloodlines	92	150 K SNP chip
GH75	Lhasa, Tibet	cows with 75%~84% Holstein bloodlines	92	150 K SNP chip
GH40	Lhasa, Tibet	crossbred cattle (40%~74% Holstein bloodlines)	119	150 K SNP chip
GJ	Lhasa, Tibet	Jersey cattle (76%~100% Jersey bloodlines)	28	150 K SNP chip
Plains dairy cattle				
PH	Beijing	purebred Holstein Cattle	57	150 K SNP chip
PJ	Dingzhou, Hebei	purebred Jersey cattle	77	150 K SNP chip
Plateau native cattle				
XZ	Shigatse, Tibet	Tibet Cattle	61	100 K SNP chip
AP	Nyingchi, Tibet	Apaijiza Cattle	39	100 K SNP chip
TF	Shigatse, Tibet	Shigatse Hump Cattle	59	100 K SNP chip

the GH100 compared to PH and GJ compared to PJ. In comparisons between crossbred dairy cattle and their possible parent breeds, the within population signatures of selection in GH40 were computed using iHS, and between-population tests were applied to identify potential sweeps that occurred in the GH40 compared to PH and XZ.

F_{ST} detection for each SNP was performed using vcftools (v0.1.16) [25]. After phasing with Beagle (v5.2) [17], iHS and XP-EHH was calculated using selscan (v1.3.0) [26], and the XP-EHH values were standardized using the norm in selscan (v1.3.0). Large positive or negative iHS values indicate unusually long haplotypes carrying the ancestral or derived alleles, respectively. SNP iHS scores were further transformed into $\pi iHS = -\log[1 - 2|\Phi(iHS) - 0.5|]$ where $\Phi(x)$ represent the Gaussian cumulative distribution function [27]. Assuming iHS are normally distributed (under neutrality), πiHS might thus be interpreted as $\log_{10}(1/P)$ where P is the two-sided P -value associated to the neutral hypothesis (no selection). Negative XP-EHH scores suggest selection happened in the reference population, otherwise happened in the observed population. The XP-EHH is transformed into $\pi XP-EHH = -\log[1 - 2|\Phi(XP-EHH) - 0.5|]$ where $\Phi(x)$ represents the Gaussian cumulative distribution function [27]. Assuming XP-EHH are normally distributed (under neutrality), $\pi XP-EHH$ may be interpreted as a two-sided P -value in a $-\log_{10}$ scale. In this study, sites within the top 0.1% of the F_{ST} detection results and $P < 0.001$ (0.1%) of the iHS and XP-EHH detection results were defined as selected or differential sites [28–31].

Gene annotation and functional enrichment

GALLO (v1.1) [32] was used for gene annotation within 200 kb upstream and downstream of selected sites. To gain a better understanding of the biological processes (BP) shared by these candidate genes, we also performed functional enrichment analysis using DAVID [33] and the ‘clusterProfiler’ package [34].

Results

Genetic diversity and linkage disequilibrium

In this study, the genetic diversity indicators were calculated for 10 populations and subpopulations, and the results are shown in Table 2. Compared with those of other breeds, the Holstein cattle populations exhibited higher heterozygosity and MAF, with heterozygosity consistently above 0.4. Jersey cattle showed the highest level of inbreeding, followed by plains Holstein cattle. GH100 and GJ showed large differences in indicators of genetic diversity. The heterozygosity, MAF and F_{ROH} of GJ were closer to those of PJ. As the proportion of Holstein ancestry decreased, the MAF increased, the F_{ROH} decreased

Table 2 Statistics for the studied populations and subpopulations calculated on the basis of the SNP genotypes

Populaiton	Number	HET	MAF	F_{ROH}	N_{E13}
GH100	129	0.4134	0.3258	0.0629	157
GH85	92	0.4299	0.3353	0.0305	194
GH75	92	0.4319	0.3426	0.0231	260
GH40	119	0.4290	0.3451	0.0162	399
GJ	28	0.3720	0.2798	0.1140	79
PH	56	0.4036	0.3075	0.0921	91
PJ	77	0.3694	0.2779	0.1190	96
XZ	61	0.3533	0.2735	0.0486	191
AP	39	0.3461	0.2496	0.0353	106
TF	59	0.2974	0.2151	0.0817	111

HET heterozygosity, MAF minor allele frequency, F_{ROH} Inbreeding coefficient based on runs of homozygosity, N_{E13} effective population size 13 generations ago, GH100 is Holstein cattle with 94%~100% Holstein bloodlines, GH85 is dairy cattle with 85%~93% Holstein bloodlines, GH75 is dairy cattle with 75%~84% Holstein bloodlines, GH40 is crossbred dairy cattle with 40%~74% Holstein bloodlines, GJ is Jersey cattle with 76%~100% Jersey bloodlines, PH is Holstein cattle in Beijing, PJ is Jersey cattle in Hebei, XZ is Tibet cattle, AP is Apeijia cattle, TF is Shigatse hump cattle

and the N_{E13} increased. GH40 and GH75 showed higher genetic diversity than plateau native cattle. Among the plateau native cattle, TF had the lowest heterozygosity, MAF, and the highest F_{ROH} . XZ had a larger N_{E13} compared to AP and TF.

Figure 1 shows the LD decay diagram within 500 kb of each population and subpopulations. The LD levels of Jersey cattle in the plains and plateaus were higher than that of other populations and subpopulations. The LD level of PH is higher than that of native cattle breeds in Tibet and Holstein cattle with different bloodline purities in Tibet. The LD decay rates of XZ and GH40 were similar and the fastest. The LD decay rate of different bloodline purities Holstein cattle subpopulations in Tibet increased as the proportion of Holstein ancestry decreased.

Population genetic structure

The MDS analysis based on the IBS distance is presented in Fig. 2(a). The first dimension separates Tibetan native cattle breeds from other Holstein and Jersey cattle populations, with XZ being genetically closest to other Holstein cattle populations, while the TF exhibits the farthest genetic distance from other Holstein cattle populations. Holstein cattle of different bloodline purities in Tibet are distributed along the first dimension according to the proportion of Holstein ancestry, with GH100 having the closest genetic distance to PH and GH40 having the closest genetic distance to the native Tibetan cattle breeds. The second dimension separates Jersey cattle, while PJ and GJ do not exhibit distinct clustering. Five individuals

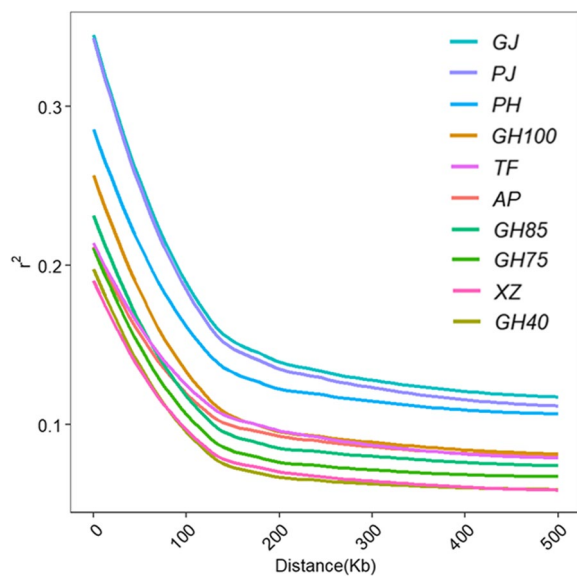


Fig. 1 LD decay plot for each population and subpopulation. GH100 is Holstein cattle with 94%~100% Holstein bloodlines; GH85 is dairy cattle with 85%~93% Holstein bloodlines; GH75 is dairy cattle with 75%~84% Holstein bloodlines; GH40 is crossbred dairy cattle with 40%~74% Holstein bloodlines; GJ is Jersey cattle with 76%~100% Jersey bloodlines; PH is Holstein cattle in Beijing; PJ is Jersey cattle in Hebei; XZ is Tibet cattle; AP is Apeijia cattle; TF is Shigatse hump cattle

from the GH40 subpopulation are positioned between the Holstein and Jersey cattle populations.

The neighbor-joining phylogenetic tree constructed based on genetic distance (Fig. 2(b)) divides 752 cattle into three distinct branches. PJ and GJ clustered in the

second branch. PH, most of the GH100 strains, a few of the GH85 strains and a few of the GH75 strains clustered in the first branch. Other cattle were distributed in the third branch, with TF being more distant from the other populations, and XZ and AP clustered together and closer to TF. Most of the GH40s were closer to Tibetan native cattle, and GH85 and GH75 did not show significant clustering.

In the population structure analysis (Fig. 3), three native Tibetan cattle breeds were separated from Holstein and Jersey cattle at $k=2$, and Jersey cattle were clearly demarcated at $k=3$. TF cattle were clearly demarcated at $k=4$, and Holstein cattle of different bloodline purities in Tibet were admixed with more XZ and AP bloodlines as the proportion of Holstein bloodlines decreased. Ancestry of Holstein cattle was added at $k=5$, and at $k=6$ TF cattle appeared to have ancestry different from AP and XZ. XZ and AP showed a clear difference in ancestry by $k=7$. GJ and PJ both had good consistency in terms of pedigree composition from $k=2$ to 7. The CV error continued to decrease from $k=1$ to $k=20$ (Additional file 1: Table S1).

