

RESEARCH

Open Access



Whole-genome analysis reveals distinct adaptation signatures to diverse environments in Chinese domestic pigs

Zhen Wang^{1,2†}, Bangmin Song^{1,3,4†}, Jianyu Yao⁵, Xingzheng Li^{1,2}, Yan Zhang⁶, Zhonglin Tang^{1,2,7*} and Guoqiang Yi^{1,2,7*} 

Abstract

Background Long-term natural and artificial selection has resulted in many genetic footprints within the genomes of pig breeds across distinct agroecological zones. Nevertheless, the mechanisms by which these signatures contribute to phenotypic diversity and facilitate environmental adaptation remain unclear.

Results Here, we leveraged whole-genome sequencing data from 82 individuals from 6 domestic pig breeds originating in tropical, high-altitude, and frigid regions. Population genetic analysis suggested that habitat isolation significantly shaped the genetic diversity and contributed to population stratification in local Chinese pig breeds. Analysis of selection signals revealed regions under selection for adaptation in tropical (55.5 Mb), high-altitude (43.6 Mb), and frigid (17.72 Mb) regions. The potential functions of the selective sweep regions were linked to certain complex traits that might play critical roles in different geographic environments, including fat coverage in frigid environments and blood indicators in tropical and high-altitude environments. Candidate genes under selection were significantly enriched in biological pathways involved in environmental adaptation. These pathways included blood circulation, protein degradation, and inflammation for adaptation to tropical environments; heart and lung development, hypoxia response, and DNA damage repair for high-altitude adaptation; and thermogenesis, cold-induced vasodilation (CIVD), and the cell cycle for adaptation to frigid environments. By examining the chromatin state of the selection signatures, we identified the lung and ileum as two candidate functional tissues for environmental adaptation. Finally, we identified a mutation (chr1: G246,175,129A) in the *cis*-regulatory region of *ABCA1* as a plausible promising variant for adaptation to tropical environments.

Conclusions In this study, we conducted a genome-wide exploration of the genetic mechanisms underlying the adaptability of local Chinese pig breeds to tropical, high-altitude, and frigid environments. Our findings shed light on the prominent role of *cis*-regulatory elements in environmental adaptation in pigs and may serve as a valuable biological model of human plateau-related disorders and cardiovascular diseases.

Keywords Environmental adaptation, Local Chinese breeds, Pig, Population genetics, Selection signals, Whole-genome resequencing

[†] Zhen Wang and Bangmin Song contributed equally to this work.

*Correspondence:

Zhonglin Tang
tangzhonglin@caas.cn
Guoqiang Yi
yiguoqiang@caas.cn

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



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Environmental stressors can indeed have negative impacts on the health, welfare, and production efficiency of domesticated animals [1]. It is well established that environmental stress can significantly influence genome diversity [2]. As a result, identifying genomic features involved in the response to environmental pressure has become a focus of evolutionary biology. In natural populations, adaptation is a dynamic and long-term evolutionary process whereby populations enhance their adaptation by accumulating beneficial alleles at gene loci that control adaptive phenotypes [3]. Compared to natural populations, domesticated animals have experienced a greatly accelerated process of environmental adaptation via evolution due to human migration and selection [4]. Therefore, domestic animals represent an excellent model for probing adaptive mutations in current genetic studies.

Approximately 10,000 years ago, the domestication of pigs occurred in multiple regions across Eurasia [5, 6]. Since then, pigs have undergone significant domestication changes and have become one of the most economically domestic animals globally [7]. Compared to modern commercial pig breeds that are under strong selective pressure, such as Duroc and Landrace, indigenous pig breeds generally exhibit superior and broader environmental adaptability. Given the current and future impacts of climate change on the planet, understanding the genetic adaptability of local pig breeds to different environments and breeding highly adaptable pig breeds that can cope with climate change has practical value in modern pig farming and energy conservation. Furthermore, pigs share a high degree of physiological similarity with humans, particularly in terms of ecology, metabolism, and immunity [8]. Thus, exploring the biological processes related to environmental adaptation in domestic pigs could provide valuable insights into treating diseases caused by environmental factors in humans.

