

# Analysis of 206 whole-genome resequencing reveals selection signatures associated with breed-specific traits in Hu sheep

Fuping Zhao<sup>1</sup>  | Rui Xie<sup>1,2</sup> | Lingzhao Fang<sup>3</sup>  | Ruidong Xiang<sup>4</sup>  | Zehu Yuan<sup>5</sup> | Yang Liu<sup>2</sup> | Lixian Wang<sup>1</sup>

<sup>1</sup>State Key Laboratory of Animal Biotech Breeding, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China

<sup>2</sup>Department of Animal Genetics, Breeding and Reproduction, National Experimental Teaching Demonstration Center of Animal Science, College of Animal Science and Technology, Nanjing Agricultural University, Nanjing, China

<sup>3</sup>Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus, Denmark

<sup>4</sup>Faculty of Veterinary and Agricultural Science, The University of Melbourne, Parkville, Victoria, Australia

<sup>5</sup>Joint International Research Laboratory of Agriculture and Agri-Product Safety of Ministry of Education, Yangzhou University, Yangzhou, China

## Correspondence

Yang Liu, Department of Animal Genetics, Breeding and Reproduction, National Experimental Teaching Demonstration Center of Animal Science, College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China.  
Email: [yangliu@njau.edu.cn](mailto:yangliu@njau.edu.cn)

Lixian Wang, State Key Laboratory of Animal Biotech Breeding, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China.  
Email: [iaswlx@263.net](mailto:iaswlx@263.net)

## Funding information

the Natural Science Foundations of China, Grant/Award Number: 32172702; National Key Research and Development Program of China, Grant/Award Number: 2021YFD1301101; Agricultural Science and Technology Innovation Program, Grant/Award Number: ASTIP-IAS02

## Abstract

As an invaluable Chinese sheep germplasm resource, Hu sheep are renowned for their high fertility and beautiful wavy lambskins. Their distinctive characteristics have evolved over time through a combination of artificial and natural selection. Identifying selection signatures in Hu sheep can provide a straightforward insight into the mechanism of selection and further uncover the candidate genes associated with breed-specific traits subject to selection. Here, we conducted whole-genome resequencing on 206 Hu sheep individuals, each with an approximate 6-fold depth of coverage. And then we employed three complementary approaches, including composite likelihood ratio, integrated haplotype homozygosity score and the detection of runs of homozygosity, to detect selection signatures. In total, 10 candidate genomic regions displaying selection signatures were simultaneously identified by multiple methods, spanning 88.54 Mb. After annotating, these genomic regions harbored collectively 92 unique genes. Interestingly, 32 candidate genes associated with reproduction were distributed in nine genomic regions detected. Out of them, two stood out as star candidates: *BMP1R1B* and *GNRH2*, both of which have documented associations with fertility, and a *HOXA* gene cluster (*HOXA1-5*, *HOXA9*, *HOXA10*, *HOXA11* and *HOXA13*) had also been linked to fertility. Additionally, we identified other genes that are related to hair follicle development (*LAMTOR3*, *EEF1A2*), ear size (*HOXA1*, *KCNQ2*), fat tail formation (*HOXA10*, *HOXA11*), growth and development (*FAF1*, *CCNDBP1*, *GJB2*, *GJA3*), fat deposition (*ACOXL*, *JAZF1*, *HOXA3*, *HOXA4*, *HOXA5*, *EBF4*), immune (*UBR1*, *FASTKD5*) and feed intake (*DAPP1*, *RNF17*, *NPBWR2*). Our results offer novel insights into the genetic mechanisms underlying the selection of breed-specific traits in Hu sheep and provide a reference for sheep genetic improvement programs.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Evolutionary Applications* published by John Wiley & Sons Ltd.

## KEYWORDS

high prolificacy, Hu sheep, selection signature, whole-genome sequencing data

## 1 | INTRODUCTION

Hu sheep is an invaluable Chinese indigenous sheep germplasm resource and well-known for its hyper-prolificacy and the beautiful wavy lambskins (CNCAGR, 2012). They originated from Mongolian sheep and were introduced from the pastoral region of North China to the Taihu Lake basin about 900 years ago (CNCAGR, 2012; Geng et al., 2002). After being introduced to South China, Hu sheep are raised indoors throughout the year and have adapted to the local hot and humid climate (Abied, Bagadi, et al., 2020). After long-time artificial and natural selection, they have possessed their breed-specific features, such as precocious puberty, perennial oestrus, and high fecundity (Feng et al., 2018). Their specific features are different from Mongolian sheep that exhibit seasonal estrous and singleton breeding. Now, Hu sheep have been utilized as the maternal line in sheep crossbreeding programs, and are almost distributed across China.

In the course of selection, the beneficial allelic frequencies can rapidly increase under positive selection pressure and be fixed within a population. The distinctive imprints were just sculpted on the genome, which are called selection signatures (Qanbari & Simianer, 2014). Scanning these selection signatures is crucially important in animal genetics. The genomic regions under selection often harbor QTLs or genes influencing some economically important traits (Ceballos et al., 2018; Rubin et al., 2010; Zhao et al., 2015). Coupling with the development of high-throughput sequencing technology, the reduction of sequencing cost makes it possible to detect selection signatures on the whole genome. Thus, the selection signature analysis can reveal genomic regions of interest for selection and explore the potential genetic mechanisms of phenotypic polymorphisms and adaptation.

The statistical methods for detecting the selection signatures within a population can be classified into three primary groups (Qanbari & Simianer, 2014; Saravanan et al., 2020). The first group is a statistic identifying the site frequency spectrum to determine whether allele frequencies are concordant with the hypothesis of genetic neutrality. Such as Tajima's D (Tajima, 1989), Fay and Wu's H statistic (Fay & Wu, 2000) and the composite likelihood ratio test (CLR) (Nielsen, 2005). The second group consists of powerful tests that analyzes the level of linkage disequilibrium (LD) associated with the target haplotypes and/or alleles of selection. Such as extended haplotype homozygosity (EHH), relatively extended haplotype homozygosity (REHH) (Sabeti et al., 2002) and integrated haplotype score (iHS) (Voight et al., 2006). The third group is established on assessment of DNA polymorphism levels, which includes runs of homozygosity (ROH) (McQuillan et al., 2008) and pooled heterozygosity ( $H_p$ ) (Rubin et al., 2010). These methods have been successfully implemented to detect positive selection signatures in domestic

animals (Liu et al., 2021; Saravanan et al., 2020; Shi et al., 2020; Zhao et al., 2015, 2016).

In previous studies, Hu sheep together with other sheep breeds, was used in inter-populations analysis to identify the genomic region showing evidence of positive selection not within Hu sheep itself (Abied, Bagadi, et al., 2020; Abied, Xu, et al., 2020; Liu et al., 2016; Wei et al., 2015). The main aim of this study is to detect selection signatures only in the Hu sheep breed. We used whole genome re-sequencing data of 206 individuals and applied three complementary methods (CLR, iHS and ROH) to identify genomic regions putatively under selection. To reduce the false positive rate, genomic regions simultaneously identified by multiple methods were considered as being under selection. Then bioinformatics analysis was also performed to explain the biological function of the candidate genes residing in genomic regions. Our findings will help to better elucidate the genetic mechanism of breed-specific traits of Hu sheep, and provide a reference for the genetic improvement of Hu sheep.

## 2 | MATERIALS AND METHODS

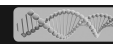
### 2.1 | Ethics statement

All animal experiments in this study were fully approved by the Animal Care and Use Committee of the Institute of Animals Science, Chinese Academy of Agricultural Sciences (IAS-CAAS) with the following reference number: IASCAAS-AE-03, September 2014.

### 2.2 | Animal sampling and whole-genome resequencing

Two hundred and six Hu sheep were randomly selected from a Hu sheep breeding farm in the Gansu province of China. Ear tissue samples were obtained. The genomic DNA of these samples was extracted using a standard phenol-chloroform protocol.

The quantity and quality of genomic DNA were examined using a NanoDrop-2000 device (Thermo Fisher Scientific, Wilmington, DE, USA) and by 1% agarose gel electrophoresis. For each sample, at least 3  $\mu$ g genomic DNA was used for sequencing library construction with an average insert size of 450bp. After the examinations, we generated paired-end libraries for each eligible sample using standard procedures. Next, sequencing was carried out using a 2  $\times$  150bp pair-end configuration on an Illumina HiSeq X. The average read sequence coverage was 6 $\times$  for all samples. The depth ensured the accuracy of variant calling and genotyping and met the requirements for next-step analyses.