Selection signatures based on dairy cattle in plains and plateaus

The genome-wide distribution of $|iHS|$ values in the GH100 is shown in Fig. 4(a). The Manhattan plots of pairwise XP-EHH and F_{ST} analysis between GH100 and PH are presented in Fig. 4(b-c). The mean F_{ST} value of GH100 in comparison to PH was 0.0123. A total of 79 significant SNPs ($p < 0.001$) distributed across 23 chromosomes were identified in iHS analysis (Additional

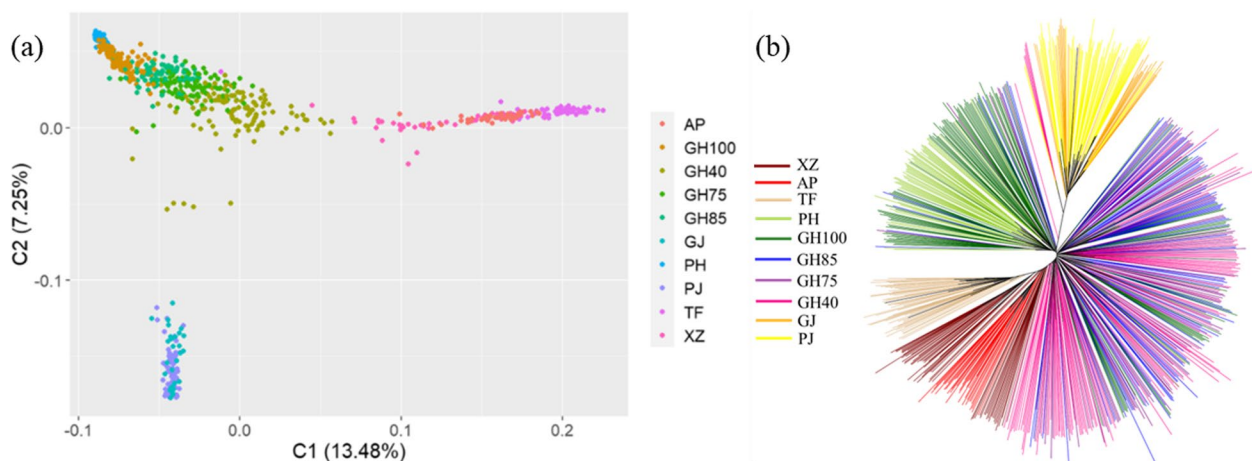


Fig. 2 Population structure, as revealed by Multidimensional scaling analysis (a) and Neighborhood phylogenetic tree (b). GH100 is Holstein cattle with 94%~100% Holstein bloodlines; GH85 is dairy cattle with 85%~93% Holstein bloodlines; GH75 is dairy cattle with 75%~84% Holstein bloodlines; GH40 is crossbred dairy cattle with 40%~74% Holstein bloodlines; GJ is Jersey cattle with 76%~100% Jersey bloodlines; PH is Holstein cattle in Beijing; PJ is Jersey cattle in Hebei; XZ is Tibet cattle; AP is Apeijia cattle; TF is Shigatse hump cattle

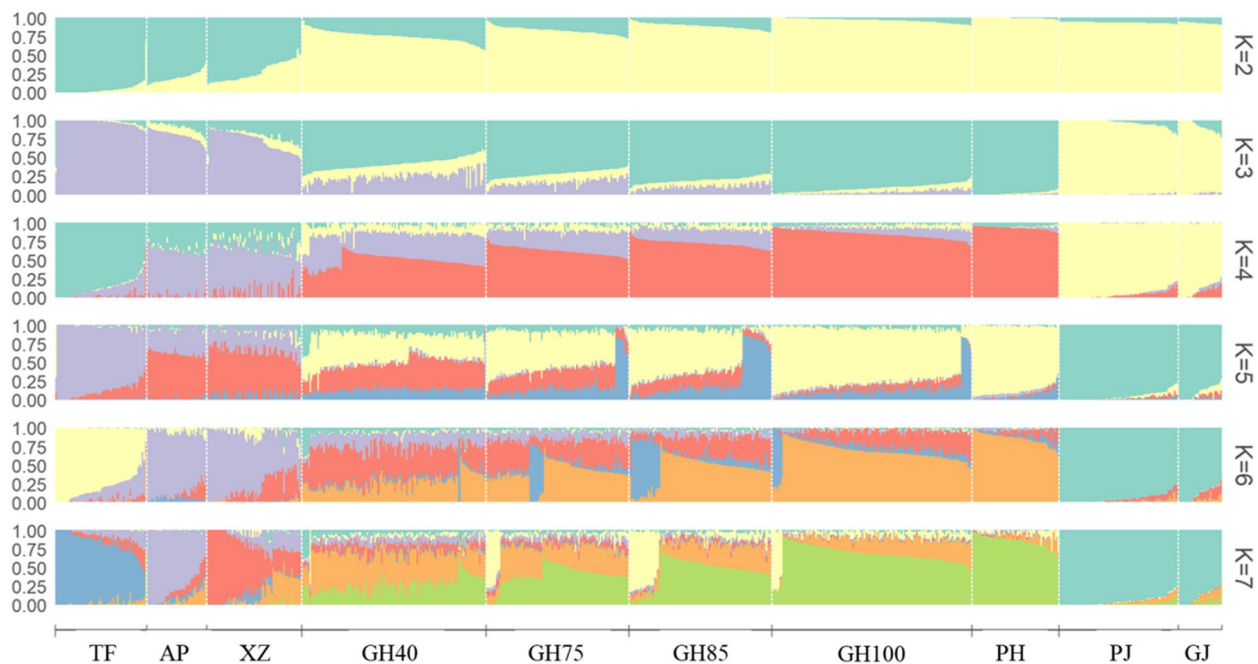


Fig. 3 Admixture plot represents the cluster structure of the studied populations and subpopulations for the number of clusters (K) 2 to 7. GH100 is Holstein cattle with 94%~100% Holstein bloodlines; GH85 is dairy cattle with 85%~93% Holstein bloodlines; GH75 is dairy cattle with 75%~84% Holstein bloodlines; GH40 is crossbred dairy cattle with 40%~74% Holstein bloodlines; GJ is Jersey cattle with 76%~100% Jersey bloodlines; PH is Holstein cattle in Beijing; PJ is Jersey cattle in Hebei; XZ is Tibet cattle; AP is Apeijia cattle; TF is Shigatse hump cattle

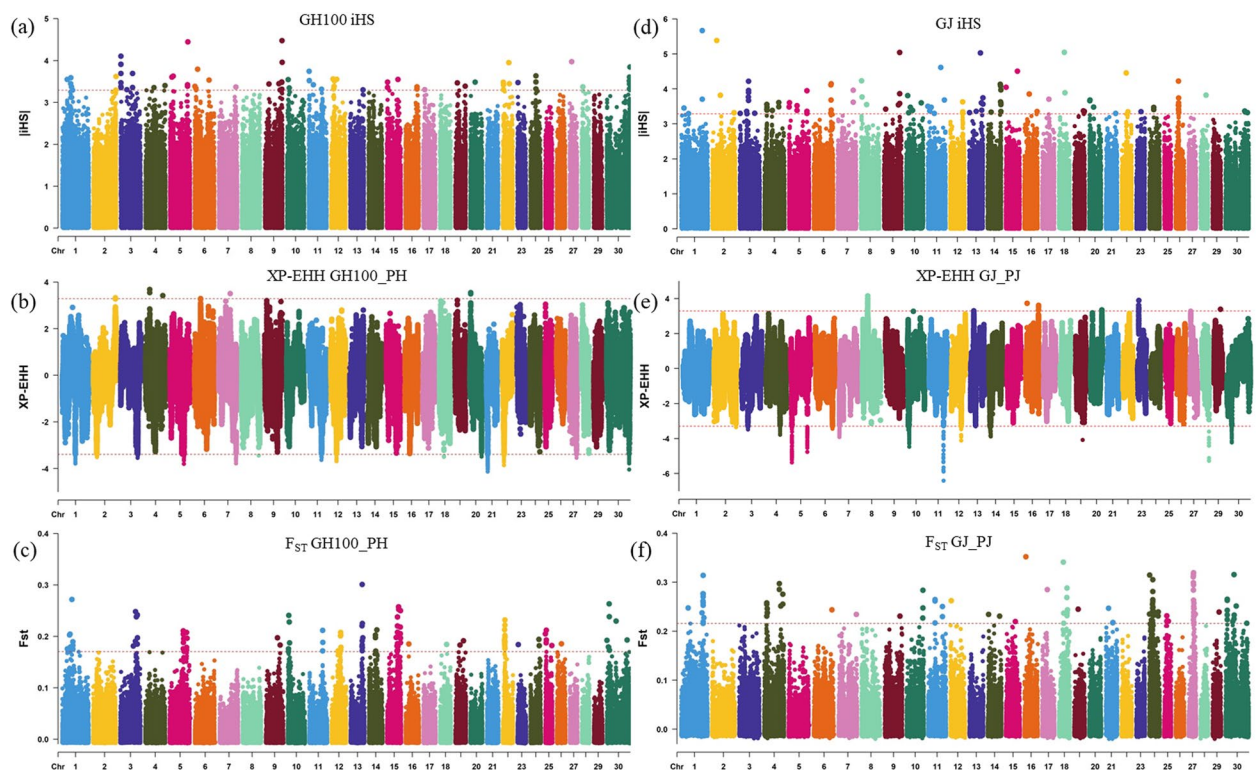


Fig. 4 **a** iHS plot for the GH100. **b** XP-EHH plot for GH100 comparison with PH. **c** F_{ST} plot for GH100 comparison with PH. **d** iHS plot for the GJ. **e** XP-EHH plot for GJ comparison with PJ. **f** F_{ST} plot for GJ comparison with PJ. GH100 is Holstein cattle in Tibet; PH is Holstein cattle in Beijing; GJ is Jersey cattle in Tibet; PJ is Jersey cattle in Hebei

file 2: Table S2). BTA3 and BTA9 were the chromosomes harboring the largest number of significant |iHS| markers (10 and 8, respectively). The XP-EHH selection signals were detected on 20 chromosomes (Additional file 2: Table S3), and positive XP-EHH scores (indicating a selection on the GH100 genome) are observed on BTA2, BTA4, BTA6, BTA7 and BTA20. F_{ST} signals were identified on 19 chromosomes (Additional file 2: Table S4). The regions on the chromosomes having highest F_{ST} values were located on BTA13 (0.30), BTA1 (0.27), BTA30 (0.26) and BTA15 (0.26), respectively.

