

Whole-Genome Resequencing of Xiangxi Cattle Identifies Genomic Diversity and Selection Signatures

Xiaoyu Luo^{1,2†}, Jianbo Li^{1,3†}, Chentong Xiao², Luyang Sun², Weixuan Xiang^{2,4}, Ningbo Chen², Chuzhao Lei², Hong Lei^{1,3}, Yun Long^{3,5}, Ting Long^{3,5}, Quji Suolang⁶ and Kangle Yi^{1,3*}

¹Hunan Institute of Animal and Veterinary Science, Changsha, China, ²Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Xianyang, China, ³Xiangxi Cattle Engineering Technology Center of Hunan Province, Huayuan, China, ⁴School of Life Science, University of Bristol, Bristol, United Kingdom, ⁵Hunan De Nong Animal Husbandry Group Co. Ltd., Huayuan, China, ⁶Institute of Animal Science, Tibet Academy of Agricultural and Animal Husbandry Science, Lhasa, China

OPEN ACCESS

Edited by:

El Hamidi Hay, United States Department of Agriculture, United States

Reviewed by:

Xiangdong Ding, China Agricultural University, China Kwan-Suk Kim, Chungbuk National University, South Korea

*Correspondence:

Kangle Yi yikangle@yeah.net

[†]These authors have contributed equally to this work and share first authorship

Specialty section:

This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics

Received: 16 November 2021 Accepted: 13 April 2022 Published: 27 May 2022

Citation:

Luo X, Li J, Xiao C, Sun L, Xiang W, Chen N, Lei C, Lei H, Long Y, Long T, Suolang Q and Yi K (2022) Whole-Genome Resequencing of Xiangxi Cattle Identifies Genomic Diversity and Selection Signatures. Front. Genet. 13:816379. doi: 10.3389/fgene.2022.816379 Understanding the genetic diversity in Xiangxi cattle may facilitate our efforts toward further breeding programs. Here we compared 23 Xiangxi cattle with 78 published genomes of 6 worldwide representative breeds to characterize the genomic variations of Xiangxi cattle. Based on clustering models in population structure analysis, we displayed that Xiangxi cattle had a mutual genome ancestor with Chinese indicine, Indian indicine, and East Asian taurine. Population genetic diversity was analyzed by four methods (nucleotide diversity, inbreeding coefficient, linkage disequilibrium decay and runs of homozygosity), and we found that Xiangxi cattle had higher genomic diversity and weaker artificial selection than commercial breed cattle. Using four testing methods ($\theta\pi$, CLR, F_{ST} , and XP-EHH), we explored positive selection regions harboring genes in Xiangxi cattle, which were related to reproduction, growth, meat quality, heat tolerance, and immune response. Our findings revealed the extent of sequence variation in Xiangxi cattle at the genome-wide level. All of our fruitful results can bring about a valuable genomic resource for genetic studies and breed protection in the future.

Keywords: Xiangxi cattle, genetic diversity, selection signatures, DNAJC8, whole genome analysis

INTRODUCTION

One of the most economically important breeds of livestock is domestic cattle, providing human beings with basic resources, such as milk, beef, draft energy, and so on. They can be artificially classified into two subspecies: humpless taurine (*Bos taurus*) and humped indicine (*Bos indicus*) (Loftus et al., 1994), which are hybrid with each other. Recently, the research shows that domestic cattle can be categorized into five groups worldwide: European taurine, Eurasian taurine, East Asian taurine, Chinese indicine, and Indian indicine (Chen et al., 2018). As a significant part of cattle variety resources, Chinese native cattle have abundant genetic diversity. Xiangxi cattle, a representative breed of cattle in southwestern China, is one of the breeds recognized by China and renowned abroad, which belongs to the beef-serving type (**Supplementary Note**). Xiangxi cattle is highly valued because of its strong immunity, excellent meat quality, high service endurance and heat resistance capacity (Wang et al., 2012; Long and Ma, 2009; Zhang et al., 2011). Previous research on genetic diversity, population polymorphism, and hybridization adopted methods, for instance, autosomal, Y

1

chromosome or mitochondrial data (Cai et al., 2019; Luoreng et al., 2014; Yao et al., 2017) to demonstrate that the Xiangxi cattle originated from *Bos taurus* X *Bos indicus*.

With the application of whole-genome sequencing, it has become possible to study the genetic diversity of various breeds on the group size (Daetwyler et al., 2014). Since the implementation of the "The 1,000 Bull Genomes" project, whole-genome sequencing has been used to investigate the world's commercial breeds (Daetwyler et al., 2014; Stothard et al., 2011) and indigenous breeds (Kim et al., 2017; Lee et al., 2013), providing precious genome resources for future molecular breeding and genetic improvement. As a hot research methodology, the whole-genome sequencing has been progressively applied to Chinese native cattle such as Qinchuan cattle, Jiaxian Red cattle, and Yanbian cattle (Choi et al., 2015; Mei et al., 2018; Xia et al., 2021).

Considering that the entire genome variation of Xiangxi cattle was largely unexplored, we performed whole-genome sequencing on 23 Xiangxi cattle. SNPs were identified in Xiangxi cattle by basing on *Bos taurus* reference genome assembly (ARS-UCD1.2). Then SNPs of Xiangxi cattle made a comparison with those of commercial and native breeds previously collected from around the world. Finally, by scanning the whole genome of Xiangxi cattle, the selective sweeping results were determined. Our analyses fully described the genomic diversity and population structure, revealed the possible signs of natural and artificial selection, as well as provided new insights for breeding Xiangxi cattle.

METHODS

Samples, DNA Extraction, and Sequencing

Ear tissue samples of Xiangxi cattle (n = 23) were collected from the Xiangxi cattle engineering technology center of Hunan Province, Huayuan, China. Genomic DNA of the ear tissue samples was extracted using a standard phenol/chloroformbased protocol. The DNA library was constructed for each sample (500 bp insert size). Sequencing via Illumina NovaSeq 6000 with 2 × 150 bp model at Novogene Bioinformatics Institute, Beijing, China, and 150 bp paired-end sequence data were generated.