## 2.3 | Variant discovery and genotyping

The raw reads were trimmed by Trimmomatic v0.39 (Bolger et al., 2014), and then BWA v0.7.17 software (Li & Durbin, 2009) was used to map high-quality trimmed reads to the domestic sheep reference genome (*Ovis aries*, Oar\_v4.0, [https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000298735.2/](https://www.ncbi.nlm.nih.gov/assembly/GCF_000298735.2/)). Samtools v1.15 (Li et al., 2009) was utilized to covert alignment sam files into bam format, sort, index and generated an mpileup file containing genotype depth and allelic depth. Picard v2.0.1 (Li & Durbin, 2009) was used to remove MarkDuplicates and avoid any influence on variant detection. After that, GATK v4.0.2.1 (DePristo et al., 2011) was performed with SNP calling. To ensure the high quality of the result, the software VCFtools v0.1.17 (Danecek et al., 2011) was used to further filter the output of GATK. Only the SNPs on the autosome were retained for further analyses and other SNPs that did not meet the following standards were excluded: (i) mean sequencing depth > 2X (over all samples); (ii) maximum missing rate < 0.05; (iii) a minor allele frequency > 0.05; (iv) *p*-value of Hardy-Weinberg equilibrium >  $1 \times 10^{-6}$ ; and (v) only keep two alleles. The reference genome link is Oar\_v4.0.

## 2.4 | Selection signature detection

In this study, we employed three complementary methods to identify the selection signatures in the Hu sheep population, which included CLR, iHS and ROH detection.

### 2.4.1 | Composite likelihood ratio (CLR) statistic

CLR test compares a neutral model for allele frequency spectrum with a selective sweep model since a recent selective sweep can cause allele frequency spectra to depart from the expectation under neutrality and reduced genetic diversity (Nielsen, 2005). In the neutral model, the probability of allele frequency spectrum is derived from the background pattern of variation in the genome. The CLR statistic does not only evaluate the skewness of the frequency spectrum across multiple loci but also incorporates information on the recombination rate to distinguish selection from other demographic events (Nielsen, 2005). In this study, the SweepFinder2 software (DeGiorgio et al., 2016) was used to calculate CLR value per locus. To increase statistical power and reduce sampling noise, the whole genome was divided into 500kb windows with a 250kb overlap and the maximum CLR value in each window was used as the test statistic. Windows at the top 1% of the empirical distribution were selected as candidate regions.

### 2.4.2 | Integrated haplotype score (iHS)

As a representative for haplotype-based methods, the iHS score is developed based on extended haplotype homozygosity (EHH), and

compares the extent of LD between haplotypes carrying the ancestral and derived alleles. The ancestral alleles of sheep SNPs were downloaded from <https://www.sheephapmap.org>. For the iHS test, we used rehh package v3.0 of R software (Gautier & Vitalis, 2012) to calculate iHS value. The unstandardized integrated haplotype score (iHS) is computed using the equation  $\text{uniHS} = \ln\left(\frac{iHH_A}{iHH_D}\right)$ , where  $iHH_A$  and  $iHH_D$  represented the integrated EHH value of the ancestor allele and derived allele, respectively. Then, the iHS standardization was performed by setting:  $iHS = \frac{\text{uniHS} - \text{Mean}(\text{uniHS})}{\text{SD}(\text{uniHS})}$ , where Mean (uniHS) and SD (uniHS) respectively represent the mean and standard deviation of uniHS restricted to those markers with a similar derived allele frequency as observed at the target marker. When the iHS value is a large positive value, it indicates that the haplotype may be the same as the ancestor, while when the iHS value is a large negative value, it indicates that there may be a new derived haplotype (Voight et al., 2006). In the same manner, the genome was also divided into 500kb windows with a 250kb overlap and the averaged |iHS| value in each window was used as its test statistic. In order to calculate the *p*-value at the genomic level, iHS score value per window was further transformed following the equation  $P_{iHS} = -\log_{10}[1 - 2|\Phi(iHS) - 0.5|]$ , where  $\Phi(x)$  represents the Gaussian cumulative distribution function (under neutrality), and  $P_{iHS}$  is the two-sided *p*-value related to the neutral hypothesis. The  $P_{iHS}$  surpassing the top 1% of the empirical distribution were considered as candidate regions.

### 2.4.3 | ROH detection

In this study, we conducted ROH detection with consecutive runs module in detectRUNS 0.9.6 R package (Biscarini et al., 2018). The following criteria were selected for ROH detection: (i) the minimum length of one ROH was set 300kb to avoid the impact of strong LD, (ii) no more than two missing SNP and one heterozygous genotype were allowed per ROH, (iii) the max gap between two consecutive ROH was set 500kb, (iv) the minimum SNP number (*l*) required in one ROH was set to 88, which was computed using the equation proposed by Lencz et al. (2007),  $l = \frac{\ln[\alpha / (n_i \times n_j)]}{\ln(1 - \text{het})}$ , where  $\alpha$  is the false positive rate of ROH identified (set to 0.05 in the present study),  $n_i$  is the number of SNPs per individual,  $n_j$  is the number of individuals and *het* is the proportion of heterozygosity across all SNPs.

To identify the high-frequency ROH regions in the Hu sheep genome, we calculated the percentage of the occurrence of SNPs in ROH by counting the times of an SNP detected in an ROH across all individuals. In this study, the threshold for identifying the genomic regions most commonly associated with ROH was the top 0.1% SNPs observed in ROH genomic regions.

## 2.5 | Functional annotation of genomic regions under selection

We retained the candidate genomic regions simultaneously detected by at least two methods mentioned above. Gene annotation

for regions under putative selection was performed based on the domestic sheep reference (Oar\_v4.0). The search for positional candidate genes was extended 50 kb up- and downstream from genomic regions identified based on the result of LD analysis. The biological function of the positional genes annotated was carried out by survey of relevant literature.

## 2.6 | Enrichment analysis

The annotated genes were further perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis using the DAVID online software (<http://david.abcc.ncifcrf.gov/home.jsp>).

## 3 | RESULTS

### 3.1 | Basic statistic of resequencing data of Hu sheep

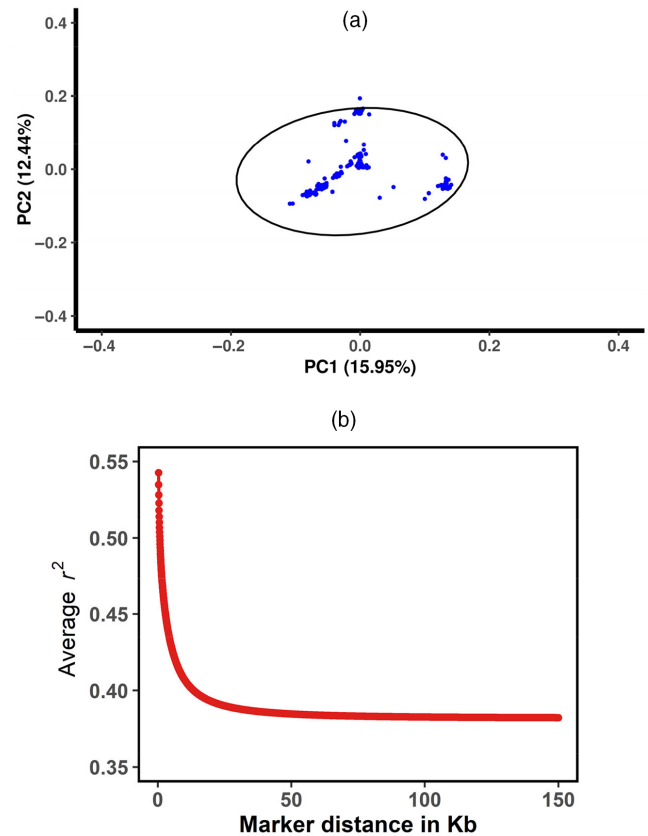
In this study, we generated an average of 53,417,468 clean reads for 206 WGS data, ranging from 26,631,650 to 148,454,839. After quality control, we obtained a total of 1,604,526 SNPs.

### 3.2 | Population analysis

PCA showed the total genetic variance in the population was explained by the first three principal components 15.95%, 12.44% and 9.78%, respectively. As seen in [Figure 1a](#), individuals resided in the region of 99% confidence ellipse. Genomic relationships between all animals analyzed were calculated by Yang's method (Yang et al., 2011), and the heatmap was presented in [Figure S1](#). [Figure 1b](#) displays the decay of average LD ( $r^2$ ) at various physical distances between two loci on all the autosomes in Hu sheep. The average  $r^2$  at pair-wise SNP distance of <5 Mb on autosomes ranged from 0.43 to 0.58. In addition, in this population, the expected and observed heterozygosity are  $0.23 \pm 0.11$  and  $0.27 \pm 0.13$ , respectively. ROH-based inbreeding values were  $0.00898 \pm 0.00823$ .