The genome-wide distribution of |iHS| values in the GJ is shown in Fig. 4(d). The Manhattan plots of pairwise XP-EHH and F_{ST} analysis between GJ and PJ are presented in Fig. 4(e–f). The mean F_{ST} value of GJ in comparison to PJ was 0.0117. A total of 107 significant SNPs ($p < 0.001$) distributed across all chromosomes except BTA25, BTA27 and BTA29 were identified in iHS analysis (Additional file 2: Table S5). The XP-EHH selection signals were detected on 19 chromosomes (Additional file 2: Table S6), and positive XP-EHH scores (indicating a selection on the GJ genome) are observed on BTA8, BTA13, BTA16, BTA20, BTA23 and BTA29. BTA5 and BTA7 were the chromosomes harboring the largest number of these SNPs (83 and 61, respectively), followed by BTA10 (36 SNPs). F_{ST} signals were identified on 20 chromosomes (Additional file 2: Table S7). The regions on the chromosomes having highest F_{ST} values were located on BTA16 (0.35), BTA18 (0.34), and BTA27 (0.32), respectively.

The top 10 SNPs with their iHS values or positive XP-EHH values and annotated genes are shown in Table 3. The enriched biological processes were shown in Additional file 3: Table S8–S9. Among the top 10 |iHS| markers in GH100, 3 are located on BTA3, in which annotated *STYXL2*. *STYXL2* (BTA3, iHS=4.10), *ALPL* (BTA2, XP-EHH=3.32) and *PPMIK* (BTA6, XP-EHH=3.29) enriched the dephosphorylation BP (GO:0016311, $p < 0.05$). Among the top 10 positive XP-EHH markers in GJ, 8 are located on BTA8, in which annotated *TYRP1*. *TYRP1* (BTA8, XP-EHH=4.16) was contained in the cellular pigmentation BP (GO:0033059, $p < 0.05$). *GRM4* (BTA23, XP-EHH=3.89) was contained in the G protein-coupled glutamate receptor signaling pathway BP (GO:0007216, $p < 0.05$). The strongest iHS signal (5.67) in GJ was found on BTA1 (BovineHD0100035549), in which annotated *DIPK2A*. *DIPK2A* (BTA1, iHS=5.67) was contained in the cardiac muscle tissue development BP (GO:0048738, $p < 0.05$). *BPI* (BTA13, iHS=5.03) and *LBP* (BTA13, iHS=5.03) enriched the regulation of macrophage activation (GO:0043030, $p < 0.01$).

A total of 29 genes on BTA3, BTA5, BTA11, BTA22 and BTA23 were determined by more than one approach in

GH100 in comparison to PH (Table 4). *SCN5A*, *SCN10A* and *SCN11A* on BTA22 enriched to the neuronal action potential BP (GO:0019228, $p < 0.01$). *SCN5A* and *SCN10A* enriched to the AV node cell action potential BP (GO:0086016, $p < 0.01$). *SCN5A* and *SCN11A* enriched to the cardiac muscle cell action potential involved in contraction BP (GO:0086002, $p < 0.05$). *RAP1GAP* (BTA2, XP-EHH=3.32), *SNX13* (BTA4, XP-EHH=3.68), *STAC* (BTA22) and *EPM2AIP1* (BTA22) enriched to the positive regulation of molecular function BP (GO:0044093, $p < 0.05$). *TTL2* (BTA9, iHS=3.96), *PRMT8* (BTA5, iHS=4.45), *HIPK2* (BTA4, XP-EHH=3.42) and *STK38L* (BTA5) enriched to the peptidyl-amino acid modification BP (GO:0018193, $p < 0.05$).

A total of 20 genes on BTA1, BTA4, BTA6, BTA10, BTA11, BTA16 and BTA29 were determined by more than one approach in GJ in comparison to PJ (Table 4). *LBP* (BTA13, iHS=5.03) and *THBS4* (BTA10) enriched to the positive regulation of neutrophil chemotaxis BP (GO:0090023, $p < 0.05$). Two genes *WDR48* (BTA22), *MAEL* (BTA3, iHS=3.91) from Holstein and three genes *ASPM* (BTA16), *APOB* (BTA11, iHS=4.61), *TDRD15* (BTA11, iHS=4.61) from Jersey were enriched for spermatogenesis BP (GO:0007283, $p < 0.05$). *MAEL* (BTA3, iHS=3.91) selected in GH100, *NCAPD2* (BTA5) and *ASPM* (BTA16) selected in GJ were enriched for meiotic cell cycle process BP (GO:1903046, $p < 0.01$).

Selection signatures based on crossbred dairy cattle in Tibet

The genome-wide distribution of |iHS| values in the GH40 is shown in Fig. 5(a). A total of 88 significant SNPs ($p < 0.001$) distributed across 20 chromosomes were identified. BTA14 was the chromosome harboring the largest number of these SNPs (33). Results of the pairwise genome-wide XP-EHH analyses between the GH40 and PH/XZ are shown in the Manhattan plots included in Fig. 5(b, d). Against PH, the XP-EHH selection signals were detected across all chromosomes except BTA19 and BTA28. Positive XP-EHH values (indicating a selection on the GH40 genome) were identified on BTA14, BTA17 and BTA24. Against XZ, the XP-EHH selection signals were detected on 23 chromosomes, and positive XP-EHH values were identified on BTA2, BTA8, BTA9, BTA14, BTA16, BTA20 and BTA30. F_{ST} values in the two between-breeds comparisons, that is GH40 v.s. PH and GH40 v.s. XZ, were 0.0451 and 0.0651, respectively. Figure 5(c, e) reports the Manhattan plots of the window-based pairwise genome-wide F_{ST} analyses of the GH40 breed against other breeds. In the GH40 v.s. PH comparison, F_{ST} signals were identified on 23 chromosomes. The SNPs with the highest F_{ST} values were on BTA25 (0.50), BTA3 (0.49) and BTA4 (0.46). In the GH40 v.s. XZ

Table 3 List of the top 10 |iHS| and positive XP-EHH for the SNP markers in GH100 and GJ, and their closest genes, with information on the Bos taurus chromosome (BTA) position

Test	SNP	BTA	Position	iHS /XP-EHH	Annotated genes
iHS (GH100)	BovineHD0900030245	9	101,767,952	4.48	<i>RPS6KA2,RNASET2,CEP43,CCR6,GPR31</i>
iHS (GH100)	BovineHD0500030656	5	106,261,952	4.45	<i>PARP11,CRACR2A,PRMT8</i>
iHS (GH100)	BovineHD0300000421	3	1,581,101	4.10	<i>POU2F1,STYXL2,GPA33</i>
iHS (GH100)	ARS-BFGL-NGS-66811	27	12,414,458	3.97	N.A
iHS (GH100)	ARS-BFGL-NGS-102044	9	102,110,374	3.96	<i>CCR6,GPR31,TLL2,UNC93A,AFDN</i>
iHS (GH100)	Hapmap35530-SCAFFOLD65737_1657	22	36,411,261	3.95	N.A
iHS (GH100)	BTA-67311-no-rs	3	1,531,782	3.91	<i>POU2F1,STYXL2,GPA33</i>
iHS (GH100)	BovineHD0300000445	3	1,645,135	3.91	<i>POU2F1,STYXL2,GPA33,MAEL</i>
iHS (GH100)	BovineHD3000044205	30	137,652,192	3.84	N.A
iHS (GH100)	ARS-BFGL-NGS-91848	6	18,195,932	3.79	<i>DKK2</i>
XP-EHH (GH100)	BovineHD0400007618	4	26,173,337	3.68	<i>SNX13</i>
XP-EHH (GH100)	ARS-BFGL-NGS-38856	20	5,397,009	3.55	<i>BOD1</i>
XP-EHH (GH100)	BTB-01383912	4	26,109,668	3.54	N.A
XP-EHH (GH100)	Hapmap31211-BTA-29377	7	68,574,155	3.51	N.A
XP-EHH (GH100)	BovineHD2000001656	20	5,392,046	3.49	<i>BOD1</i>
XP-EHH (GH100)	BovineHD0400029309	4	103,308,051	3.42	<i>HIPK2,TBXAS1,PARP12</i>
XP-EHH (GH100)	BovineHD0200038391	2	131,340,253	3.32	<i>RAP1GAP,ALPL,ECE1,EIF4G3</i>
XP-EHH (GH100)	BovineHD0600010456	6	36,395,590	3.29	<i>HERC3,PYURF,HERC5,HERC6,PPM1K,ABCG2</i>
iHS (GJ)	BovineHD0100035549	1	125,040,443	5.67	<i>DIPK2A,SLC9A9</i>
iHS (GJ)	BovineHD0200007620	2	26,291,885	5.38	<i>MYO3B,UBR3</i>
iHS (GJ)	BovineHD1800009949	18	32,766,212	5.05	<i>CDH11</i>
iHS (GJ)	BovineHD0900026973	9	93,742,851	5.04	<i>ARID1B,TMEM242</i>
iHS (GJ)	BovineHD0900026974	9	93,744,138	5.04	<i>ARID1B,TMEM242</i>
iHS (GJ)	BPI-2	13	67,172,651	5.03	<i>TGM2,KIAA1755,BPI,LBP</i>
iHS (GJ)	BovineHD1100022373	11	77,869,084	4.61	<i>TDRD15,APOB</i>
iHS (GJ)	BovineHD1500020381	15	69,760,189	4.50	N.A
iHS (GJ)	BovineHD2200007996	22	27,616,751	4.46	<i>CNTN3</i>
iHS (GJ)	ARS-BFGL-NGS-68739	8	5,147,600	4.23	N.A
XP-EHH (GJ)	BovineHD0800009655	8	31,812,018	4.16	<i>TYRP1</i>
XP-EHH (GJ)	Hapmap54720-rs29023017	8	31,818,769	4.02	<i>TYRP1</i>
XP-EHH (GJ)	BovineHD0800009650	8	31,796,418	3.99	<i>TYRP1</i>
XP-EHH (GJ)	ARS-BFGL-NGS-23687	23	8,154,656	3.89	<i>GRM4,HMGA1,SMIM29,NUDT3</i>
XP-EHH (GJ)	BTA-123382-no-rs	8	31,779,503	3.87	<i>TYRP1</i>
XP-EHH (GJ)	BovineHD0800009630	8	31,751,207	3.76	<i>LURAP1L,TYRP1</i>
XP-EHH (GJ)	BovineHD1600002717	16	9,494,040	3.73	N.A
XP-EHH (GJ)	BTA-81018-no-rs	8	31,845,307	3.72	<i>TYRP1</i>
XP-EHH (GJ)	BovineHD0800009641	8	31,850,020	3.68	N.A
XP-EHH (GJ)	Hapmap43060-BTA-81019	8	31,867,242	3.65	N.A

comparison, F_{ST} signals were identified across all chromosomes except BTA21, BTA25 and BTA27. The SNPs with the highest F_{ST} values were on BTA4 (0.59), BTA11 (0.54) and BTA30 (0.53).