Furthermore, we obtained sequence data of 78 cattle, including 25 European cattle (Hereford-14, Jersey-11), 15 Hanwoo, 31 Chinese native cattle (Wannan-5, Guangfeng-4, Ji'an-4, Leiqiong-3, Jinjiang-3, Qinchuan-12), and 7 India-Pakistan cattle (Tharparkar-1, Sahiwal-1, Hariana-1, Nelore-1, Gir-2 and unknown-1) publicly. In total, 101 whole genomes of cattle were used for the subsequent analysis.

Reads Mapping and SNP Calling

The clean reads were mapped onto the *Bos taurus* reference genome assembly ARS-UCD1.2 using BWA-MEM (0.7.13r1126) with default parameters (Li and Durbin, 2009). After mapping, the single nucleotide polymorphisms (SNPs) were detected by using Samtools (Li et al., 2009), Picard tools (http://broadinstitute.github.io/picard), and Genome Analysis Toolkit (GATK, version 3.6–0-g89b7209). All SNPs were filtered using the module "Variant Filtration" of GATK to obtain high-quality SNPs. By using the ANNOVAR(Wang et al., 2010), SNPs were annotated based on the latest reference assembly (ARS-UCD1.2).

Population Genetics Analysis

After pruning in PLINK with the parameter (--indep-pair-wise 50 5 0.2), a set of SNPs were generated for the following analyses. An unrooted neighbor-joining (NJ) was constructed based on the matrix of pairwise genetic distances using MEGA v7.0 (Kumar et al., 2016) and iTOL v5 (Letunic and Bork, 2021). Principal component analysis (PCA) was performed using the smartPCA of the EIGENSOFT v5.0 package (Patterson et al., 2006). Population structure analysis was assessed with genetic clusters K ranging from 2 to 7 using the ADMIXTURE v1.3 (Alexander and Large, 2011).

Linkage disequilibrium (LD) decay was calculated using PopLDdecay (Zhang et al., 2019) with default parameters. By using the VCFtools (Danecek et al., 2011), we respectively calculated the inbreeding coefficient (--het) and the Nucleotide diversity (π). Additionally, we identified the runs of homozygosity (ROHs) using the--homozyg option implemented in the PLINK. The number and length of ROH for each breed were calculated and divided into four categories (0.5–1 Mb, 1–2 Mb, 2–4 Mb, >4 Mb). The plot as mentioned above was depicted using the R script (http://www. r-project.org).

Selective Sweep Test

To identify the selective sweep regions in Xiangxi cattle, two different statistics were used, including the nucleotide diversity ($\theta\pi$) with 50 kb sliding window and 20 kb step in VCFtools (Danecek et al., 2011) and the composite likelihood ratio (CLR) with 50 kb windows in SweepFinder2 (DeGiorgio et al., 2016). We took the top 1% of the overlapping genes obtained by the two above methods.

For comparison in Xiangxi cattle and Qinchuan cattle, we separately calculated fixation index ($F_{\rm ST}$) and cross-population extended haplotype homozygosity (XP-EHH). We estimated the $F_{\rm ST}$ values using VCFtools (50 kb windows with 20 kb step) and the XP-EHH using selscan v1.1 (Szpiech et al., 2014) with default settings. It is noteworthy that the normalized XP-EHH score was tested about each 50-kb region in the XP-EHH statistic. Significant genomic regions were identified (p < 0.005). In order to make the results more reliable, we used two or more methods to determine that the overlapping regions had been used as candidate regions. We calculated the Tajima's D statistic to consolidate our results by using VCFtools.

Moreover, KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/) was used to understand the function and complex pathways of candidate genes, including the Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Gene Ontology (GO) in the study (corrected *p*-value < 0.05).

RESULTS

Sequencing and Identification of Single Nucleotide Polymorphisms

Using BWA-MEM software, the clean reads were aligned with the *B. taurus* reference genome (ARS-UCD1.2), resulting in a total of



~4.7 billion reads and ~10.9 × coverage each (**Supplementary Table S1**). Xiangxi cattle were jointly analyzed with the public genomic data of five "core" populations (Chen et al., 2018; Xia et al., 2021)and Qinchuan cattle (native cattle in North-central China). Five "core" populations: European taurine (Hereford), Eurasian taurine (Jersey), East Asian taurine (Hanwoo), Chinese indicine (Leiqiong, Guangfeng, Ji'an, Jinjiang and Wannan), and Indian indicine (Tharparkar, Nelore, Sahiwal, Hariana and Gir) (**Supplementary Table S2**).

Xiangxi cattle have been discovered with 40573969 biallelic SNPs. The functional annotation of the polymorphic loci displayed that the vast majority of SNPs existed in the intergenic region (60.1%) or intron region (38.2%). Exon accounted for 0.7% of the total SNPs, including 112572 nonsynonymous and 179112 synonymous SNPs (Supplementary Figure S1). In addition, 0.9% of the SNPs were detected in an untranslated region (UTR), and the remaining 0.1% in splice. Xiangxi cattle was found to have the largest number of SNPs which was consistent with Chinese indicine on the whole, whereas Qinchuan was only second to them (Supplementary Table S3).

Population Structure and Characterization

Based on the study of autosomal SNPs, we discussed the phylogenetic relationships among 101 samples from 7 populations (**Figure 1**). The NJ (neighbor-joining) tree showed

that all "core" cattle populations form separate clusters. Qinchuan cattle are located between taurine cattle and indicine cattle (Mei et al., 2018), while Xiangxi cattle are aggregated near Chinese indicine and Indian indicine (**Figure 1A**). Principal component analysis (PCA) provides similar results to the above conclusions (**Figure 1B**).