### 3.3 | Genome-wide scan of selection signature detection

We employed three complementary methods to detect positive selection signatures within the Hu sheep population. The whole genome was split into a total of 9758 windows using the sliding window approach. [Figure 2](#) visualized the genome-wide distribution of the maximum CLR value per window. The average CLR value across the whole genome was 2.03, ranging from 0.00030 to 10.07. The highest CLR value (10.07) resided in OAR3: 105.5–106 Mb.



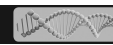
**FIGURE 1** Population structure analysis. (a) Principal component analysis. (b) The decay of  $r^2$  with pair-wise SNP marker distances in Hu sheep.

Ninety-seven windows surpassed the threshold value as candidate regions of positive selection among all windows in the Hu sheep genome (see [Table S1](#)).

[Figure 3](#) shows the distribution of  $|iHS|$  value per window along the Hu sheep genome. The mean  $|iHS|$  value of the whole genome was 0.68, with a range of 0.12 to 2.94. The highest value of mean  $|iHS|$  (2.94) occurred on OAR13: 51.25–51.75 Mb. Among all windows, 98 windows at the top 1% of the empirical distribution were selected as candidate regions (see [Table S2](#)).

Furthermore, we detected 9973 ROHs with a length of over 300 Kb among the 206 individuals. The number of detected ROHs varied among individuals (ranging from 2 to 356), with an average of 48.41 per individual. Among all these ROHs, the average length of ROH per individual was 22.00 Mb, accounting for 0.90% of the entire genome. The longest ROH was 3.25 Mb, which occurred on chromosome 21 and consisted of 362 SNPs, while the shortest ROH was 0.30 Mb, which occurred on chromosome 22 and consisted of 174 SNPs (see [Table 1](#)).

[Figure 4](#) displayed the percentage of SNPs occurrence in ROHs along the genome of Hu sheep. As seen in [Figure 4](#), ROH islands were mainly distributed on OARs 4, 6 and 13. ROH identified were unevenly distributed in the Hu sheep genome. Among all these ROHs, a total of 18 candidate genome regions were identified as



high-frequency ROH regions, representing 1497 SNPs with a range of 1 to 216 SNPs per region, and the total length of these regions was 5.17 Kb (see [Table S3](#)).

### 3.4 | Identification and functional annotation of genes under positive selection

[Figure 5](#) showed the distribution of genomic regions identified by three complementary statistics across the Ovine chromosomes. A total of 88.54 Mb genome regions were identified by at least one method and corresponding to 3.62% of the autosome in the Hu sheep genome. Among them, a total of 5.56 Mb genome regions were identified by both ROH and iHS tests; and a total of 0.5 Mb genome regions were identified by both iHS and CLR tests. However, there were no common genomic regions identified by both ROH and CLR tests ([Figure 5](#)).

Candidate regions for selection were defined as those genomic positions identified by multiple methods. Finally, 10 genome regions were retained for gene annotation, and the total length of these regions was 6.06 Mb ([Table 2](#)). After annotation, 93 unique genes were located in the selected genomic regions. We found 28 candidate genes (*CDKN2C*, *ACOXL*, *BCL2L11*, *HIBADH*, *HoxA* gene cluster (*HOXA1-5*, *HOXA9*, *HOXA10*, *HOXA11* and *HOXA13*), *H2AFZZ*, *DNAJB14*, *UNC5C*, *BMPR1B*, *STARD9*, *PSPC1*, *RNF17*, *ATP12A*, *AVP*, *OXT*, *MRPS26*, *GNRH2*, *OPRL1*, *TCEA2*, *NPBWR2*, *SOX18*, *ABHD16B*, *ZGPAT* and *STMN3*) associated with fertility localized on nine genomic regions (OARs 1, 3, 4, 7 and 13) ([Table 2](#)). Among of them, two important star genes (*BMPR1B* and *GNRHER*) and the *HOXA* gene cluster were associated with reproductive resided in the OARs 4, 6 and 13, respectively.

Moreover, we identified other genes that are related to hair follicle development (*LAMTOR3*, *EEF1A2*), ear size (*HOXA1*, *KCNQ2*), fat tail formation (*HOXA10*, *HOXA11*), growth and development (*FAF1*, *CCNDBP1*, *GJB2*, *GJA3*), fat deposition (*ACOXL*, *JAZF1*, *HOXA3*, *HOXA4*, *HOXA5*, *EBF4*), immune (*UBR1*, *FASTKD5*) and feed intake (*DAPP1*, *RNF17*, *NPBWR2*). Some of them were also enriched in the uterus development (see [Table S4](#)).

## 4 | DISCUSSION

In this study, we employed three complementary statistics to identify the selection signatures within Hu sheep. These test statistical approaches are on the basis of different principles. The CLR test is based on hypothesis testing that compares the neutral model for the site frequency spectrum of a genomic window with the selective sweep model. This test had higher detection efficiency for fixed or to-be-fixed selection signals (Williamson et al., 2007), and is extremely sensitive to detecting selection signatures across multiple sites in a single population rather than analyzing each SNP separately. Compared to CLR, the iHS method is used to measure how the haplotypes are unusual around a core SNP relative to

the whole genome. Hence, this method needs haplotype phasing, recombination map, genomic position, and ancestral and derived allelic information for each SNP. It should be mentioned that iHS method has a higher detection efficiency when the selected alleles are at moderate frequencies. Selecting for a beneficial mutation can emerge in a high-frequency haplotype with extended linkage disequilibrium (LD). Along with the rapid increase in the frequency of a haplotype carrying a beneficial mutation, ROHs can emerge and lead to a regional reduction in genetic variation up- and downstream of the target locus. In a population, ROHs always enrich one genomic region carrying the target locus to form an ROH island. Therefore, the origin of ROH islands is also considered the result of recent positive selection (Ceballos et al., 2018; Pemberton et al., 2012). Hence, ROH can be also used to identify selection signatures, since the individuals that have been subjected to the selection pressure will exhibit long ROH around the target locus. Although few overlapping genomic regions were simultaneously identified by the three methods, we integrated these three different methods and found the overlapping genomic regions with selection signatures to increase the power of detection.

The length of ROHs can reflect the inbreeding history for one population, because the possibility of ROH being interrupted due to recombination is increased with the increase in the number of generation. The shorter ROHs are generally produced by more ancient inbreeding, while the longer ROHs indicate more recent inbreeding (Ceballos et al., 2018). In this study, long ROHs were not obtained using the whole-genome sequencing data. The whole-genome sequencing data can obtain all SNPs on the whole genome compared to the SNP chip. It was reported that high-density chip data can detect shorter ROH fragments than low-density chip data, while low-density chips cannot detect most ROH fragments of 500 kb~1 Mb (Purfield et al., 2012). In the present study, the minimum length of ROH was set to 300 Kb to avoid the impact of strong LD between genes. The lengths of ROHs centralized on the 300 Kb to 1 Mb, which was shorter than using SNP chip data. These short ROHs may indicate that the inbreeding events of Hu sheep in recent generations have not happened, or this pattern would be an effect of admixture or crossbreeding, which may be contributed to the breakdown of long ROH segments on domestic lines.

As a valuable sheep germplasm resource, Hu sheep is a typical Chinese indigenous hyper-prolificacy sheep breed. Now, they are distributed across the whole country and are used as maternal lines in sheep crossbreeding programs. In this study, we identified two star genes and one gene cluster related to reproduction traits in sheep. One star genes is *BMPR1B* gene which is a paramount regulator of ovulation and litter size in sheep. *BMPR1B* had been documented to be involved in increasing the ovulation rate of ewes, and played an important role in follicle development (Liu et al., 2014). Moreover, this gene was previously detected associated to reproduction traits in the Hu sheep breed (Abied, Xu, et al., 2020; Liu et al., 2016; Wei et al., 2015; Zhang et al., 2017). The other star gene is *GNRH2* critical for reproduction. *GNRH2* gene encoded the gonadotropin-releasing hormone (GNRH) which triggered the