The top 10 SNPs with their iHS values or positive XP-EHH values and annotated genes are shown in Table 5. The enriched biological processes were shown in Additional file 3: Table S10 – S11. The highest iHS value

on autosomes was on BTA20, annotated to *NKX2-5*. *NKX2-5* (BTA20, iHS=4.47) and *SMAD7* (BTA24, XP-EHH_{PH}=3.32) enriched the ventricular septum morphogenesis BP (GO:0060412, $p < 0.05$). Among the top 10 positive XP-EHH values in the GH40 and XZ comparisons, 5 are located on BTA2. *PINK1* (BTA2, XP-EHH_{XZ}=4.00), *HP1BP3* (BTA2, XP-EHH_{XZ}=4.00) and *MTOR* (BTA16, XP-EHH_{XZ}=3.60) were all shown to be

Table 4 Genes identified by more than one test in plateau dairy cattle

Breeds	Tests	BTA	Annotated genes
Holstein	F _{ST} -XP-EHH	3	<i>ERI3</i>
Holstein	F _{ST} -XP-EHH	5	<i>PPFIBP1,C5H12orf71,ARNTL2,STK38L,MED21,TM7SF3</i>
Holstein	F _{ST} -iHS	5	<i>NCAPD2,MRPL51,VAMP1,TAPBPL,CD27,VWF</i>
Holstein	F _{ST} -XP-EHH	11	<i>OSR1</i>
Holstein	F _{ST} -XP-EHH	22	<i>STAC,LRRFIP2,MLH1,EPM2AIP1,TRANK1,DCLK3,GOLGA4,SCN5A,SCN10A,SCN11A,WDR48,GORASP1,TTC21A,CSRNP1</i>
Holstein	F _{ST} -iHS	23	<i>HCRT2</i>
Jersey	F _{ST} -iHS	1	<i>SLC9A9</i>
Jersey	F _{ST} -iHS	4	<i>TAC1,ASNS,COL28A1</i>
Jersey	XP-EHH;iHS	6	<i>HSD17B11,NUDT9,SPARCL1</i>
Jersey	XP-EHH;iHS	10	<i>CMYA5,MTX3,THBS4,SERINC5</i>
Jersey	F _{ST} -XP-EHH	11	N.A
Jersey	XP-EHH;iHS	16	<i>F13B,CFHR5,ASPM,ZBTB41</i>
Jersey	F _{ST} -XP-EHH	29	<i>PAG4,PAG14,PAG16,PAG20,PAG21</i>

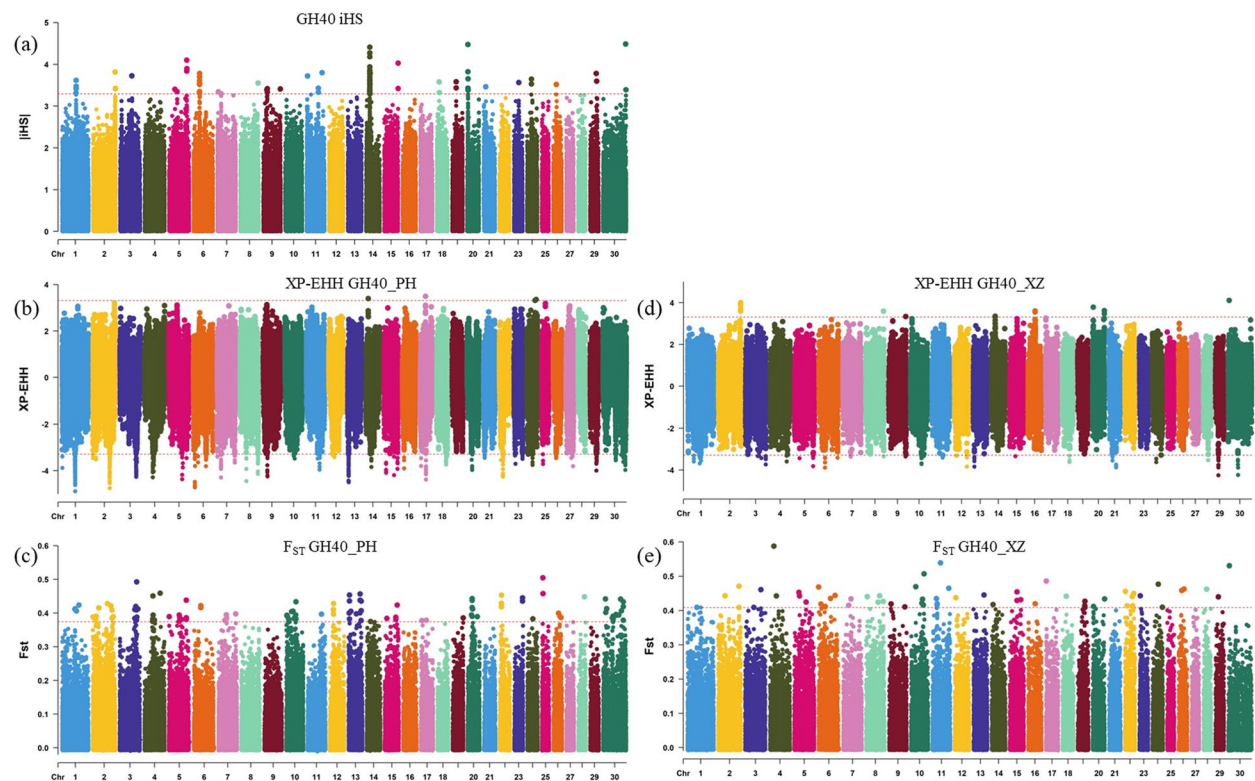


Fig. 5 **a** iHS plot for the GH40. **b** XP-EHH plot for GH40 comparison with PH. **c** F_{ST} plot for GH40 comparison with PH. **d** XP-EHH plot for GH40 comparison with XZ. **e** F_{ST} plot for GH40 comparison with XZ. GH40 is crossbred dairy cattle in Tibet; PH is Holstein cattle in Beijing; XZ is Tibet cattle

selected in GH40 compared to XZ, enriched for cellular response to hypoxia BP (GO:0071456, $p=0.01$). *TERT* (BTA20, XP-EHH_{XZ}=3.62) enriched the RNA-templated DNA biosynthetic process BP (GO:0006278, $p<0.05$).

A total of 52 genes and 35 genes were determined by more than one approach in GH40 in comparison to PH and XZ, respectively (Table 6). In comparison with PH, *RAB27B* (BTA24, XP-EHH_{PH}=3.35) and *OCA2* (BTA2) enriched to the pigmentation BP (GO:0043473, $p<0.01$).

Table 5 List of the top 10 |iHS| and positive XP-EHH for the SNP markers in GH40 and their closest genes, with information on the Bos taurus chromosome (BTA) position

Test	SNP	BTA	Position	iHS /XP-EHH	Annotated genes
iHS (GH40)	Hapmap53788-rs29014096	30	135,508,916	4.49	N.A
iHS (GH40)	BTB-01546179	20	5,027,370	4.47	CREBRF, BNIP1, NKX2-5, STC2
iHS (GH40)	BovineHD1400007102	14	22,813,366	4.41	XKR4
iHS (GH40)	BTB-01532239	14	22,781,305	4.27	XKR4
iHS (GH40)	BovineHD1400007101	14	22,809,167	4.18	XKR4
iHS (GH40)	BovineHD0500030627	5	106,200,983	4.10	PARP11, CRACR2A, PRMT8
iHS (GH40)	BTA-37923-no-rs	15	83,673,161	4.03	NCAPD3, VPS26B, THYN1, ACAD8, GLB1L3, GLB1L2, B3GAT1
iHS (GH40)	BovineHD1400007103	14	22,814,269	3.94	XKR4
iHS (GH40)	BovineHD0500030663	5	106,291,238	3.90	PARP11, CRACR2A, PRMT8, TSPAN11
iHS (GH40)	BovineHD1400007253	14	23,321,577	3.86	LYN, RPS20, MOS, PLAG1, CHCHD7, SDR16C5, SDR16C6
XP-EHH (GH40v.s.PH)	BTA-102389-no-rs	17	31,170,793	3.49	N.A
XP-EHH (GH40v.s.PH)	BTB-01110396	14	14,314,815	3.40	N.A
XP-EHH (GH40v.s.PH)	ARS-BFGL-NGS-6396	24	54,432,213	3.35	RAB27B, CCDC68, TCF4
XP-EHH (GH40v.s.PH)	BovineHD2400013686	24	48,501,459	3.32	CTIF, SMAD7
XP-EHH (GH40v.s.XZ)	ARS-USDA-AGIL-chrX-4347630-000803	30	5,540,748	4.11	N.A
XP-EHH (GH40v.s.XZ)	ARS-BFGL-NGS-118497	2	131,944,866	4.00	EIF4G3, HP1BP3, SH2D5, KIF17, DDOST, PINK1, CDA, FAM43B, MUL1, CAMK2N1
XP-EHH (GH40v.s.XZ)	BovineHD0200038536	2	131,931,144	3.95	EIF4G3, HP1BP3, SH2D5, KIF17, DDOST, PINK1, CDA, FAM43B, MUL1
XP-EHH (GH40v.s.XZ)	BovineHD0200038542	2	131,952,913	3.87	EIF4G3, HP1BP3, SH2D5, KIF17, DDOST, PINK1, CDA, FAM43B, MUL1, CAMK2N1
XP-EHH (GH40v.s.XZ)	MS-rs208701048	20	4,708,873	3.79	DUSP1, ERGIC1, RPL26L1, ATP6V0E1, CREBRF, BNIP1
XP-EHH (GH40v.s.XZ)	ARS-BFGL-NGS-60038	2	131,987,746	3.68	EIF4G3, HP1BP3, SH2D5, KIF17, DDOST, PINK1, CDA, FAM43B, MUL1, CAMK2N1
XP-EHH (GH40v.s.XZ)	ARS-BFGL-NGS-95318	20	70,975,538	3.62	IRX4, NDUFS6, MRPL36, LPCAT1, SLC6A3, CLPTM1L, TERT, SLC6A18
XP-EHH (GH40v.s.XZ)	BovineHD1600012100	16	42,390,809	3.60	DISP3, UBIAD1, MTOR, ANGPTL7, EXOSC10, SRM, MASP2
XP-EHH (GH40v.s.XZ)	BovineHD0200038520	2	131,850,464	3.60	EIF4G3, HP1BP3, SH2D5, KIF17, DDOST, PINK1, CDA
XP-EHH (GH40v.s.XZ)	ARS-BFGL-NGS-20859	8	106,212,293	3.59	ASTN2