We estimated the ancestral populations of all cattle samples by using clustering models. When K = 2, the lineage of cattle can be fundamentally distinguished from that of taurine and indicine cattle. When K = 5, five "core" cattle herds were separated, while Xiangxi cattle exhibited mixed phenomena (**Figure 1C**). The result displayed that it had a mutual genome ancestor with Chinese indicine, Indian indicine, and East Asian taurine. Moreover, a more dramatic genetic influence of Chinese indicine than the others was shown in the results.

Population Genetic Diversity

As shown in **Figure 2A**, the nucleotide diversity of Xiangxi cattle (0.00339) was second only to Chinese indicine (0.00357) and close to Qinchuan cattle (0.00247) and Indian indicine (0.00238), while that of Eurasian taurine (0.00095) was the lowest. The results of the linkage disequilibrium analysis agreed with these findings. The lowest LD level was found at short distances in Xiangxi cattle, followed by Chinese indicine, and the Eurasian taurine (Jersey) showed a higher LD level (**Figure 2B**).



Meanwhile, the ROH of each breed was divided by length (0.5-1 Mb, 1-2 Mb, 2-4 Mb and >4 Mb). In order of the number of ROH, it demonstrated that the European taurine took the largest number. By contrast, the number of the Indian indicine is lowest, that of the Chinese Crossbreed and Chinese indicine are in the middle respectively (**Figure 2C**). The Xiangxi cattle we were concerned with showed more extensive amounts of short/ medium ROH (0.5-2 Mb) but there were no long ROH (2-4 Mb). The number of ROH in European taurine cattle is much larger than in Xiangxi cattle due to high breeding. It is noteworthy that highly selective breeding may be at risk of inbreeding depression. As expected, the inbreeding coefficient also proved these findings (**Figure 2D**).

Genome-Wide Selective Sweep Test

For the sake of screening out the associated genetic variation, the nucleotide diversity analysis ($\theta\pi$) and the composite likelihood ratio (CLR) were used to explore the selected genomic regions in Xiangxi cattle (**Supplementary Tables S4, S5**) (**Figure 3**). We found that partial genome regions in Xiangxi cattle might have been selected during domestication. A total of 2225 ($\theta\pi$) and 182(CLR) genes were identified. Among them, 127 genes were overlapped which were considered to be candidate genes (**Supplementary Table S6**). The annotations of candidate genes revealed the functions that may be associated with economic traits, including reproduction (*KHDRBS2, GOLGA4, BBX, BANF2, INSL6*) (Cai et al., 2017; Guo et al., 2018; Ortega et al., 2016; Pini et al., 2020; Wang et al., 2016), growth

(*RALGAPA1*, *NCK1*, *VGLL2*, *FGFR3*) (Aryal et al., 2015; Honda et al., 2017; Sun et al., 2020; Wen et al., 2016), meat quality (*PLIN4*) (Pena et al., 2019), and somatotype (*SYN3*) (An et al., 2019). We also obtained genes (*ELANE*, *AZU1*) (Cui et al., 2021; Zhang et al., 2021) related to immune response, which may result from long-term natural selection.

Furthermore, we implemented two methods, F_{ST} (p < 0.005, $F_{\rm ST} \ge$ 0.29979) and XP-EHH (p < 0.005, XP-EHH \ge 2.27), to elucidate further the positive selection characteristics between Xiangxi cattle and Qinchuan cattle (Figure 4), which contained 1,264 and 553 hypothetical selection genes, respectively (Supplementary Tables S7, S8). 102 overlapping genes were also tested in both methods (Supplementary Table S9). Among them, regions containing known candidate genes related to heat tolerance (DNAJC8) (Li G. et al., 2020) showed intense differentiation signals. As for functional enrichment analysis (Figure 5), the KEGG pathway had only one significant pathway called "Regulation of actin cytoskeleton" (corrected p-value < 0.05) and 5 genes were performed (PIP4K2A, FGF22, DIAPH1, WASF2, SLC9A1), which were related to meat quality, growth efficiency, and the balance of ionic concentrations (Supplementary Table S10). GO terms were enriched, in particular including the terms of skin barrier and epidermal cell differentiation ("establishment of skin barrier, GO: 0061436" "positive regulation of epidermal cell differentiation, GO:0045606" "negative regulation of keratinocyte proliferation, GO:0010839"). These several genes (SFN, PALLD, KDF1, FA2H) were found in Xiangxi cattle which were associated with the



development of sebum, sebaceous glands, and epidermal cells. In addition, we also detected significant GO terms responsible for meat quality ("positive regulation of lipophagy, GO:1904504" "calmodulin binding, GO:0005516") (**Supplementary Table S11**) involving relevant genes (*ADRB2*, *SESN2*, *MYH10*, *ATP5IF1*, *SMTNL1*, *SLC9A1*).

Above all, among the above four applied methods, 23 candidate genes were involved in production and immunization traits (**Supplementary Table S12**). e.g., The *ADRB2* gene was associated with muscle pH (Murani et al., 2013). The *MYH10* gene was discovered to potentially affect the beef quality of the longissimus lumborum muscle following castration (Li et al., 2020b). *WASF2* gene affects cytoskeleton arrangement. The *EYA3* gene was revealed to have a significant correlation with muscle development (Xia et al., 2021). *MYBPC1* genes are well-known marbling-related genes, which were first identified in Japanese Black beef cattle (Li et al., 2020c). The *GPR3* gene was relative to fat development in mice (Godlewski et al., 2015). The *WDTC1* gene was involved in the regulation of fat-

related gene transcription (Ducos et al., 2017). PTAFR was also reported to be involved in inflammatory responses in cattle (Trovato et al., 2015; Vanvanhossou et al., 2020; Xerxa et al., 2016). *JAK2* mediates essential signaling events in innate and adaptive immunity (Lukashova et al., 2003).