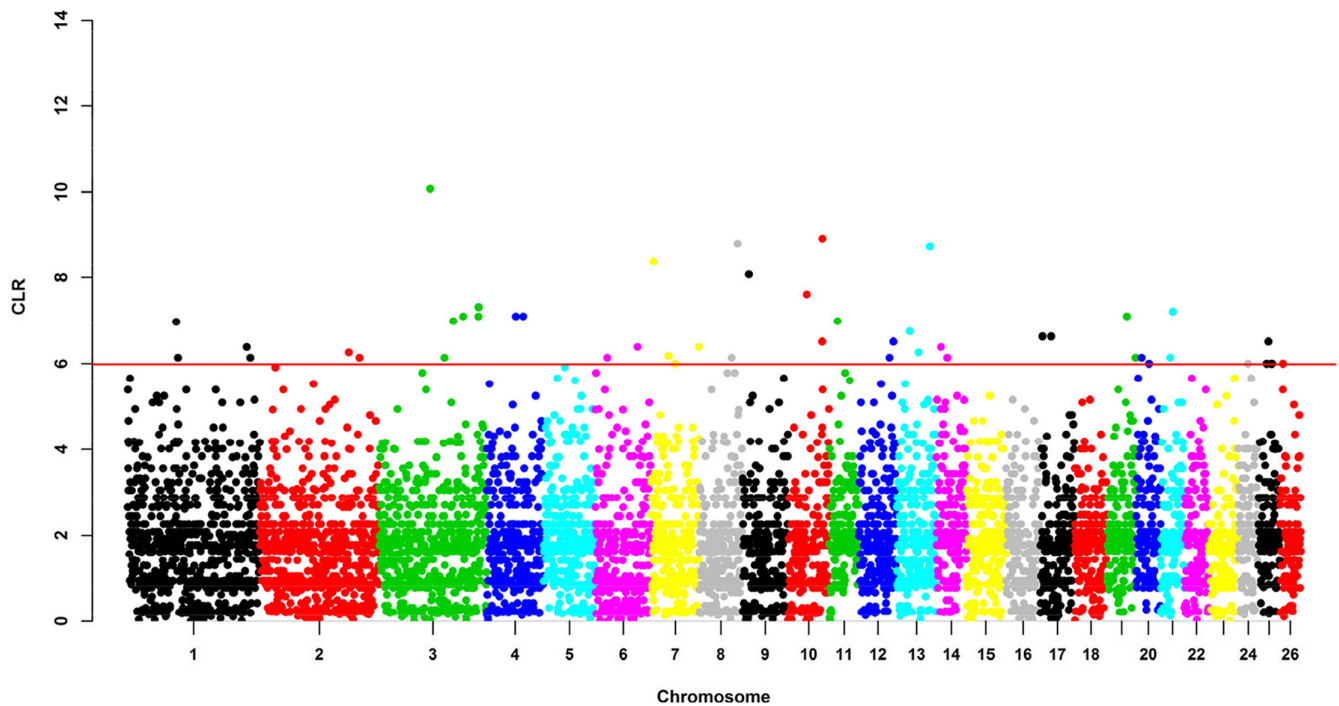


FIGURE 2 The distribution of selection signatures detected by CLR across the genome. The red line (5.98) was the threshold line of the candidate windows at the top 1% of the empirical distribution.

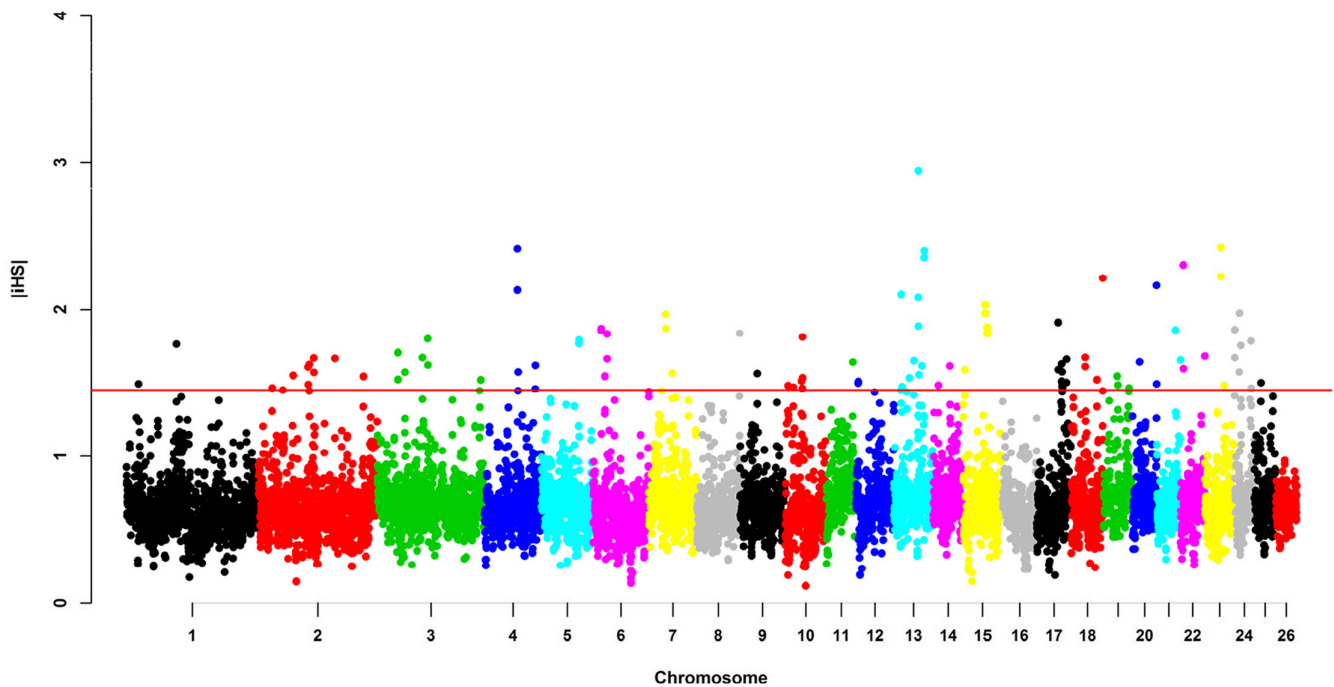


FIGURE 3 The distribution of selection signatures detected by iHS across the genome. The red line (1.45) was the threshold line of the candidate windows at the top 1% of the empirical distribution.

synthesis and secretion of gonadotropins and played an important role in regulating the entire hypothalamic–pituitary–gonadal axis (Desaulniers et al., 2017; Mijiddorj et al., 2017). MacColl et al. (2002) found that if the *GNRH* gene was mutated or its expression was suppressed, the secretion of various hormones in the hypothalamus–pituitary–gonadal axis would be blocked, which

could lead to infertility. One gene cluster also identified on OAR4: 68.25–69.00 Mb is the *HoxA* gene cluster, which included 9 *HoxA* genes (*HOXA1-5*, *HOXA9*, *HOXA10*, *HOXA11* and *HOXA13*). The *HOXA* gene cluster is an essential contributor to morphological differentiation of the female reproductive tract, cyclic endometrial changes, and embryo implantation (Du & Taylor, 2015). *HOXA3*

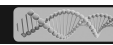


TABLE 1 Descriptive statistics for ROH in Hu sheep genome.

Sample size	SNP number	Average length (Mb)		Average number	
		Mean $\pm$ SE	Range	Mean $\pm$ SE	Range
206	174–362	22.00 $\pm$ 1.40	1.05–191.5	48.41 $\pm$ 2.67	2–356

Abbreviations: SE, standard error; SNP Number, the range of SNP involved in ROH.