TMEM109 (BTA29) enriched to the cellular response to radiation BP (GO:0071478, $p < 0.05$). *TOE1* (BTA3) and *CPSF7* (BTA29) enriched to the RNA 3'-end processing BP (GO:0031123, $p < 0.05$). *P4HA2* (BTA7) was contained in the peptidyl-proline hydroxylation BP (GO:0019511, $p < 0.05$). *BTBD3* (BTA13), *LAMA3* (BTA24) and *ZFYVE27* (BTA26) enriched to the cell morphogenesis involved in neuron differentiation BP (GO:0048667, $p < 0.05$). *NRCAM* (BTA4) and *BTBD3* (BTA13) are genes that were selected in PH and enriched for brain development BP (GO:0007420, $p < 0.05$). *MOS* (BTA14, iHS = 3.86), *LYN* (BTA14, iHS = 3.86) and *FGF10* (BTA20) enriched for positive regulation of MAPK cascade BP (GO:0043410, $p < 0.05$). *SMAD7* (BTA24, XP-EHH_{PH} = 3.32), *FGF10* (BTA20) and *ZFYVE27* (BTA26) enriched for response to growth factor (GO:0070848, $p < 0.05$). *FGF10* (BTA20) and *SFRP5* (BTA26) enriched

for regulation of canonical Wnt signaling pathway (GO:0060828, $p < 0.05$).

In comparison with XZ, *CREBRF* (BTA20) was detected to be selected in GH40 by 3 methods, contained in response to unfolded protein BP (GO:0006986). *NCAPD3* (BTA15, iHS = 4.03) and *NCAPG* (BTA6) enriched to chromosome condensation BP (GO:0030261, $p < 0.01$). *IRX4* (BTA20, XP-EHH_{XZ} = 3.62), *ASTN2* (BTA8, XP-EHH_{XZ} = 3.59), *TENM4* (BTA29), *ZFYVE27* (BTA26) and *CRTAC1* (BTA26) enriched to generation of neurons (GO:0048699, $p < 0.01$).

XKR4 (BTA14, iHS = 4.31) was detected in selection signals in both comparison groups and was annotated by 27 SNPs under selection in the iHS analysis, contained in plasma membrane phospholipid scrambling (GO:0017121, $p < 0.05$). *ASTN2* (BTA8) was detected in XP-EHH in both comparison groups, differing in that selection occurred in GH40 compared to XZ

Table 6 Genes identified by more than one test in crossbred dairy cattle in Tibet

Breeds	Tests	BTA	Annotated genes
GH40v.s.PH	XP-EHH; F_{ST}	1	<i>IL1RAP,TMEM207,TPRG1</i>
GH40v.s.PH	XP-EHH; F_{ST}	2	<i>OCA2,HERC2</i>
GH40v.s.PH	XP-EHH; F_{ST}	3	<i>TESK2,TOE1,MUTYH,HPDL,ZSWIM5</i>
GH40v.s.PH	XP-EHH; F_{ST}	4	<i>COG5,GPR22,NRCAM</i>
GH40v.s.PH	XP-EHH; iHS	7	<i>P4HA2</i>
GH40v.s.PH	XP-EHH; iHS	8	<i>ASTN2</i>
GH40v.s.PH	XP-EHH; iHS	9	<i>CENPW</i>
GH40v.s.PH	XP-EHH; F_{ST}	12	<i>COG6,LHFPL6</i>
GH40v.s.PH	XP-EHH; F_{ST}	13	<i>BTBD3</i>
GH40v.s.PH	F_{ST} ; iHS	14	<i>XKR4,TMEM68,TGS1,LYN</i>
GH40v.s.PH	XP-EHH; F_{ST}	20	<i>FGF10</i>
GH40v.s.PH	XP-EHH; F_{ST}	22	<i>STAC</i>
GH40v.s.PH	XP-EHH; F_{ST}	24	<i>LAMA3,ANKRD29,NPC1,RMC1</i>
GH40v.s.PH	XP-EHH; iHS	26	<i>ZFYVE27,SFRP5</i>
GH40v.s.PH	XP-EHH; iHS	29	<i>MS4A10,CCDC86,PTGDR2,PRPF19,TMEM109,TMEM132A,SLC15A3,CD6,PAG8,P GA5,VWCE,DDB1,TKFC,CYB561A3,TMEM138,TMEM216,CPSF7,SDHAF2,PPP1R32 ,LRRC10B,SYT7</i>
GH40v.s.XZ	XP-EHH; iHS	2	<i>EIF4G3,HP1BP3,SH2D5,KIF17,DDOST,PINK1,CDA,FAM43B,MUL1,CAMK2N1</i>
GH40v.s.XZ	XP-EHH; F_{ST}	3	<i>EPS15</i>
GH40v.s.XZ	XP-EHH; iHS	6	<i>LAP3,MED28,FAM184B,NCAPG,DCAF16</i>
GH40v.s.XZ	XP-EHH; F_{ST}	7	<i>CATSPER3,PITX1</i>
GH40v.s.XZ	XP-EHH; iHS	8	<i>ASTN2</i>
GH40v.s.XZ	XP-EHH; iHS	9	<i>CCR6,GPR31,TTL2,UNC93A</i>
GH40v.s.XZ	XP-EHH; iHS	14	<i>XKR4</i>
GH40v.s.XZ	XP-EHH; F_{ST}	20	<i>DUSP1,ERGIC1,RPL26L1,ATP6V0E1</i>
GH40v.s.XZ	XP-EHH; F_{ST} ; iHS	20	<i>CREBRF</i>
GH40v.s.XZ	XP-EHH; iHS	20	<i>BNIP1</i>
GH40v.s.XZ	F_{ST} ; iHS	26	<i>ZFYVE27,SFRP5,GOLGA7B,CRTAC1</i>
GH40v.s.XZ	XP-EHH; F_{ST}	29	<i>TENM4</i>

(GH40v.s.xz) and in PH compared to PH (GH40v.s.PH). *ZFYVE27* and *SFRP5* on BTA26 were detected in the selection signals of both comparison groups. The generation of neurons (GO:0048699) containing *ASTN2* and *ZFYVE27* was significantly enriched in both comparison groups.

Discussion

Tibet has not historically had a specialized milk-producing dairy herd. Therefore, the dairy cattle in the present intensive farms are partly introduced from the plains and partly the offspring of grading up of native cattle. Based on genetic diversity analysis and LD decay, dairy cattle with different Holstein bloodline purity in Tibet exhibited low inbreeding levels, high levels of population polymorphism, while the LD levels were similar to those of native cattle breeds in Tibet. After the Holstein cattle were grouped by bloodline, significant differences were observed in the LD levels in

each subpopulation. Studies have shown that artificial selection promotes the increase in LD levels within the population [35, 36]. GJ showed similar LD decay as PH and PJ that have undergone a long period of artificial selection and mating. GH40 and XZ showed a faster LD decay relative to the other populations.

Through the MDS we found that GH40 and XZ were more genetically connected compared to AP and TF. Therefore, we chose XZ as a representative of local breeds to participate in the analysis of selection signatures. Combining MDS, the NJ phylogenetic tree, and Admixture population structure analysis, GH85, GH75, and GH40 exhibited clear signs of historical hybridization, whereas GH100 and PH were not completely separate. This suggested lower genomic consistency among different-purity Holstein cattle subpopulations in Tibet. Consequently, in subsequent research, these populations should not be treated as a single group. In

contrast, GJ, which has not undergone hybridization, exhibited complete genomic admixture with PJ.

It is through selection and mating systems that people in Tibet are trying to solve the problem of insufficient milk production in dairy cattle living on the plateau. This includes the use of semen from dairy cattle to mate with native cattle to improve milk production performance. At the same time, some purebred dairy cattle that could not adapt to the plateau environment were culled due to disease or death. According to the MDS, the first dimension component separated the native Tibetan breeds from the other populations, indicating that the high-altitude environment led to distinctive genomic characteristics in native Tibetan cattle. In contrast to commercial high-bred dairy cattle, Tibetan dairy cattle retain greater genomic variability, which may be crucial for addressing current and future challenges related to environmental adaptability and meeting breeding goals.

In Admixture GH100 appears in the same areas as native cattle in the Plateau. Due to the lack of pedigree, we were unable to determine the mating history of GH100 and the origin of these regions. The size of the overlap region between the first Holstein cattle introduced to the plateau and the plateau native cattle is unknown. We were also unable to determine what percentage of these regions were recognized as Holstein bloodlines in the calculations due to the absence of BBR values for plateau native cattle.