DISCUSSION

In order to figure out the population structure and genetic diversity of Xiangxi cattle, we utilized the representative worldwide cattle dataset, making a comparison in a number of contexts and analyses as follows. Ancestor Component Analysis can reflect the communicating extent of genetic information. Xiangxi cattle was complex in the lineage composition which included Chinese indicine (~73%), East Asian taurine (~20%), and Indian indicine (~7%) (**Supplementary Figure S2**). The heterogeneity of Chinese cattle was induced by different taurine and indicine lineage proportions. Since Xiangxi cattle



was situated in the southern regions of China, its ancestry was affected by indicine more than taurine cattle. On the contrary, Qinchuan cattle was situated in North-central China. Using the principal component and neighbor-joining tree analytical methods, we observed an accurate division of seven populations. In particular, Xiangxi cattle were markedly separated into two branches (**Figure 1A**), which may be caused by the short-term hybridization of foreign commercial cattle, or the self-domesticated Indian zebu which have been migrated from the Indus Valley in the East and entered China from Yunnan Province about three thousand years ago. It may subsequently have an effect on southern Chinese cattle.

In the present study, the genomic variation parameters of the populations all had a similar trend. We ranked all breeds according to their nucleotide diversity (Chinese indicine > Xiangxi cattle > Qinchuan cattle > Indian indicine > taurine cattle), which was consistent with the previous study (Liu et al., 2020). By contrasting the two Chinese breeds, it was found that the genetic diversity of Xiangxi cattle exceeded the Qinchuan cattle, possibly because of the fact that Xiangxi cattle contained more Chinese indicine cattle lineage. The highest genetic diversity observed in Chinese indicine was induced by introgression of other bovines (*Bos javanicus*) (Chen et al., 2018). Similarly, the analysis of SNPs, inbreeding coefficient, and LD decay was highly concurrent with the abundant genetic diversity of Xiangxi cattle. The length and distribution density of ROH revealed that the degree of Xiangxi cattle was obviously less than taurine (**Figure 2**). Our results showed that Xiangxi cattle had more



genetic diversity than taurine cattle due to insufficient breeding. After a long period of intensive breeding, taurine cattle have formed a mature commercial system, whereas the breeding system of Xiangxi cattle was imperfect and has tremendous potential.

Xiangxi cattle is an advantageous resource for the development of animal husbandry because of its excellent characteristics, especially for its disease-resistance and environmental adaptability. To increase the assay efficiency and decrease false positives, four selective sweeping methods were performed on Xiangxi cattle. If the gene was conspicuously detected by at least two methods, then it will be served as an actual candidate gene. Xiangxi cattle is muscular, and have long been used as draft animals. According to $\theta\pi$ and CLR (Figure 3A), the overlapped candidate gene (VGLL2) played a vital role in the process of locomotion. The VGLL2 gene can affect the development of skeletal muscles (Gabriel et al., 2016; Honda et al., 2017,2019). Subsequently the positive selection region including the VGLL2 gene was intensively confirmed by the lower values of Tajima'D and nucleotide diversity analysis in Xiangxi cattle (Figures 3B,C).

The surroundings of Xiangxi cattle are hotter than Qinchuan cattle. In the long-term natural and artificial selection, Xiangxi cattle has adapted to the thermal climate in southern China. Using Qinchuan cattle as the reference population, we calculated F_{ST} and XP-EHH (**Figure 4A**). The *DNAJC8* gene, a member of DnaJ Heat Shock Protein Family, was observed to act as an important part of heat stress. Previous studies suggested that *DNAJC8* protected bees from heat stress by regulating heat-

inducible and antioxidant genes (Li G. et al., 2020). Heat stress triggers the production of reactive oxygen species (ROS) (Slimen et al., 2014). Previous studies have shown that DnaJ protein plays an important role in protecting antioxidant enzyme activity and removing excess ROS(Kong et al., 2014; Wang et al., 2014). In order to avoid false positives, the verification analysis was carried out on the gene region. *DNAJC8* was located on chromosome 2 of cattle, about 0.025 Mbp. This region showed extremely apparent differentiation and distinct haplotype patterns in two populations. Also, Tajima'D analysis exhibited that it was significantly lower in Xiangxi cattle (**Figure 4B**). These results provided forceful evidence that the region containing *DNAJC8* may have a crucial effect on heat resistance in Xiangxi cattle.

In addition, GO terms (e.g., keratinocyte development, establishment of skin barrier, positive regulation of epidermal cell differentiation) were enriched, involving epidermal cell related genes (e.g., *SFN*, *PALLD*, *KDF1*, *FA2H*) (Figure 5). The *SFN* gene can affect the expression of key matrix metalloproteinase (MMPs) in skin fibroblasts, thereby controlling the process of cutaneous wound healing (Kilani et al., 2008; Medina et al., 2007). *FA2H* is essential for the synthesis of HFA-sphingolipids in various organs which is highly expressed in the skin and crucial for the formation of the epidermal barrier (Maier et al., 2011). Xiangxi cattle strengthen the resistance to external toxins and prevent excessive loss of body fluids through the action of keratinocytes, thus adapting to the local heat and humid environment better. A KEGG pathway called "Regulation of

actin cytoskeleton" enriched genes (PIP4K2A, FGF22, DIAPH1, WASF2, SLC9A1) associated with the muscle contraction. SLC9A1 improves the adjustment capacity of muscle pH value and maintains body endurance (Iaia et al., 2008; Skovgaard et al., 2014). It was worth noting that the WASF2 gene, which was overlapped in the four selection methods, was a downstream effector molecule that was involved in the signal transduction from tyrosine kinase receptors and small GTPases to actin cytoskeleton and promoted actin filament formation (Miki et al., 1998). The gene can respond to changes of the external environment by reregulating its expression or distribution to further influence the cytoskeleton arrangement (Qian et al., 2009). Besides, WASF2 was also related to innate immunity, which may be attributed to the pleiotropy of the gene (Liu et al., 2021; Nolz et al., 2006). Compared to Qinchuan cattle, Xiangxi cattle conserved more WASF2. The haplotype patterns of the WASF2 gene were different between the two populations (Figure 4C). Therefore, it was speculated that the WASF2 gene was selected in Xiangxi cattle. In the past, due to the local geographical environment and harsh production conditions, Xiangxi cattle were mainly selected for draft, resulting in the characteristics of high endurance and strong assault draft. The above genes might be related to these characteristics.