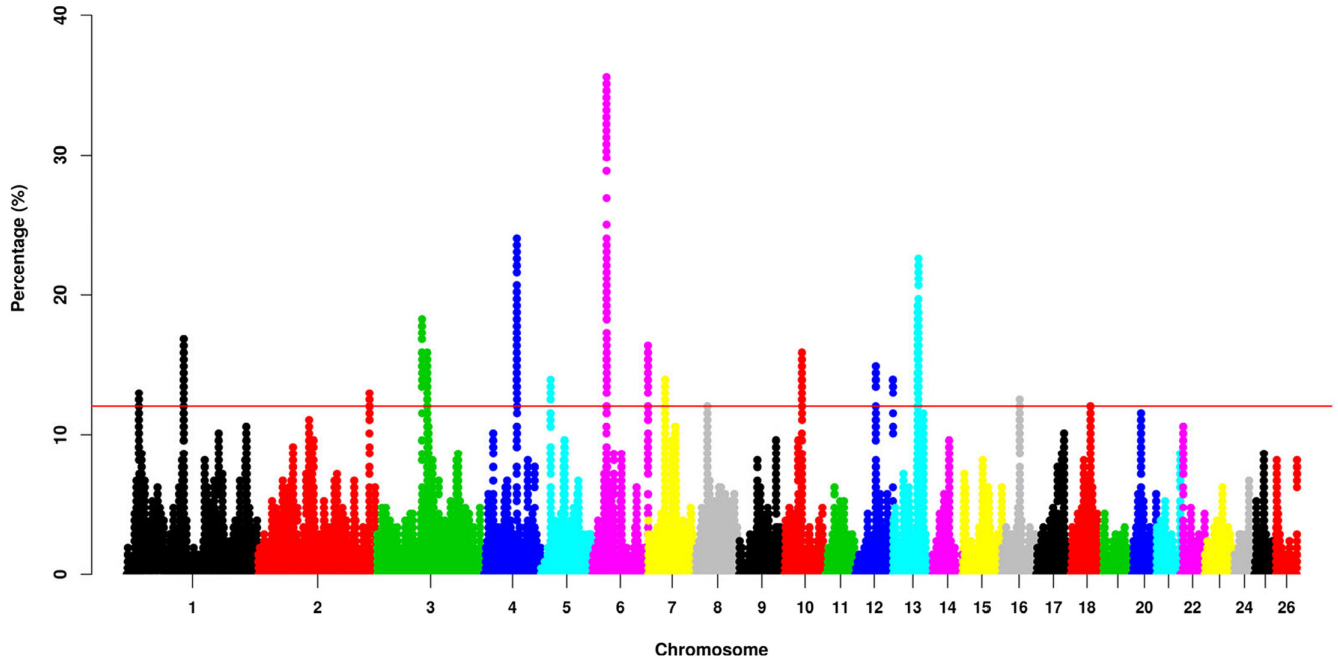
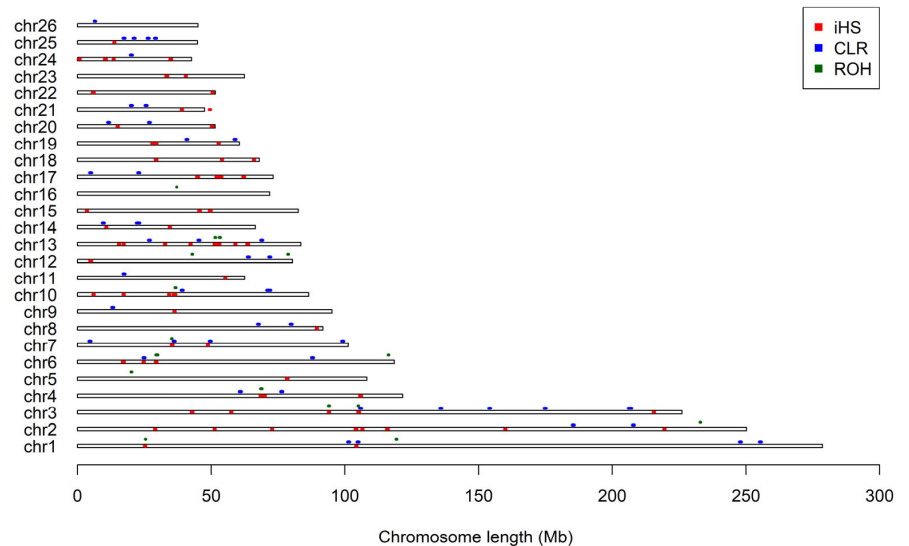


FIGURE 4 Manhattan plot of the occurrence (%) of SNPs in ROH across all individuals. The red line was the top 0.1% of SNPs observed in ROHs.

FIGURE 5 Distribution of genomic regions identified by three complementary test statistics across the Ovine chromosomes.



gene was reported to be expressed in the bovine oocytes and early-stage embryos and may influence oocyte maturation and the first stages of embryonic development (Paul et al., 2011). This gene also resided in ROH island in Chinese indigenous sheep breeds (Liu et al., 2021). *HOXA9*, *HOXA10*, *HOXA11*, and *HOXA13* are responsible for modulating female reproductive organs in mammalian

species. *HOXA9* modulates the development of the oviduct, both *HOXA10* and *HOXA11* are jointly culpable for the uterus formation, while *HOXA11* and *HOXA13* are responsible for the development of the anterior vagina and cervix (Ekanayake et al., 2022).

Other candidate genes related to reproduction were found as well. *AVP* may associate with the abnormal secretion of GNRH (Tan

Chr	Start	End	Methods	Candidate genes
OAR1	25000001	25500001	iHS, ROH	<i>FAF1</i> , <i>CDKN2C</i> , <i>TRNAY-GUA</i> , <i>TTC39A</i> , <i>C1H1orf185</i> , <i>RNF11</i>
OAR3	93750001	94250001	iHS, ROH	<i>EXOC6B</i>
OAR3	104750001	105500001	iHS, ROH	<i>BCL2L11</i> , <i>ACOXL</i>
OAR4	68250001	69002893	iHS, ROH	<i>HIBADH</i> , <i>TAX1BP1</i> , <i>JAZF1</i> , <i>EVX1</i> , <i>HOXA1</i> , <i>HOXA2</i> , <i>HOXA3</i> , <i>HOXA4</i> , <i>HOXA5</i> , <i>HOXA9</i> , <i>HOXA10</i> , <i>HOXA11</i> , <i>HOXA13</i>
OAR6	24500001	25000001	iHS, CLR	<i>TRNAC-GCA</i> , <i>H2AFZ</i> , <i>LAMTOR3</i> , <i>DDIT4L</i> , <i>DAPP1</i> , <i>DNAJB14</i>
OAR6	29250001	30022764	iHS, ROH	<i>UNC5C</i> , <i>BMP1R1B</i> , <i>PDLIM5</i>
OAR7	35000001	35500001	iHS, ROH	<i>CDAN1</i> , <i>CCNDBP1</i> , <i>EPB42</i> , <i>STARD9</i> , <i>TTBK2</i> , <i>UBR1</i> , <i>TMEM62</i>
OAR10	36250001	36831220	iHS, ROH	<i>GJA3</i> , <i>TRNAG-UCC</i> , <i>GJB2</i> , <i>ZMYM2</i> , <i>PSPC1</i> , <i>PARP4</i> , <i>CENPJ</i> , <i>MPHOSPH8</i> , <i>RNF17</i> , <i>ATP12A</i>
OAR13	51250001	51750001	iHS, ROH	<i>AVP</i> , <i>OXT</i> , <i>MRPS26</i> , <i>GNRH2</i> , <i>TMEM239</i> , <i>C13H20orf141</i> , <i>UBOX5</i> , <i>FASTKD5</i> , <i>VPS16</i> , <i>PCED1A</i> , <i>CPXM1</i> , <i>EBF4</i> , <i>PTPRA</i>
OAR13	52750001	53456146	iHS, ROH	<i>TRNAC-GCA</i> , <i>GINS1</i> , <i>PCMTD2</i> , <i>MYT1</i> , <i>OPRL1</i> , <i>TCEA2</i> , <i>UCKL1</i> , <i>DNAJC5</i> , <i>TPD52L2</i> , <i>ZBTB46</i> , <i>PRPF6</i> , <i>NPBWR2</i> , <i>LKAAEAR1</i> , <i>RGS19</i> , <i>SOX18</i> , <i>SAMD10</i> , <i>ZNF512B</i> , <i>ABHD16B</i> , <i>LIME1</i> , <i>ARFRP1</i> , <i>TNFRSF6B</i> , <i>SRMS</i> , <i>PPDPF</i> , <i>EEF1A2</i> , <i>ZGPAT</i> , <i>RTEL1</i> , <i>STMN3</i> , <i>C13H20orf195</i> , <i>PTK6</i> , <i>KCNQ2</i> , <i>GMEB2</i>

Note: The genes in black and bold indicated associated with reproduction traits.

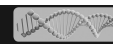
et al., 2021). *OXT* was related to litter size in Yunnan semi-fine wool sheep (Guo et al., 2022). *ZGPAT* may influence protein synthesis, processing and secretion so as to regulate milk production in yak (Xin et al., 2020). *ACOXL* gene was reported that played a pivotal role in fertility and precocity (Sbardella et al., 2021). Some candidate genes are reported with spermatogenesis. *RNF17* was essential for spermiogenesis (Pan et al., 2005) and *OPRL1* was exclusively expressed in the spermatocytes (Eto, 2015). *CDKN2C* gene played an important role in spermatogenesis in the adult testis (Xu et al., 2013) and *TCEA2* is expressed primarily in the testis in humans (Weaver & Kane, 1997). *SOX18* was more strongly expressed in spermatogonial cells and spermatocytes than in spermatids and testicular somatic cells of sheep (Yang et al., 2021). *ATP12A* played a role in acrosome reactions (Favia et al., 2022). *BCL2L11* affected sperm apoptosis (Ding et al., 2021). *HIBADH* was involved in the mitochondrial function of spermatozoa, and maintained sperm motility related to sperm motility (Tasi et al., 2013). *STARD9* was associated with asthenospermia (Mao et al., 2011). *PSPC1* regulated nuclear events during spermatogenesis (Myojin et al., 2004). *NPBWR2* played an essential role in hormone secretion in pigs (Yang et al., 2018). *DNAJB14* was related to heat shock protein in spermatozoa (Bansal et al., 2015).