When characterizing the genetic background of strategies for surviving at high altitude, one has to differentiate between the evolutionary genetic adaptations of native highland species and populations, e.g. yaks and blue sheep, and the acclimatization of lowland species brought to high altitude short term, for which phenotypic plasticity is the decisive factor [37, 38]. Strategies and mechanisms of both acclimatization and adaptation have a genetic background and are thus linked to each other. In this study, by comparing dairy cattle living at different altitudes, we are able to access the genetic differences caused by short-term environmental changes over several generations. By comparing crossbred cattle with Holstein and native cattle, it is possible to identify genes that differed between breeds in crossbred cattle after crossbreeding. The F_{ST} method is best suited for detection in events occurring in the more distant past [39] whereas the iHS test is best suited for detection in recent positive selection [40]. The XP-EHH is useful for detection of entirely or approximately fixed loci [41]. As the F_{ST} , iHS, and XP-EHH approaches are complementary, incorporating such measures in the study of genetic differentiation and selection would strongly contribute to understanding gene flow and genetic makeup of plateau dairy cattle (purebred and hybrid) populations.

Our study did not annotate the same genes near selection signals in Holstein and Jersey cattle, but the functions associated with these genes are similar. Various within- and between-species analyses of altitude adaptation have also shown that in most cases different genetic architectures and physiological mechanisms underlie these adaptations [38]. Chronic hypoxia with severely reduced oxygen supply is a major stressor in high-altitude conditions and limits the respiratory and cardiovascular systems from functioning properly [42]. Glover and Newsom first reported in 1915 that native cattle in Colorado, USA, at elevations higher than 1524 m, suffered from altitude sickness, which was characterized by right heart failure and chest edema due to pulmonary hypertension [43]. In this study, we have identified several heart-related genes in GH100. Genetic variants of *SCN5A* are involved in a number of inherited cardiac channelopathies and are also found to be correlated with myocardial contractile dysfunction, dilated cardiomyopathy, and heart failure [44]. Variants in *SCN10A* are associated with alterations of cardiac conduction parameters and the cardiac rhythm disorder Brugada syndrome [45]. Rap1GAP plays a critical role in mediating Ang II-induced cardiomyocyte hypertrophy through its regulation on autophagy and oxidative stress [46]. Rap1GAP was verified to regulate the AMPK SIRT1/NF- κ B signaling pathway and exacerbate the damage to myocardial cells caused by ischemia and hypoxia [47]. *SNX13* deficiency promotes cardiomyocyte apoptosis and heart failure [48]. Kun Lian found that the cardiac function of diabetes mice overexpressing *PPM1k* was improved, and the myocardial infarction area and apoptosis were reduced [49]. Exercise-induced HIPK2 suppression attenuates cardiomyocytes apoptosis [50]. The serine/threonine kinase HIPK2 as novel regulatory protein participating in hypoxic gene regulation can affect apical as well as downstream events during the hypoxic response [51]. *THBS4* and *DIPK2A* were identified in the selection signatures of GJ in this study. Studies in *THBS4*-knockout mice revealed that deletion of *THBS4* resulted in a significantly higher heart weight to body weight ratio, severe excessive cardiac fibrosis, and aggravated cardiac dysfunction than those in the WT mice in response to pressure overload [52, 53]. *DIPK2A* protects cardiomyocytes from ischemia-induced cell death by activating PRKCE (protein kinase C epsilon) [54]. Many genes associated with cardiac development are subject to selection may differ in the ability of GH40 to cope with hypoxia compared to PH and XZ. *NKX2-5* is involved in cardiac morphogenesis, including rightward looping and subsequent chamber specification and septation and in functional maturation and maintenance of working myocardium and conduction system [55]. *SMAD7* is required

for cardiac development in embryonic stage and cardiac function in adult [56]. SFRP5 diminishes cardiac inflammation and protects the heart from ischemia/reperfusion injury [57]. Serum SFRP5 levels were significantly associated with coronary artery disease in humans [58]. *HP1BP3* was up-regulated to initiate a cardioprotective stress-handling response [59]. *IRX4* may play a critical role in establishing chamber-specific gene expression in the developing heart [60].

In a previous study, angiogenesis is one of the key mechanisms of high-altitude acclimatization [61]. Hypoxia-stimulated vasculogenesis and angiogenesis may increase microcirculatory blood flow and capillary density, thus restoring the oxygen supply in tissues exposed to hypoxia to compensate for oxygen insufficiency [62]. *ALPL* and *APOB* were identified in the selection signatures of Gh100 and GJ in this study, respectively. *ALPL* might be as a distinct regulator to control exosomes secretion and subsequently regulate the pro-angiogenic potential of mesenchymal stem cells [63]. Apolipoprotein B levels can be a predictor of ischemic cardiovascular disease, and mutations and polymorphisms in *APOB* are associated with plasma apolipoprotein B levels, and risk of ischemic cardiovascular disease [64]. *APOB* and *TDRD15* found to be selected in Welsh sheep breeds living on upland pasture [65]. Among genes selected in GH40, *LYN* has well-established functions in most haematopoietic cells, viz. progenitors via influencing c-kit signaling, through to mature cell receptor/integrin signaling, e.g. erythrocytes, platelets, mast cells and macrophages [66]. *ZFYVE27* was recently found to promote endothelial cell migration and angiogenesis in humans, with its knockdown affecting these processes [67]. *P53* activates *P4HA2* expression exerting anti-angiogenic effects [68]. Activation of mTOR potentially enhances the activity of HIF1 α and vascular endothelial growth factor (VEGF)-A secretion during hypoxia [69]. *TERT* mutations are associated with hematopoietic disorders [70]. *CREBRF* was significantly upregulated by reconstituted high-density lipoproteins (rHDL) in response to both hypoxia and inflammation, and is novel target of rHDL in response to angiogenic stimuli [71].

The brain shows to be more sensitive to the fluctuation of oxygen content, brain damage caused by hypoxia impaired cognitive physiology and psychological function, and lead to learning and memory defects [72]. Among the candidate genes for GH100 and GJ, *PRMT8* is an important regulatory component of membrane phospholipid composition, short-term memory function, mitochondrial function, and neuroinflammation in response to hypoxic stress [73]. *GRM4* is crucial in ion channel regulation, neuronal excitation and neurotransmitter release [74]. Great concern

has been attached to the mechanism of *GRM4* in cancer biology in recent decades [75]. *ASPM* expression is critical for proper neurogenesis and neuronal migration, and *ASPM* is a positive regulator of Wnt signaling [76]. A large number of neuron-related genes are subject to selection among different breeds, possibly related to the different environments and life patterns in which different breeds are living. *TOE1* and *BTBD3* plays a crucial role in neurodevelopment [77, 78]. *NRCAM* plays a wide variety of roles in neural development, including cell proliferation and differentiation, axon growth and guidance, synapse formation, and the formation of the myelinated nerve structure [79]. There is extensive evidence to confirm that *PINK1* functions to protect neurons against stress-induced cell death, and *PINK1* deficiency has been shown to have a profound effect on respiration and reactive oxygen species (ROS) production [80]. Aberrant expression of *TENM4* may leads to lower learning ability, sleep reduction, and increased aggressiveness in animal models [81]. *ASTN2* has pleiotropic effects on cardiometabolic and psychiatric traits, and associations between genetic variants in the *ASTN2* locus and blood pressure, total and central obesity, neuroticism, anhedonia and mood instability were identified [82].

High-altitude environmental factors (such as UV exposure, cold and hypobaric hypoxia) can affect the immune system and increase susceptibility to cancer, various infections, and even autoimmune disease [83]. *BPI* and *LBP* were identified in the selection signatures of GJ. In mammals, bactericidal permeability/increasing protein (BPI) and lipopolysaccharide (LPS) binding protein (LBP) are two important pattern-recognition receptors involved in the innate immune response to LPS [84]. Mammalian BPI has been shown to play bactericidal, opsonic and anti-inflammatory roles [85]. The strong solar ultraviolet radiation (UVR) increased the risk of skin cancer in plateau animals. Melanin in the skin as the main component shielding UVR, prevents skin and organ damage by absorbing UV photons and free radicals to protect DNA from UV damage [86, 87]. *TYRP1* is related to melanin synthesis, the biology of melanocytes and melanosomes, and the migration and survival of melanocytes during development [88]. The expression of *RAB27A* and *RAB27B* in both melanocytes and platelets makes them candidates for involvement in mouse and human disorders characterized by the combination of pigment dilution and a platelet storage pool defect [89]. *OCA2* encodes the protein P, a transmembrane protein, and has been shown to play a role in pigmentation in both humans and mice [90]. In humans, it has been implicated in iris, skin, and hair pigmentation [91]. *TYRP1* was identified in the selection signatures of GJ.

RAB27B and *OCA2* were identified in the comparisons of different breeds.

In addition, the selective signatures between breeds were annotated to a number of other growth and development-related genes. *FGF10* serves as an essential regulator of lung and limb formation [92]. The *LAMA3* gene belongs to the laminin family and is responsive to several epithelial-mesenchymal regulators, including keratinocyte growth factor, epidermal growth factor and skin fibrosis [93]. SNP loci within the *CRTAC1* had a positive impact on milk fat, lactose, and total solids (TS) content in Gannan yak milk [94]. Variation in the *XKR4* gene is significantly associated with subcutaneous rump fat thickness in indicine and composite cattle [95], and is associated with feed intake and growth phenotypes in cattle [96]. *PITX1* as a key specificity factor in HIF-1 α dependent responses is required for the survival and proliferation responses in hypoxia [97]. *PITX1* is required for pituitary gland and hind limb development [98].

In this study, a series of genes related to cardiac health and development, neuronal development and function, angiogenesis and hematopoietic were found through genome analysis of dairy cattle at Tibet, and these genes showed significant selection pressure in the plateau environment. These findings provide potential molecular markers for the breeding and genetic improvement of dairy cows, which can help select superior dairy cattle breeds that are better adapted to the plateau environment and improve their survival, fertility and milk yield in low oxygen environment. There is still a need to find or validate variant loci in candidate genes associated with plateau acclimatization and adaptation in combination with phenotypic data, as well as other sequencing and analytical methods, and with reference to results in other species. More genotyping data of Jersey cattle can help to identify whether some of the variation is breed-specific, and provide data to support the exploration of acclimatization and adaptation mechanisms in the two breeds.