In four selected methods ($\theta\pi$, CLR, F_{ST} , and XP-EHH), six genes (*ADRB2*, *WDTC1*, *MYBPC1*, *EYA3*, *MYH10*, *GPR3*) also showed positive selection, which may be related to superior meat quality traits in Xiangxi cattle. Two genes (*PTAFR*, *JAK2*) were related to the disease-resistance of Xiangxi cattle. *PTAFR* gene, platelet activating factor receptor, can regulate inflammatory, smooth-muscle contractile and hypotensive activity (Honda et al., 2002; Marathe et al., 1999). Additionally, *JAK2* is one of the tyrosine kinase family genes related to cytokine receptors, which has a certain correlation with *PTAFR* gene in immune response (Lukashova et al., 2003). Therefore, they were vital for the disease resistance of Xiangxi cattle.

In summary, Xiangxi Cattle is one of the famous local excellent breeds in China with superior traits (disease-resistance, environmental adaptability, meat quality). Our selective sweep identification of Xiangxi cattle may promote the cognition of the hereditary mechanism of potential population characteristics, and hopefully serve as a theoretical reference for the breeding in the future.

CONCLUSION

In this study, we used WGS data to study the population structure of Xiangxi cattle, so as to provide the first in-depth study of its gene diversity, phylogenetic relationship, and Selective Sweep

REFERENCES

Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE Algorithm for Individual Ancestry Estimation. *BMC Bioinforma*. 12, 246. [Journal Article; Research Support, N.I.H., Extramural]. doi:10.1186/1471-2105-12-246 Test. Some genes were identified that will not only be helpful to us for a better understanding of the features of Xiangxi cattle, but also further study the characteristics of other native cattle of China. In conclusion, the revelation of the genetic diversity of Xiangxi cattle will establish a sound foundation for conservation breeding in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi. nlm.nih.gov/genbank/, accession number: PRJNA779877.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee of Northwest A&F University.

AUTHOR CONTRIBUTIONS

XL and JL execution of this manuscript. CX, LS, WX, NC, and CL revised the manuscript. HL, YL, TL, and QS contributed to the sample collections. KY provided the laboratories for statistical analysis and the funding for the research.

FUNDING

This work was supported by the key research and development project of Hunan Province, China (No. 2020NK 2066), Innovation platform and talent plan of Hunan Province, China (No. 2019TP 2010) and the China Agriculture Research System of MOF and MARA (CARS-37).

ACKNOWLEDGMENTS

We thank the High-Performance Computing of Northwest A&F University for providing computing resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.816379/ full#supplementary-material

An, B., Xia, J., Chang, T., Wang, X., Xu, L., Zhang, L., et al. (2019). Genome-wide Association Study Reveals Candidate Genes Associated with Body Measurement Traits in Chinese Wagyu Beef Cattle. Anim. Genet. 50 (4), 386–390. doi:10.1111/age.12805

Aryal, A. C, S., Miyai, K., Izu, Y., Hayata, T., Notomi, T., Noda, M., et al. (2015). Nck Influences Preosteoblastic/osteoblastic Migration and Bone Mass. P. Natl. Acad. Sci. Usa. 112 (50), 15432–15437. doi:10.1073/pnas.1518253112