*MRPS26* was consistently expressed throughout early embryogenesis with little stage or tissue specificity (Cheong et al., 2020). *H2AFZ* gene had been reported as candidate gene for boar semen quality traits (Zhao et al., 2019). *UNC5C* exerted significant effects on the conception rate in Holsteins (Sugimoto et al., 2015). *ABHD16B* was involved in bovine conception rate and sperm plasma membrane lipid composition (Shan et al., 2020). *STMN3* was identified to be potentially related to egg production in chicken (Chen et al., 2021). These results were helpful to explain the genetic mechanism of hyper-prolificacy trait in Hu sheep.

In addition to hyper prolificacy, Hu sheep have other breed-specific traits. Hu sheep are also well recognized for the beautiful wavy lambskins, and their fleeces consist of white coarse wool and heterotypical fibers. Moreover, they possess a short fat tail and are round with the tip upward-pointing. In this study, two genes (*LAMTOR3* and *EEF1A2*) identified were associated with hair follicle development. *LAMTOR3* gene was reported to be associated with the development of hair follicles and cashmere growth in Liaoning cashmere goats (Jin et al., 2019). *EEF1A2* gene has been demonstrated to be related to wool-related traits in sheep by a phenome-wide association study in fine-wooled Merino sheep (Zhao et al., 2021).

TABLE 2 Genes located in genomic regions showing evidence of selection simultaneously identified by two methods in Hu sheep.





Genes *HOXA10* and *HOXA11* identified in OAR4 may play important roles in fat tail formation. Both two genes were identified by selection signature detection in Chinese indigenous sheep breeds with different types of tails (Li et al., 2020; Liu et al., 2021; Yuan et al., 2017; Zhao et al., 2020) and *HOXA10* was further validated by RNA Seq as a candidate gene strongly linked with fat deposition in sheep tails (Bakhtiarzadeh & Alamouti, 2020). *HOXA1* and *KCNQ2* have been documented as being associated with ear size in sheep (Li et al., 2020; Qiao et al., 2015).

Hu sheep can also be used for producing mutton. Some of the candidate genes have been documented as important candidate genes for development and fat metabolism as well. *FAF1* gene was associated with dysfunctions of bone and joints in pigs and can lead to osteochondrosis (Rangkasenee et al., 2013). *CCNDBP1* was considered as a new positive regulator of skeletal myogenesis (Huang et al., 2016). *GJB2* and *GJA3* had been reported to be related to body size and developments, were identified by selection signature detection in sheep (Kim et al., 2016) and also resided in the ROH islands in Chinese sheep breeds (Abied, Xu, et al., 2020; Liu et al., 2021). *ACOXL* had documented association with fat metabolism (Oteru et al., 2013). *JAZF1* was shown to regulate lipid metabolism (Ming et al., 2014) and was also reported to be related to height in humans (Johansson et al., 2009). Moreover, this gene was identified by genome-wide selection signature detection in cattle (Zhao et al., 2015) and sheep (Zhao et al., 2020). Both *HOXA3* and *HOXA5* genes were involved in controlling the development of the subcutaneous abdominal adipose tissues (Karpe & Pinnick, 2015). *HOXA4* was considered as a transcriptional regulator of terminal brown fat cell differentiation (Pradhan et al., 2017), *EBF4* was identified to be associated with body fat traits in humans by a genome-wide methylation study (Cao et al., 2021).

In addition, we identified other candidate genes related to immune and feed intake. *UBR1* was potentially related to trypanotolerance, since the previous study had shown that E3-ubiquitin ligases played a decisive role in the immune response (Serrano et al., 2021). *FASTKD5* interacted with *NLRX1* to promote HIV1 replication (Guo et al., 2021). *DAPP1* has been reported to be associated with behavior and/or feed intake traits in pigs (Do et al., 2013), and *RNF17* has also been reported to be involved in feed intake in Pekin Ducks (Zhu et al., 2019). *NPBWR2* could inhibit the release of growth hormone and prolactin in chicken pituitary cells and played a significant role in the hypothalamic control of feed intake and hence energy homeostasis (Scanes, 2016).

## 5 | CONCLUSION

In this study, we applied three complementary test statistics to identify the selection signatures within Hu sheep using the whole genome sequencing data. A total of 10 significant genomic regions were simultaneously identified by multiple methods. These genomic regions harbored 92 candidate genes. Out of them, 32 genes associated with reproduction were resided in 9 genomic regions under

selection. Among them, two star genes (*BMPR1B* and *GNRH2*) and one *HOXA* gene cluster (*HOXA1-5*, *HOXA9*, *HOXA10*, *HOXA11* and *HOXA13*) have been documented to be associated with fertility. Moreover, we identified other genes that are related to hair follicle development, ear size, fat tail formation, growth and development, fat deposition, immune and feed intake. Our results could contribute to understanding the genetic mechanisms underlying the selection of breed-specific traits in Hu sheep, and provide a reference for genetic improvement program in Hu sheep.

## ACKNOWLEDGMENTS

We greatly appreciate three anonymous reviewers and the associated editor for their comments on an early version of the manuscript and their suggestions in revision of the manuscript.

## FUNDING INFORMATION

This research was funded by the Natural Science Foundations of China (No. 32172702), the National Key Research and Development Program of China (2021YFD1301101), and the Agricultural Science and Technology Innovation Program (ASTIP-IAS02).

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

Data available from the Dryad Digital Repository: <https://DOI:10.5061/dryad.rbnzs7hkv>.

## ORCID

Fuping Zhao <https://orcid.org/0000-0001-9582-5993>

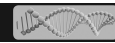
Lingzhao Fang <https://orcid.org/0000-0003-1103-3679>

Ruidong Xiang <https://orcid.org/0000-0002-1584-7605>

## REFERENCES

- Abied, A., Bagadi, A., Bordbar, F., Pu, Y., Augustino, S. M. A., Xue, X., Xing, F., Gebreselassie, G., Mwacharo, J., Ma, Y., & Zhao, Q. (2020). Genomic diversity, population structure, and signature of selection in five Chinese native sheep breeds adapted to extreme environments. *Genes (Basel)*, *11*, 494.
- Abied, A., Xu, L., Sahlou, B. W., Xing, F., Ahbara, A., Pu, Y., Lin, J., Berihulay, H., Islam, R., He, X., Mwacharo, J. M., Zhao, Q., & Ma, Y. (2020). Genome-wide analysis revealed homozygosity and demographic history of five Chinese sheep breeds adapted to different environments. *Genes*, *11*, 1480.
- Bakhtiarzadeh, M. R., & Alamouti, A. A. (2020). RNA-Seq based genetic variant discovery provides new insights into controlling fat deposition in the tail of sheep. *Scientific Reports*, *10*, 13525.
- Bansal, S. K., Gupta, N., Sankhwar, S. N., & Rajender, S. (2015). Differential genes expression between fertile and infertile spermatozoa revealed by transcriptome analysis. *PLoS One*, *10*, e0127007.
- Biscarini, F., Marras, G., Cozzi, P. A., & Gaspa, G. (2018). detectRUNS: An R package to detect runs of homozygosity and heterozygosity in diploid genomes.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120.
- Cao, V. T., Lea, R. A., Sutherland, H. G., Benton, M. C., Pishva, R. S., Haupt, L. M., & Griffiths, L. R. (2021). A genome-wide methylation study of