Farmers can gradually improve the plateau adaptability and production performance of dairy herds by selecting cows with favorable genes for breeding, regardless of whether the variation is acquired through crossbreeding or natural selection. Traditional breeding methods and modern genome selection techniques can be combined to more effectively select individuals with anti-hypoxia and enhanced immunity through marker-assisted selection (MAS) or whole-genome selection (GS). Scientists need to continuously monitor the performance of these genes in different environments, optimize breeding programs, and develop nutrition and management strategies suitable for the plateau environment to help farmers improve the growing environment of dairy cows. At the same time, these results also laid a foundation for other

livestock adaptation research in the plateau area, with a wide range of application prospects.

Conclusion

The composition of the Tibet dairy cattle population is complex. There are significant genetic background differences between Tibet dairy cattle and Tibet native cattle populations. A number of genes associated with cardiac health and development (*SCN5A*, *SCN10A*, *RAP1GAP*, *SNX13*, *PPM1k*, *HIPK2*, *THBS4*, *DIPK2A*, *NKX2-5*, *SMAD7*, *SFRP5*, *HP1BP3*, *IRX4*), neuronal development and function (*PRMT8*, *GRM4*, *ASPM*, *TOE1*, *BTBD3*, *NRCAM*, *PINK1*, *TENM4*, *ASTN2*), angiogenesis and hematopoietic (*ALPL*, *APOB*, *LYN*, *ZFYVE27*, *P4HA2*, *MTOR*, *TERT*, *CREBRF*), and pigmentation (*TYRP1*, *RAB27B*, *OCA2*) are selected for in plateau dairy cattle. In addition, genes related to immune response (*BPI*, *LBP*) were selected for plateau Jersey cattle, and genes related to growth and development (*FGF10*, *LAMA3*, *CRTAC1*, *XKR4*, *PITX1*) were selected for in crossbred cattle. These candidate genes provide valuable information for further fine-mapping studies and functional genomic studies as a starting point to identify causal genetic variants that control plateau adaptation traits for the implementation of breeding schemes aiming genetic improvement of plateau dairy cattle populations.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11335-0>.

Additional file 1: Table S1. Cross validation (CV) error for ADMIXTURE analysis.

Additional file 2: Table S2. List of the significant SNP markers for iHS analysis in GH100. Table S3. List of the significant SNP markers for XP-EHH analysis in GH100 and PH. Table S4. List of the significant SNP markers for F_{ST} analysis in GH100 and PH. Table S5. List of the significant SNP markers for iHS analysis in GJ. Table S6. List of the significant SNP markers for XP-EHH analysis in GH100 and PH. Table S7. List of the significant SNP markers for F_{ST} analysis in GJ and PJ.

Additional file 3: Table S8. Gene enrichment analysis over the DAVID Gene Ontology (GO; Biological Process) resource of genes under selection in GH100_PH and GJ_PJ analysis. Table S9. Gene enrichment analysis over the cluster profiler Gene Ontology (GO; Biological Process) resource of genes under selection in GH100_PH and GJ_PJ analysis. Table S10. Gene enrichment analysis over the DAVID Gene Ontology (GO; Biological Process) resource of genes under selection in GH40_PH and GH40_XZ analysis. Table S9. Gene enrichment analysis over the cluster profiler Gene Ontology (GO; Biological Process) resource of genes under selection in GH40_PH and GH40_XZ analysis.

Acknowledgements

Thanks to Wenqi Lou, Jingyang Ning, Hailiang Zhang, Lei Wang, Rui Shi, Yao Chang and Changwei Qu for their contributions to this study.

Authors' contributions

S.H. and L.M. collected the blood samples, analyzed the data and draft the article; B.L. designed the experiment and managed the cattle; J.D. and Q.X.

revised the article; Y.W. designed the experiment and revised the article. All authors reviewed the manuscript.

Funding

This study was supported by the Key Research Project of Tibet Autonomous Region (No. CGZH2023000257). Special Item of Regional Collaborative Innovation in Tibet Autonomous Region (No. QYXTZX-LS2021-01). The earmarked fund for CARS-36. The Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62).

Data availability

The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request. Data are located in controlled access data storage at College of Animal Science and Technology, China Agricultural University.

Declarations

Ethics approval and consent to participate

The cattle selected in this study obtained informed consent from the owners, who agreed that we can use these animals in this study. All procedures involving animals were approved by the Institutional Animal Care and Use Committee at the China Agricultural University, China.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹National Engineering Laboratory for Animal Breeding, Laboratory of Animal Genetics, Breeding and Reproduction, MARA, State Key Laboratory of Animal Biotech Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China. ²Institute of Animal Husbandry and Veterinary Medicine, Academy of Agriculture and Animal Husbandry of Tibet Autonomous Region, Lhasa 850000, China. ³Animal Science and Technology College, Beijing University of Agriculture, Beijing 102206, China. ⁴Institute of Life Science and Bioengineering, Beijing Jiaotong University, Beijing 100044, China.

Received: 16 July 2024 Accepted: 7 February 2025

Published online: 12 March 2025

References

- Ramirez JM, Folkow LP, Blix AS. Hypoxia tolerance in mammals and birds: from the wilderness to the clinic. *Annu Rev Physiol*. 2007;69:113–43.
- Cao BK, Xu XM. Reflections on the introduction of dairy cattle and the development of dairy industry in Lhasa. *J Anim Sci Vet Med*. 2007;26(6):50–51. 53 (in Chinese).
- Basang ZZ, Danzeng LS, Zhao L, et al. Effect analysis of improving Tibetan yellow cattle by Holstein frozen semen. *Chin J Anim Sci*. 2023;59(1):149–53 (in Chinese).
- Lang Y, Wang GX, Wu PR. The dilemma and countermeasures of milk industry development in China. *Heilongjiang Anim Sci Vet Med*. 2020(4):12–6. 147 (in Chinese).
- Suolang QJ, Basang ZZ, Zhao L, et al. Observation and research on plateau adaptability of introduced breeding cows in Tibet. *China Feed*. 2019;2:11–5 (in Chinese).
- Gnecchi-Ruscone GA, et al. Evidence of polygenic adaptation to high altitude from Tibetan and Sherpa genomes. *Genome Biol Evol*. 2018;10(11):2919–30.
- Yang J, et al. Genetic signatures of high-altitude adaptation in Tibetans. *Proc Natl Acad Sci U S A*. 2017;114(16):4189–94.
- Yang J, et al. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Mol Biol Evol*. 2016;33(10):2576–92.
- Gorkhali NA, et al. Genomic analysis identified a potential novel molecular mechanism for high-altitude adaptation in sheep at the Himalayas. *Sci Rep*. 2016;6:29963.
- Edea Z, et al. Linkage disequilibrium and genomic scan to detect selective loci in cattle populations adapted to different ecological conditions in Ethiopia. *J Anim Breed Genet*. 2014;131(5):358–66.
- Zhang Y, et al. Population structure, and selection signatures underlying high-altitude adaptation inferred from genome-wide copy number variations in Chinese indigenous cattle. *Front Genet*. 2019;10:1404.
- Guan X, et al. The three missense mutations of EPAS1, IL37 and EEF1D genes associated with high-altitude adaptation in Chinese cattle. *Anim Genet*. 2020;51(6):987–8.
- Wang X, et al. Introgression, admixture, and selection facilitate genetic adaptation to high-altitude environments in cattle. *Genomics*. 2021;113(3):1491–503.
- Iqbal N, et al. Genomic variants identified from whole-genome resequencing of indicine cattle breeds from Pakistan. *PLoS One*. 2019;14(4):e0215065.
- Norman HD, et al. 0324 Breed base representation in dairy animals of five breeds. *J Anim Sci*. 2016;94(suppl_5):155–6.
- Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform*. 2013;14(2):144–61.
- Browning BL, Zhou Y, Browning SR. A One-Penny imputed genome from next-generation reference panels. *Am J Hum Genet*. 2018;103(3):338–48.
- Chang CC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.
- Zhang C, et al. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*. 2019;35(10):1786–8.
- Barbato M, et al. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front Genet*. 2015;6:109.
- Lukicheva S, Mardulyn P. Whole-genome sequencing reveals asymmetric introgression between two sister species of cold-resistant leaf beetles. *Mol Ecol*. 2021;30(16):4077–89.
- Lefort V, Desper R, Gascuel O. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol*. 2015;32(10):2798–800.
- Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49(W1):W293–6.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655–64.
- Danecek P, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156–8.
- Szpiech ZA, Hernandez RD. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Mol Biol Evol*. 2014;31(10):2824–7.
- Gautier M, Naves M. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Mol Ecol*. 2011;20(15):3128–43.
- Zinovieva NA, et al. Selection signatures in two oldest Russian native cattle breeds revealed using high-density single nucleotide polymorphism analysis. *PLoS One*. 2020;15(11):e0242200.
- Singh A, et al. Signatures of selection in composite vrindavani cattle of India. *Front Genet*. 2020;11:589496.
- Maiorano AM, et al. Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLoS One*. 2018;13(8):e0200694.
- Cardoso DF, et al. Genome-wide scan reveals population stratification and footprints of recent selection in Nelore cattle. *Genet Sel Evol*. 2018;50(1):22.
- Fonseca P, et al. GALLO: an R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *Gigascience*. 2020;9(12):giaa149.
- Sherman BT, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022;50(W1):W216–21.
- Yu G, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–7.