- Cai, C., Ning, Q., Chen, Q., Zhu, Y., Chen, N., Chen, H., et al. (2019). Genetic Diversity of Y-SNPs in Xiangxi Cattle (In Chinese). *China Cattle Sci.* 45 (05), 34–36.
- Cai, X., Yu, S., Mipam, T., Yang, F., Zhao, W., Liu, W., et al. (2017). Comparative Analysis of Testis Transcriptomes Associated with Male Infertility in Cattleyak. *Theriogenology.* 88, 28–42. doi:10.1016/j.theriogenology.2016.09.047
- Chen, N., Cai, Y., Chen, Q., Li, R., Wang, K., Huang, Y., et al. (2018). Wholegenome Resequencing Reveals World-wide Ancestry and Adaptive Introgression Events of Domesticated Cattle in East Asia. *Nat. Commun.* 9 (1), 2337. doi:10.1038/s41467-018-04737-0
- Choi, J. W., Choi, B. H., Lee, S. H., Lee, S. S., Kim, H. C., Yu, D., et al. (2015). Whole-Genome Resequencing Analysis of Hanwoo and Yanbian Cattle to Identify Genome-wide SNPs and Signatures of Selection. *Mol. Cells.* 38 (5), 466–473. [Journal Article.; Research Support, Non-U.S. Gov't]. doi:10.14348/ molcells.2015.0019
- Cui, C., Chakraborty, K., Tang, X. A., Zhou, G., Schoenfelt, K. Q., Becker, K. M., et al. (2021). Neutrophil Elastase Selectively Kills Cancer Cells and Attenuates Tumorigenesis. *Cell.* 184 (12), 3163–3177. doi:10.1016/j.cell.2021.04.016
- Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brøndum, R. F., et al. (2014). Whole-genome Sequencing of 234 Bulls Facilitates Mapping of Monogenic and Complex Traits in Cattle. *Nat. Genet.* 46 (8). doi:10.1038/ng.3034
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The Variant Call Format and VCFtools. *Bioinformatics*. 27 (15), 2156–2158. [Journal Article; Research Support, N.I.H., Extramural; Research Support, N.I.H., Intramural; Research Support, Non-U.S. Gov't]. doi:10.1093/ bioinformatics/btr330
- DeGiorgio, M., Huber, C. D., Hubisz, M. J., Hellmann, I., and Nielsen, R. (2016). SweepFinder2: Increased Sensitivity, Robustness and Flexibility. [Journal Article]. *Bioinformatics.* 32 (12), 1895–1897. doi:10.1093/bioinformatics/ btw051
- Ducos, E., Vergès, V., Dugé de Bernonville, T., Blanc, N., Giglioli-Guivarc'h, N., and Dutilleul, C. (2017). Remarkable Evolutionary Conservation of Antiobesity ADIPOSE/WDTC1 Homologs in Animals and Plants. *Genetics.* 207 (1), 153–162. doi:10.1534/genetics.116.198382
- Gabriel, B. M., Hamilton, D. L., Tremblay, A. M., and Wackerhage, H. (2016). The Hippo Signal Transduction Network for Exercise Physiologists. J. Appl. physiology. 1985120 (10), 1105–1117. doi:10.1152/japplphysiol.01076.2015
- Godlewski, G., Jourdan, T., Szanda, G., Tam, J., Cinar, R., Harvey-White, J., et al. (2015). Mice Lacking GPR3 Receptors Display Late-Onset Obese Phenotype Due to Impaired Thermogenic Function in Brown Adipose Tissue. *Sci. Rep.-UK.* 5, 14953. doi:10.1038/srep14953
- Guo, J., Tao, H., Li, P., Li, L., Zhong, T., Wang, L., et al. (2018). Whole-genome Sequencing Reveals Selection Signatures Associated with Important Traits in Six Goat Breeds. Sci. Rep.-UK. 8 (1), 10405. doi:10.1038/s41598-018-28719-w
- Honda, M., Hidaka, K., Fukada, S., Sugawa, R., Shirai, M., Ikawa, M., et al. (2017). Vestigial-like 2 Contributes to Normal Muscle Fiber Type Distribution in Mice. *Sci. Rep.-UK*. 7 (1), 7168. doi:10.1038/s41598-017-07149-0
- Honda, M., Tsuchimochi, H., Hitachi, K., and Ohno, S. (2019). Transcriptional Cofactor Vgll2 Is Required for Functional Adaptations of Skeletal Muscle Induced by Chronic Overload. *J. Cell. Physiol.* 234 (9), 15809–15824. doi:10. 1002/jcp.28239
- Honda, Z., Ishii, S., and Shimizu, T. (2002). Platelet-activating Factor Receptor. J. Biochem. 131 (6), 773–779. doi:10.1093/oxfordjournals.jbchem.a003164
- Iaia, F. M., Thomassen, M., Kolding, H., Gunnarsson, T., Wendell, J., Rostgaard, T., et al. (2008). Reduced Volume but Increased Training Intensity Elevates Muscle Na+-K+ Pump Alpha1-Subunit and NHE1 Expression as Well as Short-Term Work Capacity in Humans. [Journal Article; Research Support, Non-U.S. Gov't]. Am. J. Physiol. Regul. Integr. Comp. Physiol. 294 (3), R966–R974. doi:10.1152/ajpregu.00666.2007
- Kilani, R. T., Medina, A., Aitken, A., Jalili, R. B., Carr, M., and Ghahary, A. (2008). Identification of Different Isoforms of 14-3-3 Protein Family in Human Dermal and Epidermal Layers. *Mol. Cell. Biochem.* 314 (1-2), 161–169. doi:10.1007/ s11010-008-9777-6
- Kim, J., Hanotte, O., Mwai, O. A., Dessie, T., Bashir, S., Diallo, B., et al. (2017). The Genome Landscape of Indigenous African Cattle. [Journal Article; Research Support, Non-U.S. Gov't]. *Genome Biol.* 18 (1), 34. doi:10.1186/s13059-017-1153-y