- body fat traits in the Norfolk Island isolate. *Nutrition, Metabolism, and Cardiovascular Diseases*, 31, 1556–1563.
- Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews. Genetics*, 19, 220–234.
- Chen, X., Sun, X., Chimbaka, I. M., Qin, N., Xu, X., Liswaniso, S., Xu, R., & Gonzalez, J. M. (2021). Transcriptome analysis of ovarian follicles reveals potential pivotal genes associated with increased and decreased rates of chicken egg production. *Frontiers in Genetics*, 12, 622751.
- Cheong, A., Lingutla, R., & Mager, J. (2020). Expression analysis of mammalian mitochondrial ribosomal protein genes. *Gene Expression Patterns*, 38, 119147.
- CNCAGR (China National Commission of Animal Genetic Resources). (2012). *Animal genetic resources in China – Sheep and goats*. China Agriculture Press.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Genomes Project Analysis G. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.
- DeGiorgio, M., Huber, C. D., Hubisz, M. J., Hellmann, I., & Nielsen, R. (2016). SweepFinder2: Increased sensitivity, robustness and flexibility. *Bioinformatics*, 32, 1895–1897.
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernysky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43, 491–498.
- Desaulniers, A. T., Cederberg, R. A., Lents, C. A., & White, B. R. (2017). Expression and role of gonadotropin-releasing hormone 2 and its receptor in mammals. *Front Endocrinol (Lausanne)*, 8, 269.
- Ding, Y., Ding, N., Zhang, Y., Xie, S., Huang, M., Ding, X., Dong, W., Zhang, Q., & Jiang, L. (2021). MicroRNA-222 transferred from semen extracellular vesicles inhibits sperm apoptosis by targeting BCL2L1. *Frontiers in Cell and Development Biology*, 9, 736864.
- Do, D. N., Strathe, A. B., Ostensen, T., Jensen, J., Mark, T., & Kadarmideen, H. N. (2013). Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping. *PLoS One*, 8, e71509.
- Du, H., & Taylor, H. S. (2015). The role of Hox genes in female reproductive tract development, adult function, and fertility. *Cold Spring Harbor Perspectives in Medicine*, 6, a023002.
- Ekanayake, D. L., Malopolska, M. M., Schwarz, T., Tuz, R., & Bartlewski, P. M. (2022). The roles and expression of HOXA/Hoxa10 gene: A prospective marker of mammalian female fertility? *Reproductive Biology*, 22, 100647.
- Eto, K. (2015). Nociceptin and meiosis during spermatogenesis in postnatal testes. *Vitamins and Hormones*, 97, 167–186.
- Favia, M., Gerbino, A., Notario, E., Tragni, V., Sgobba, M. N., Dell'Aquila, M. E., Pierri, C. L., Guerra, L., & Ciani, E. (2022). The non-gastric H(+)/K(+) ATPase (ATP12A) is expressed in mammalian spermatozoa. *International Journal of Molecular Sciences*, 23, 1048.
- Fay, J. C., & Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics*, 155, 1405–1413.
- Feng, X., Li, F., Wang, F., Zhang, G., Pang, J., Ren, C., Zhang, T., Yang, H., Wang, Z., & Zhang, Y. (2018). Genome-wide differential expression profiling of mRNAs and lncRNAs associated with prolificacy in Hu sheep. *Bioscience Reports*, 38, BSR20171350.
- Gautier, M., & Vitalis, R. (2012). Reh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, 28, 1176–1177.
- Geng, R. Q., Chang, H., Yang, Z. P., Sun, W., & Ren, Z. J. (2002). Study on origin and phylogeny status of Hu sheep. *Journal of Northwest Sci-Tech University of Agriculture and Forestry*, 16, 743–747.
- Guo, H., Wang, Q., Ghneim, K., Wang, L., Rampanelli, E., Holley-Guthrie, E., Cheng, L., Garrido, C., Margolis, D. M., Eller, L. A., Robb, M. L., Sekaly, R. P., Chen, X., Su, L., & Ting, J. P. (2021). Multi-omics analyses reveal that HIV-1 alters CD4(+) T cell immunometabolism to fuel virus replication. *Nature Immunology*, 22, 423–433.
- Guo, Y., Liang, J., Lv, C., Wang, Y., Wu, G., Ding, X., & Quan, G. (2022). Sequencing reveals population structure and selection signatures for reproductive traits in Yunnan semi-fine wool sheep (*Ovis aries*). *Frontiers in Genetics*, 13, 812753.
- Huang, Y., Chen, B., Ye, M., Liang, P., Zhangfang, Y., Huang, J., Liu, M., Songyang, Z., & Ma, W. (2016). Ccndbp1 is a new positive regulator of skeletal myogenesis. *Journal of Cell Science*, 129, 2767–2777.
- Jin, M., Cao, M., Chu, M., Piao, J., Guo, X., Zhao, F., & Piao, J. (2019). Gene expression and localization of LAMTOR3 in the skin cells of Liaoning cashmere goats. *Animal Biotechnology*, 30, 36–42.
- Johansson, A., Marroni, F., Hayward, C., Franklin, C. S., Kirichenko, A. V., Jonasson, I., Hicks, A. A., Vitart, V., Isaacs, A., Axenovich, T., Campbell, S., Dunlop, M. G., Floyd, J., Hastie, N., Hofman, A., Knott, S., Kolcic, I., Pichler, I., Polasek, O., ... Consortium, E. (2009). Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. *Human Molecular Genetics*, 18, 373–380.
- Karpe, F., & Pinnick, K. E. (2015). Biology of upper-body and lower-body adipose tissue—Link to whole-body phenotypes. *Nature Reviews. Endocrinology*, 11, 90–100.
- Kim, E. S., Elbeltagy, A. R., Aboul-Naga, A. M., Rischkowsky, B., Sayre, B., Mwacharo, J. M., & Rothschild, M. F. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity (Edinb)*, 116, 255–264.
- Lencz, T., Lambert, C., DeRosse, P., Burdick, K. E., Morgan, T. V., Kane, J. M., Kucherlapati, R., & Malhotra, A. K. (2007). Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 19942–19947.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078–2079.
- Li, X., Yang, J., Shen, M., Xie, X. L., Liu, G. J., Xu, Y. X., Lv, F. H., Yang, H., Yang, Y. L., Liu, C. B., Zhou, P., Wan, P. C., Zhang, Y. S., Gao, L., Yang, J. Q., Pi, W. H., Ren, Y. L., Shen, Z. Q., Wang, F., ... Li, M. H. (2020). Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. *Nature Communications*, 11, 2815.
- Liu, J., Shi, L., Li, Y., Chen, L., Garrick, D., Wang, L., & Zhao, F. (2021). Estimates of genomic inbreeding and identification of candidate regions that differ between Chinese indigenous sheep breeds. *Journal of Animal Science and Biotechnology*, 12, 95.
- Liu, Q., Pan, Z., Wang, X., Wenping, H. U., Ran, D. I., Yao, Y., & Chu, M. (2014). Progress on major genes for high fecundity in ewes. *Front Agr Sci Eng*, 1, 282–290.
- Liu, Z., Ji, Z., Wang, G., Chao, T., Hou, L., & Wang, J. (2016). Genome-wide analysis reveals signatures of selection for important traits in domestic sheep from different ecoregions. *BMC Genomics*, 17, 863.
- MacColl, G., Quinton, R., & Bouloux, P. M. (2002). GnRH neuronal development: Insights into hypogonadotropic hypogonadism. *Trends in Endocrinology and Metabolism*, 13, 112–118.
- Mao, X. M., Xing, R. W., Jing, X. W., Zhou, Q. Z., Yu, Q. F., Guo, W. B., Wu, X. Q., Chu, Q. J., & Feng, C. Q. (2011). Differentially expressed genes in asthenospermia: A bioinformatics-based study. *Zhonghua Nan Ke Xue*, 17, 694–698.
- McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., Macleod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright,