35. Zhao P, et al. Evidence of evolutionary history and selective sweeps in the genome of Meishan pig reveals its genetic and phenotypic characterization. *Gigascience*. 2018;7(5):gij058.
36. Perez OA, et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet Sel Evol*. 2014;46(1):19.
37. Storz JF, Scott GR. Phenotypic plasticity, genetic assimilation, and genetic compensation in hypoxia adaptation of high-altitude vertebrates. *Comp Biochem Physiol A Mol Integr Physiol*. 2021;253:110865.
38. Friedrich J, Wiener P. Selection signatures for high-altitude adaptation in ruminants. *Anim Genet*. 2020;51(2):157–65.
39. Cadzow M, et al. A bioinformatics workflow for detecting signatures of selection in genomic data. *Front Genet*. 2014;5:293.
40. Gautier M, Klassmann A, Vitalis R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol Ecol Resour*. 2017;17(1):78–90.
41. Sabeti PC, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 2007;449(7164):913–8.
42. Ivy CM, Scott GR. Control of breathing and the circulation in high-altitude mammals and birds. *Comp Biochem Physiol A Mol Integr Physiol*. 2015;186:66–74.
43. Holt TN, Callan RJ. Pulmonary arterial pressure testing for high mountain disease in cattle. *Vet Clin North Am Food Anim Pract*. 2007;23(3):575–96, vii.
44. Li W, et al. SCN5A variants: association with cardiac disorders. *Front Physiol*. 2018;9:1372.
45. van den Boogaard M, et al. A common genetic variant within SCN10A modulates cardiac SCN5A expression. *J Clin Invest*. 2014;124(4):1844–52.
46. Gao Y, et al. Rap1GAP mediates angiotensin II-induced cardiomyocyte hypertrophy by inhibiting autophagy and increasing oxidative stress. *Oxid Med Cell Longev*. 2021;2021:7848027.
47. Shan T, et al. Rap1GAP exacerbates myocardial infarction by regulating the AMPK/SIRT1/NF-kappaB signaling pathway. *Cell Signal*. 2024;117:111080.
48. Li J, et al. SNX13 reduction mediates heart failure through degradative sorting of apoptosis repressor with caspase recruitment domain. *Nat Commun*. 2014;5:5177.
49. Lian K, et al. PP2Cm overexpression alleviates MI/R injury mediated by a BCAA catabolism defect and oxidative stress in diabetic mice. *Eur J Pharmacol*. 2020;866:172796.
50. Zhou Q, et al. Exercise downregulates HIPK2 and HIPK2 inhibition protects against myocardial infarction. *EBioMedicine*. 2021;74:103713.
51. Calzado MA, et al. From top to bottom: the two faces of HIPK2 for regulation of the hypoxic response. *Cell Cycle*. 2009;8(11):1659–64.
52. Frolova EG, et al. Thrombospondin-4 regulates fibrosis and remodeling of the myocardium in response to pressure overload. *FASEB J*. 2012;26(6):2363–73.
53. Cingolani OH, et al. Thrombospondin-4 is required for stretch-mediated contractility augmentation in cardiac muscle. *Circ Res*. 2011;109(12):1410–4.
54. Huang J, et al. HASF is a stem cell paracrine factor that activates PKC epsilon mediated cytoprotection. *J Mol Cell Cardiol*. 2014;66:157–64.
55. Akazawa H, Komuro I. Cardiac transcription factor Csx/Nkx2-5: Its role in cardiac development and diseases. *Pharmacol Ther*. 2005;107(2):252–68.
56. Chen Q, et al. Smad7 is required for the development and function of the heart. *J Biol Chem*. 2009;284(1):292–300.
57. Nakamura K, et al. Secreted frizzled-related protein 5 diminishes cardiac inflammation and protects the heart from ischemia/reperfusion injury. *J Biol Chem*. 2016;291(6):2566–75.
58. Miyoshi T, et al. Low serum level of secreted frizzled-related protein 5, an anti-inflammatory adipokine, is associated with coronary artery disease. *Atherosclerosis*. 2014;233(2):454–9.
59. Chan CX, Tan W, Foo R. Abstract 458: role of heterochromatin protein 1 binding partner 3, HP1BP3, in cardiomyocyte stress response. *Circ Res*. 2018;123(Suppl_1):A458–A458.
60. Bao ZZ, et al. Regulation of chamber-specific gene expression in the developing heart by Irx4. *Science*. 1999;283(5405):1161–4.
61. Gassmann NN, et al. Pregnancy at high altitude in the Andes leads to increased total vessel density in healthy newborns. *J Appl Physiol* (1985). 2016;121(3):709–15.
62. Carmeliet P. Mechanisms of angiogenesis and arteriogenesis. *Nat Med*. 2000;6(4):389–95.
63. Dong J, et al. ALPL regulates pro-angiogenic capacity of mesenchymal stem cells through ATP-P2X7 axis controlled exosomes secretion. *J Nanobiotechnology*. 2024;22(1):172.
64. Benn M. Apolipoprotein B levels, APOB alleles, and risk of ischemic cardiovascular disease in the general population, a review. *Atherosclerosis*. 2009;206(1):17–30.
65. Sweet-Jones J, et al. Genotyping and whole-genome resequencing of welsh sheep breeds reveal candidate genes and variants for adaptation to local environment and socioeconomic traits. *Front Genet*. 2021;12:612492.
66. Ingley E. Functions of the Lyn tyrosine kinase in health and disease. *Cell Commun Signal*. 2012;10(1):21.
67. Arora A, et al. Protrudin regulates FAK activation, endothelial cell migration and angiogenesis. *Cell Mol Life Sci*. 2022;79(4):220.
68. Teodoro JG, et al. p53-mediated inhibition of angiogenesis through up-regulation of a collagen prolyl hydroxylase. *Science*. 2006;313(5789):968–71.
69. Land SC, Tee AR. Hypoxia-inducible factor 1alpha is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif. *J Biol Chem*. 2007;282(28):20534–43.
70. Yamaguchi H, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. *N Engl J Med*. 2005;352(14):1413–24.
71. Wong N, et al. Exploring the roles of CREBRF and TRIM2 in the regulation of angiogenesis by high-density lipoproteins. *Int J Mol Sci*. 2018;19(7):1903.
72. Zhang ZA, et al. Insight into the effects of high-altitude hypoxic exposure on learning and memory. *Oxid Med Cell Longev*. 2022;2022:4163188.
73. Couto ESA, et al. Protein arginine methyltransferase 8 modulates mitochondrial bioenergetics and neuroinflammation after hypoxic stress. *J Neurochem*. 2021;159(4):742–61.
74. Muhle H, et al. Role of GRM4 in idiopathic generalized epilepsies analysed by genetic association and sequence analysis. *Epilepsy Res*. 2010;89(2–3):319–26.
75. Zhang Z, et al. GRM4 inhibits the proliferation, migration, and invasion of human osteosarcoma cells through interaction with CBX4. *Biosci Biotechnol Biochem*. 2020;84(2):279–89.
76. Buchman JJ, Durak O, Tsai LH. ASPM regulates Wnt signaling pathway activity in the developing brain. *Genes Dev*. 2011;25(18):1909–14.
77. Deng T, et al. Toe1 promotes proliferation and differentiation of neural progenitor cells. *Heliyon*. 2024;10(20):e39535.
78. Matsui A, et al. BTBD3 controls dendrite orientation toward active axons in mammalian neocortex. *Science*. 2013;342(6162):1114–8.
79. Sakurai T. The role of NrCAM in neural development and disorders—beyond a simple glue in the brain. *Mol Cell Neurosci*. 2012;49(3):351–63.
80. Deas E, Plun-Favreau H, Wood NW. PINK1 function in health and disease. *EMBO Mol Med*. 2009;1(3):152–65.
81. Yi X, et al. Genetic and functional analysis reveals TENM4 contributes to schizophrenia. *iScience*. 2021;24(9):103063.
82. Burt O, et al. Genetic variation in the ASTN2 locus in cardiovascular, metabolic and psychiatric traits: evidence for pleiotropy rather than shared biology. *Genes (Basel)*. 2021;12(8):1194.
83. Mishra KP, Ganju L. Influence of high altitude exposure on the immune system: a review. *Immunol Invest*. 2010;39(3):219–34.
84. Miguel MA, Mingala CN. Screening of Pig (*Sus scrofa*) Bactericidal Permeability-Increasing Protein (BPI) gene as marker for disease resistance. *Anim Biotechnol*. 2019;30(2):146–50.
85. Theprungsirikul J, Skopelja-Gardner S, Rigby W. Killing three birds with one BPI: bactericidal, opsonic, and anti-inflammatory functions. *J Transl Autoimmun*. 2021;4:100105.
86. Mitchell DL, Paniker L, Douki T. DNA damage, repair and photoadaptation in a Xiphophorus fish hybrid. *Photochem Photobiol*. 2009;85(6):1384–90.
87. Wu C, et al. Drivers of plateau adaptability in cashmere goats revealed by genomic and transcriptomic analyses. *BMC Genomics*. 2023;24(1):428.
88. Rajawat D, et al. Uncovering genes underlying coat color variation in indigenous cattle breeds through genome-wide positive selection. *Anim Biotechnol*. 2023;34(8):3920–33.
89. Chen D, et al. Molecular cloning and characterization of rab27a and rab27b, novel human rab proteins shared by melanocytes and platelets. *Biochem Mol Med*. 1997;60(1):27–37.

90. Frudakis T, et al. Sequences associated with human iris pigmentation. *Genetics*. 2003;165(4):2071–83.
91. Sulem P, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*. 2007;39(12):1443–52.
92. Sekine K, et al. Fgf10 is essential for limb and lung formation. *Nat Genet*. 1999;21(1):138–41.
93. Pesch M, König S, Aumailley M. Targeted disruption of the Lama3 gene in adult mice is sufficient to induce skin inflammation and fibrosis. *J Invest Dermatol*. 2017;137(2):332–40.
94. Zhang J, et al. Polymorphisms within the IQGAP2 and CRTAC1 genes of Gannan yaks and their association with milk quality characteristics. *Foods*. 2024;13(23):3720.
95. Porto NL, et al. Variation in the XKR4 gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. *Anim Genet*. 2012;43(6):785–9.
96. Lindholm-Perry AK, et al. A region on BTA14 that includes the positional candidate genes LYPLA1, XKR4 and TMEM68 is associated with feed intake and growth phenotypes in cattle(1). *Anim Genet*. 2012;43(2):216–9.
97. Mudie S, et al. PITX1, a specificity determinant in the HIF-1alpha-mediated transcriptional response to hypoxia. *Cell Cycle*. 2014;13(24):3878–91.
98. Tremblay JJ, Goodyer CG, Drouin J. Transcriptional properties of Ptx1 and Ptx2 isoforms. *Neuroendocrinology*. 2000;71(5):277–86.

Publisher's Note

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