- Kong, F., Deng, Y., Wang, G., Wang, J., Liang, X., and Meng, Q. (2014). LeCDJ1, a Chloroplast DnaJ Protein, Facilitates Heat Tolerance in Transgenic Tomatoes. *J. Integr. Plant Biol.* 56 (1), 63–74. doi:10.1111/jipb.12119
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. [Journal Article; Research Support, Non-U.S. Gov't]. doi:10.1093/ molbev/msw054
- Lee, K. T., Chung, W. H., Lee, S. Y., Choi, J. W., Kim, J., Lim, D., et al. (2013). Whole-genome Resequencing of Hanwoo (Korean Cattle) and Insight into Regions of Homozygosity. *BMC Genomics.* 14, 519. [Journal Article; Research Support, Non-U.S. Gov't]. doi:10.1186/1471-2164-14-519
- Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) V5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* 49 (W1), W293–W296. [Journal Article; Research Support, Non-U.S. Gov't]. doi:10. 1093/nar/gkab301
- Li, G., Zhao, H., Guo, H., Wang, Y., Cui, X., Li, H., et al. (2020a). Analyses of the Function of DnaJ Family Proteins Reveal an Underlying Regulatory Mechanism of Heat Tolerance in Honeybee. [Journal Article]. *Sci. Total Environ.* 716, 137036. doi:10.1016/j.scitotenv.2020.137036
- Li, H., and Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics*. 25 (14), 1754–1760. [Journal Article.; Research Support, Non-U.S. Gov't]. doi:10.1093/bioinformatics/ btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map Format and SAMtools. *Bioinformatics*. 25 (16), 2078–2079. [Journal Article; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't]. doi:10.1093/bioinformatics/btp352
- Li, Y., Cheng, G., Yamada, T., Liu, J., Zan, L., and Tong, B. (2020b). Effect of Expressions and SNPs of Candidate Genes on Intramuscular Fat Content in Qinchuan Cattle. Animals. 10 (8), 1370. doi:10.3390/ani10081370
- Li, Y., Wang, M., Li, Q., Gao, Y., Li, Q., Li, J., et al. (2020c). Transcriptome Profiling of Longissimus Lumborum in Holstein Bulls and Steers with Different Beef Qualities. *PLoS One*. 15 (6), e235218. doi:10.1371/journal.pone.0235218
- Liu, M., Zhang, J., Pinder, B. D., Liu, Q., Wang, D., Yao, H., et al. (2021). WAVE2 Suppresses mTOR Activation to Maintain T Cell Homeostasis and Prevent Autoimmunity. Sci. (New York, N.Y.). 371 (6536). doi:10.1126/science.aaz4544
- Liu, Y., Xu, L., Yang, L., Zhao, G., Li, J., Liu, D., et al. (2020). Discovery of Genomic Characteristics and Selection Signatures in Southern Chinese Local Cattle. [Journal Article]. Front. Genet. 11, 533052. doi:10.3389/fgene.2020.533052
- Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M., and Cunningham, P. (1994). Evidence for Two Independent Domestications of Cattle. *Comp. Study.* 91, 2757–2761. (Reprinted. doi:10.1073/pnas.91.7.2757
- Long, Z., and Ma, W. (2009). Characteristics and Utilization of Xiangxi Cattle Breeds (In Chinese). Mod. Agric. Sci. Technol. (16), 274–275.
- Lukashova, V., Chen, Z., Duhé, R. J., Rola-Pleszczynski, M., and Stanková, J. (2003). Janus Kinase 2 Activation by the Platelet-Activating Factor Receptor (PAFR): Roles of Tyk2 and PAFR C Terminus. J. Immunol. 1950 171 (7), 3794–3800. doi:10.4049/jimmunol.171.7.3794
- Luoreng, Z. M., Wang, X. P., Ma, Y., Li, F., Guo, D. S., Li, N., et al. (2014). Three Novel SNPs in the Coding Region of the Bovine MC3R Gene and Their Associations with Growth Traits. [Journal Article; Research Support, Non-U.S. Gov't]. *Biochem. Genet.* 52 (1-2), 116–124. doi:10. 1007/s10528-013-9632-7
- Maier, H., Meixner, M., Hartmann, D., Sandhoff, R., Wang-Eckhardt, L., Zöller, I., et al. (2011). Normal Fur Development and Sebum Production Depends on Fatty Acid 2-hydroxylase Expression in Sebaceous Glands. J. Biol. Chem. 286 (29), 25922–25934. doi:10.1074/jbc.M111.231977
- Marathe, G. K., Davies, S. S., Harrison, K. A., Silva, A. R., Murphy, R. C., Castro-Faria-Neto, H., et al. (1999). Inflammatory Platelet-Activating Factor-like Phospholipids in Oxidized Low Density Lipoproteins Are Fragmented Alkyl Phosphatidylcholines. J. Biol. Chem. 274 (40), 28395–28404. [Journal Article; Research Support, Non-U.S. Gov't; Research Support, U.S. Gov't, P.H.S.]. doi:10.1074/jbc.274.40.28395
- Medina, A., Ghaffari, A., Kilani, R. T., and Ghahary, A. (2007). The Role of Stratifin in Fibroblast-Keratinocyte Interaction. *Mol. Cell. Biochem.* 305 (1-2), 255–264. doi:10.1007/s11010-007-9538-y
- Mei, C., Wang, H., Liao, Q., Wang, L., Cheng, G., Wang, H., et al. (2018). Genetic Architecture and Selection of Chinese Cattle Revealed by Whole Genome

Resequencing. [Journal Article]. Mol. Biol. Evol. 35 (3), 688–699. doi:10.1093/ molbev/msx322

- Miki, H., Suetsugu, S., and Takenawa, T. (1998). WAVE, a Novel WASP-Family Protein Involved in Actin Reorganization Induced by Rac. *EMBO J.* 17 (23), 6932–6941. doi:10.1093/emboj/17.23.6932
- Murani, E., Ponsuksili, S., Reyer, H., Wittenburg, D., and Wimmers, K. (2013). Expression Variation of the Porcine ADRB2 Has a Complex Genetic Background. *Mol. Genet. Genomics.* 288 (11), 615–625. doi:10.1007/s00438-013-0776-2
- Nolz, J. C., Gomez, T. S., Zhu, P., Li, S., Medeiros, R. B., Shimizu, Y., et al. (2006). The WAVE2 Complex Regulates Actin Cytoskeletal Reorganization and CRAC-Mediated Calcium Entry during T Cell Activation. *Curr. Biol. Cb.* 16 (1), 24–34. doi:10.1016/j.cub.2005.11.036
- Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J., and Hansen, P. J. (2016). Use of Single Nucleotide Polymorphisms in Candidate Genes Associated with Daughter Pregnancy Rate for Prediction of Genetic Merit for Reproduction in Holstein Cows. Anim. Genet. 47 (3), 288–297. doi:10.1111/age.12420
- Patterson, N., Price, A. L., and Reich, D. (2006). Population Structure and Eigenanalysis, [Journal Article; Research Support. N.I.H., Extramural; Research Support, Non-U.S. Gov't]. *PLoS Genet.* 2 (12), e190. doi:10.1371/ journal.pgen.0020190
- Pena, R. N., Noguera, J. L., García-Santana, M. J., González, E., Tejeda, J. F., Ros-Freixedes, R., et al. (2019). Five Genomic Regions Have a Major Impact on Fat Composition in Iberian Pigs. *Sci. Rep.-UK.* 9 (1), 2031. doi:10.1038/s41598-019-38622-7
- Pini, T., Parks, J., Russ, J., Dzieciatkowska, M., Hansen, K. C., Schoolcraft, W. B., et al. (2020). Obesity Significantly Alters the Human Sperm Proteome, with Potential Implications for Fertility. J. Assist. Reprod. Gen. 37 (4), 777–787. doi:10.1007/s10815-020-01707-8
- Qian, A., Di, S., Gao, X., Zhang, W., Tian, Z., Li, J., et al. (2009). CDNA Microarray Reveals the Alterations of Cytoskeleton-Related Genes in Osteoblast under High Magneto-Gravitational Environment. Acta bioch. bioph. Sin. 41 (7), 561–577. doi:10.1093/abbs/gmp041
- Skovgaard, C., Christensen, P. M., Larsen, S., Andersen, T. R., Thomassen, M., and Bangsbo, J. (2014). Concurrent Speed Endurance and Resistance Training Improves Performance, Running Economy, and Muscle NHE1 in Moderately Trained Runners. J. Appl. Physiol. 117 (10), 1097–1109. [Journal Article; Randomized Controlled Trial; Research Support, Non-U.S. Gov't]. doi:10.1152/japplphysiol.01226.2013
- Slimen, I. B., Najar, T., Ghram, A., Dabbebi, H., Ben Mrad, M., and Abdrabbah, M. (2014). Reactive Oxygen Species, Heat Stress and Oxidative-Induced Mitochondrial Damage. A Review. Int. J. Hyperth. official J. Eur. Soc. Hyperthermic Oncol. North Am. Hyperth. Group. 30 (7), 513–523. doi:10. 3109/02656736.2014.971446
- Stothard, P., Choi, J. W., Basu, U., Sumner-Thomson, J. M., Meng, Y., Liao, X., et al. (2011). Whole Genome Resequencing of Black Angus and Holstein Cattle for SNP and CNV Discovery. *BMC Genomics.* 12, 559. [Comparative Study; Journal Article; Research Support, Non-U.S. Gov't]. doi:10.1186/1471-2164-12-559
- Sun, T., Huang, G., Wang, Z., Teng, S., Cao, Y., Sun, J., et al. (2020). Selection Signatures of Fuzhong Buffalo Based on Whole-Genome Sequences. *BMC Genomics*. 21 (1), 674. doi:10.1186/s12864-020-07095-8
- Szpiech, Z. A., and Hernandez, R. D. (2014). Selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection. *Mol. Biol. Evol.* 31 (10), 2824–2827. [Letter; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't]. doi:10.1093/molbev/msu211
- Trovato, A., Panelli, S., Strozzi, F., Cambulli, C., Barbieri, I., Martinelli, N., et al. (2015). Expression of Genes Involved in the T Cell Signalling Pathway in Circulating Immune Cells of Cattle 24 Months Following Oral Challenge with Bovine Amyloidotic Spongiform Encephalopathy (BASE). *BMC Vet. Res.* 11, 105. doi:10.1186/s12917-015-0412-y
- Vanvanhossou, S. F. U., Scheper, C., Dossa, L. H., Yin, T., Brügemann, K., and König, S. (2020). A Multi-Breed GWAS for Morphometric Traits in Four Beninese Indigenous Cattle Breeds Reveals Loci Associated with Conformation,