- A. F., ... Wilson, J. F. (2008). Runs of homozygosity in European populations. *American Journal of Human Genetics*, 83, 359–372.
- Mijiddorj, T., Kanasaki, H., Sukhbaatar, U., Oride, A., Hara, T., & Kyo, S. (2017). Mutual regulation by GnRH and kisspeptin of their receptor expression and its impact on the gene expression of gonadotropin subunits. *General and Comparative Endocrinology*, 246, 382–389.
- Ming, G. F., Xiao, D., Gong, W. J., Liu, H. X., Liu, J., Zhou, H. H., & Liu, Z. Q. (2014). JAZF1 can regulate the expression of lipid metabolic genes and inhibit lipid accumulation in adipocytes. *Biochemical and Biophysical Research Communications*, 445, 673–680.
- Myojin, R., Kuwahara, S., Yasaki, T., Matsunaga, T., Sakurai, T., Kimura, M., Uesugi, S., & Kurihara, Y. (2004). Expression and functional significance of mouse paraspeckle protein 1 on spermatogenesis. *Biology of Reproduction*, 71, 926–932.
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39, 197–218.
- Onteru, S. K., Gorbach, D. M., Young, J. M., Garrick, D. J., Dekkers, J. C., & Rothschild, M. F. (2013). Whole Genome association studies of residual feed intake and related traits in the pig. *PLoS One*, 8, e61756.
- Pan, J., Goodheart, M., Chuma, S., Nakatsuji, N., Page, D. C., & Wang, P. J. (2005). RNF17, a component of the mammalian germ cell nuage, is essential for spermiogenesis. *Development*, 132, 4029–4039.
- Paul, D., Bridoux, L., Rezsöhazi, R., & Donnay, I. (2011). HOX genes are expressed in bovine and mouse oocytes and early embryos. *Molecular Reproduction and Development*, 76, 436–449.
- Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *American Journal of Human Genetics*, 91, 275–292.
- Pradhan, R. N., Bues, J. J., Gardeux, V., Schwalie, P. C., Alpern, D., Chen, W., Russeil, J., Raghav, S. K., & Deplancke, B. (2017). Dissecting the brown adipogenic regulatory network using integrative genomics. *Scientific Reports*, 7, 42130.
- Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics*, 13, 70.
- Qanbari, S., & Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livestock Science*, 166, 133–143.
- Qiao, R., He, Y., Pan, B., Xiao, S., Zhang, X., Li, J., Zhang, Z., Hong, Y., Xing, Y., & Ren, J. (2015). Understanding the molecular mechanisms of human microtia via a pig model of HOXA1 syndrome. *Disease Models & Mechanisms*, 8, 611–622.
- Rangkasenee, N., Murani, E., Brunner, R., Schellander, K., Cinar, M. U., Scholz, A. M., Luther, H., Hofer, A., Ponsuksili, S., & Wimmers, K. (2013). KRT8, FAF1 and PTH1R gene polymorphisms are associated with leg weakness traits in pigs. *Molecular Biology Reports*, 40, 2859–2866.
- Rubin, C. J., Zody, M. C., Eriksson, J., Meadows, J. R., Sherwood, E., Webster, M. T., Jiang, L., Ingman, M., Sharpe, T., Ka, S., Hallbook, F., Besnier, F., Carlborg, O., Bed'hom, B., Tixier-Boichard, M., Jensen, P., Siegel, P., Lindblad-Toh, K., & Andersson, L. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 464, 587–591.
- Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., Ackerman, H. C., Campbell, S. J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., & Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419, 832–837.
- Saravanan, K. A., Panigrahi, M., Kumar, H., Bhushan, B., & Mishra, B. P. (2020). Selection signatures in livestock genome: A review of concepts, approaches, and applications. *Livestock Science*, 241, 104257.
- Sbardella, A. P., Watanabe, R. N., da Costa, R. M., Bernardes, P. A., Braga, L. G., Baldi Rey, F. S., Lobo, R. B., & Munari, D. P. (2021). Genome-wide association study provides insights into important genes for reproductive traits in Nelore cattle. *Animals*, 11, 1386.
- Scanes, C. G. (2016). Opening a new door: Neuropeptide W (NPW) is a novel inhibitory secretagogue for GH and prolactin acting via the Gi protein-coupled NPBWR2. *Endocrinology*, 157, 3394–3397.
- Serranito, B., Taurisson-Mouret, D., Harkat, S., Laoun, A., Ouchene-Khelifi, N. A., Pompanon, F., Benjelloun, B., Cecchi, G., Thevenon, S., Lenstra, J. A., & Da Silva, A. (2021). Search for selection signatures related to Trypanosomosis tolerance in African goats. *Frontiers in Genetics*, 12, 715732.
- Shan, S., Xu, F., Bleyer, M., Becker, S., Melbaum, T., Wemheuer, W., Hirschfeld, M., Wacker, C., Zhao, S., Schutz, E., & Brenig, B. (2020). Association of alpha/beta-hydrolase D16B with bovine conception rate and sperm plasma membrane lipid composition. *International Journal of Molecular Sciences*, 21, 627.
- Shi, L., Wang, L., Liu, J., Deng, T., Yan, H., Zhang, L., Liu, X., Gao, H., Hou, X., Wang, L., & Zhao, F. (2020). Estimation of inbreeding and identification of regions under heavy selection based on runs of homozygosity in a large White pig population. *Journal of Animal Science and Biotechnology*, 11, 46.
- Sugimoto, M., Gotoh, Y., Kawahara, T., & Sugimoto, Y. (2015). Molecular effects of polymorphism in the 3'UTR of Unc-5 homolog C associated with conception rate in Holsteins. *PLoS One*, 10, e0131283.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.
- Tan, S., Zhou, Y., Zhao, H., Wu, J., Yu, H., Yang, Y., Yang, Y., Zhao, H., & Li, H. (2021). Comprehensive transcriptome analysis of hypothalamus reveals genes associated with disorders of sex development in pigs. *The Journal of Steroid Biochemistry and Molecular Biology*, 210, 105875.
- Tasi, Y. C., Chao, H. C., Chung, C. L., Liu, X. Y., Lin, Y. M., Liao, P. C., Pan, H. A., Chiang, H. S., Kuo, P. L., & Lin, Y. H. (2013). Characterization of 3-hydroxyisobutyrate dehydrogenase, HIBADH, as a sperm-motility marker. *Journal of Assisted Reproduction and Genetics*, 30, 505–512.
- Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, 4, e72.
- Weaver, Z. A., & Kane, C. M. (1997). Genomic characterization of a testis-specific TFIIIS (TCEA2) gene. *Genomics*, 46, 516–519.
- Wei, C., Wang, H., Liu, G., Wu, M., Cao, J., Liu, Z., Liu, R., Zhao, F., Zhang, L., Lu, J., Liu, C., & Du, L. (2015). Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC Genomics*, 16, 194.
- Williamson, S. H., Hubisz, M. J., Clark, A. G., Payseur, B. A., Bustamante, C. D., & Nielsen, R. (2007). Localizing recent adaptive evolution in the human genome. *PLoS Genetics*, 3, e90.
- Xin, J., Chai, Z., Zhang, C., Zhang, Q., Zhu, Y., Cao, H., Yangji, C., Chen, X., Jiang, H., Zhong, J., & Ji, Q. (2020). Methylome and transcriptome profiles in three yak tissues revealed that DNA methylation and the transcription factor ZGPAT co-regulate milk production. *BMC Genomics*, 21, 731.
- Xu, S., Xia, W., Zohar, Y., & Gui, J. F. (2013). Zebrafish *dmrta2* regulates the expression of *cdkn2c* in spermatogenesis in the adult testis. *Biology of Reproduction*, 88, 14.
- Yang, H., Ma, J., Wan, Z., Wang, Q., Wang, Z., Zhao, J., Wang, F., & Zhang, Y. (2021). Characterization of sheep spermatogenesis through single-cell RNA sequencing. *The FASEB Journal*, 35, e21187.
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for Genome-wide complex trait analysis. *American Journal of Human Genetics*, 88, 76–82.
- Yang, S., Ma, Z., Suo, C., Cheng, L., Su, J., & Lei, Z. (2018). Cloning and mRNA expression of NPB and its effect on hormone secretion of the reproductive cells in the pig. *General and Comparative Endocrinology*, 261, 97–103.
- Yuan, Z., Liu, E., Liu, Z., Kijas, J. W., Zhu, C., Hu, S., Ma, X., Zhang, L., Du, L., Wang, H., & Wei, C. (2017). Selection signature analysis reveals genes associated with tail type in Chinese indigenous sheep. *Animal Genetics*, 48, 55–66.

- Zhang, Y., Li, F., Feng, X., Yang, H., Zhu, A., Pang, J., Han, L., Zhang, T., Yao, X., & Wang, F. (2017). Genome-wide analysis of DNA methylation profiles on sheep ovaries associated with prolificacy using whole-genome bisulfite sequencing. *BMC Genomics*, *18*, 759.
- Zhao, B., Luo, H., He, J., Huang, X., Chen, S., Fu, X., Zeng, W., Tian, Y., Liu, S., Li, C. J., Liu, G. E., Fang, L., Zhang, S., & Tian, K. (2021). Comprehensive transcriptome and methylome analysis delineates the biological basis of hair follicle development and wool-related traits in merino sheep. *BMC Biology*, *19*, 197.
- Zhao, F., Deng, T., Shi, L., Wang, W., Zhang, Q., Du, L., & Wang, L. (2020). Genomic scan for selection signature reveals fat deposition in Chinese indigenous sheep with extreme tail types. *Animals*, *10*, 773.
- Zhao, F., McParland, S., Kearney, F., Du, L., & Berry, D. P. (2015). Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics, Selection, Evolution*, *47*, 49.
- Zhao, F., Wei, C., Zhang, L., Liu, J., Wang, G., Zeng, T., & Du, L. (2016). A genome scan of recent positive selection signatures in three sheep populations. *Journal of Integrative Agriculture*, *15*, 162–174.
- Zhao, L., Ma, C., Song, H., Feng, Y., Li, Y., Xia, X., Li, J., & Li, F. (2019). H2AFZ, RNF4 and NR4A1 loci are associated with boar semen quality by population association studies. *Animal Biotechnology*, *30*, 311–316.
- Zhu, F., Cheng, S. R., Yang, Y. Z., Hao, J. P., Yang, F. X., & Hou, Z. C. (2019). Genome-wide association study of growth and feeding traits in Pekin ducks. *Frontiers in Genetics*, *10*, 702.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Zhao, F., Xie, R., Fang, L., Xiang, R., Yuan, Z., Liu, Y., & Wang, L. (2024). Analysis of 206 whole-genome resequencing reveals selection signatures associated with breed-specific traits in Hu sheep. *Evolutionary Applications*, *17*, e13697. <https://doi.org/10.1111/eva.13697>