Carcass and Adaptive Traits. *BMC Genomics*. 21 (1), 783. doi:10.1186/s12864-020-07170-0

- Wang, C., Tang, M., Chang, W., Furushima, K., Jang, C., Behringer, R. R., et al. (2016). PiggyBac Transposon-Mediated Mutagenesis in Rats Reveals a Crucial Role of Bbx in Growth and Male Fertility. *Biol. Reprod.* 95 (3), 51. doi:10.1095/ biolreprod.116.141739
- Wang, G., Cai, G., Kong, F., Deng, Y., Ma, N., and Meng, Q. (2014). Overexpression of Tomato Chloroplast-Targeted DnaJ Protein Enhances Tolerance to Drought Stress and Resistance to Pseudomonas Solanacearum in Transgenic Tobacco. *Plant physiology Biochem. PPB.* 82, 95–104. doi:10.1016/j.plaphy.2014.05.011
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* 38 (16), e164. [Evaluation Study; Journal Article; Research Support, N.I.H., Extramural]. doi:10.1093/nar/gkq603
- Wang, X. P., Luoreng, Z. M., Li, F., Wang, J. R., Li, N., and Li, S. H. (2012). Genetic Polymorphisms of Lipoprotein Lipase Gene and Their Associations with Growth Traits in Xiangxi Cattle. [Journal Article; Research Support, Non-U.S. Gov't]. *Mol. Biol. Rep.* 39 (12), 10331–10338. doi:10.1007/s11033-012-1910-7
- Wen, X., Li, X., Tang, Y., Tang, J., Zhou, S., Xie, Y., et al. (2016). Chondrocyte FGFR3 Regulates Bone Mass by Inhibiting Osteogenesis. J. Biol. Chem. 291 (48), 24912–24921. doi:10.1074/jbc.m116.730093
- Xerxa, E., Barbisin, M., Chieppa, M. N., Krmac, H., Vallino Costassa, E., Vatta, P., et al. (2016). Whole Blood Gene Expression Profiling in Preclinical and Clinical Cattle Infected with Atypical Bovine Spongiform Encephalopathy. *PLoS One.* 11 (4), e153425. [Journal Article; Research Support, Non-U.S. Gov't]. doi:10. 1371/journal.pone.0153425
- Xia, X., Zhang, S., Zhang, H., Zhang, Z., Chen, N., Li, Z., et al. (2021). Assessing Genomic Diversity and Signatures of Selection in Jiaxian Red Cattle Using Whole-Genome Sequencing Data. *BMC Genomics*. 22 (1), 43. [Journal Article; Review]. doi:10.1186/s12864-020-07340-0
- Yao, Y., Jia, Y., Xia, X., Zhu, Y., Lan, X., Huang, Y., et al. (2017). Genetic Diversity of Mitochondrial DNA D-Loop Region in Xiangxi Cattle (In Chinese). *China Cattle Sci.* 43 (05), 11–13.
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., and Yang, T. L. (2019). PopLDdecay: A Fast and Effective Tool for Linkage Disequilibrium Decay Analysis Based on Variant Call Format Files. [Journal Article; Research Support, Non-U.S. Gov't]. *Bioinformatics.* 35 (10), 1786–1788. doi:10.1093/bioinformatics/bty875
- Zhang, R., Li, Y., and Xing, X. (2021). Comparative Antler Proteome of Sika Deer from Different Developmental Stages. Sci. Rep.-UK. 11 (1), 10484. doi:10.1038/ s41598-021-89829-6
- Zhang, Y., Xu, S., and Wang, Z. (2011). Animal Genetic Resources in China-bovines (In Chinese). Beijing: China Agriculture Press.

Conflict of Interest: Authors YL and TL is employed by Hunan De Nong Animal Husbandry Group Co. Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Luo, Li, Xiao, Sun, Xiang, Chen, Lei, Lei, Long, Long, Suolang and Yi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.