Recently, several genetic variations associated to environmental adaptation traits have been reported in multiple farm animals, including chickens [9], horses [10], sheep [11], cattle [12], bactrian camels [13, 14], and pigs [15–18]. These studies have provided valuable insights into the environmental adaptability and evolution of domestic animals, yet there has been limited research on the relationship between noncoding regions and adaptability. Given the ability of noncoding regions to regulate gene expression in a spatially and temporally specific manner, comprehensive studies integrating multi-tissue and multi-omics analyses are required to gain a deeper understanding of environmental adaptations.

In this study, whole-genome sequencing data were collected from 82 individuals from 6 Chinese native pig

breeds according to our previous study. Based on these data, we conducted comprehensive analyses of the population structure and genome diversity of these pig breeds. Furthermore, we explored the potential relationships among swept genes, functional tissues, and candidate variants associated with adaptations to tropical, high-altitude, and frigid environments.

Materials and methods

Resequencing data collection

This study utilized a subset of resequencing data from project PRJNA754250 [19], comprising 82 individuals selected based on the environmental characteristics of their geographic origins. The study included Tibetan pigs from the Tibet plateau; Wuzhishan pigs, Ding'an pigs, and Tunchang pigs from Hainan Island; Hetao pigs from Inner Mongolia; and Min pigs from northeastern regions. Climate information associated to their habitats was retrieved from the study conducted by Zhao et al. [20]. We employed SOAPnuke v2.1.6 [21], a widely recognized tool for filtering raw reads, to remove adaptor and sequencing errors using the default parameters.

Reads mapping of whole-genome sequencing data to the reference

The reads were first trimmed by fastp v0.23.0 [22] with default parameters. Next, all clean reads, including our newly generated samples, were aligned to the Sscrofa11.1 reference genome using the BWA-MEM pipeline [23]. The mapped reads were then sorted, and duplicates were removed by Picard tools v2.26.0 (<https://broadinstitute.github.io/picard/>) and SAMtools v1.14 [24].

Genome-wide screening of SNPs and INDELS

The genome-wide variants were called for each sample by the GATK UnifiedGenotyper [25] with *-glm BOTH -rf BadCigar -sample_ploidy 2* option. To ensure high accuracy of variants calling, SNPs with $QD < 2.0$ || $FS > 60.0$ || $MQ < 40.0$ || $MQRankSum < -12.5$ || $ReadPosRankSum < -8.0$ were filtered. We then filtered out non-biallelic SNPs. After the quality screening, all the identified SNPs were further annotated using SnpEff v4.3t [26] based on the gene annotations of the pig reference genome Sscrofa11.1. Based on the genome annotation, SNPs were categorized as occurring in exonic regions, 5' or 3' untranslated regions, intronic regions, splicing sites (within 2 bp of a splicing junction), upstream and downstream regions (within a 1 kb region upstream or downstream from the transcription start site), or intergenic regions. SNPs in coding exons were further grouped as either synonymous SNPs or nonsynonymous SNPs. To check the confidence of SNPs called, we

compared the SNPs identified with the dbSNP (<https://www.ncbi.nlm.nih.gov/snp>, last accessed Feb 23, 2021).

Phylogenetic and population genetic analyses

To gain insights into phylogenetic relationships among pig breeds, we further filtered the SNP data set by applying criteria of minor allele frequency < 0.01, Hardy–Weinberg equilibrium < 0.001, and a proportion of missing genotypes > 10%. Subsequently, pruning was performed with the PLINK v1.90 [27] option "-indep-pairwise 50 5 0.2." Following filtering and linkage disequilibrium (LD) pruning, 1,358,458 SNPs were retained for subsequent population genetics analyses. A neighbor-joining (NJ) tree was constructed using the VCF2Dis v1.44 (<https://github.com/BGI-shenzhen/VCF2Dis>; accessed on 23 March 2022). The tree was displayed using the Interactive Tree Of Life (iTOL) [28]. To infer the population structure, we used ADMIXTURE v1.3.0 [29], which implements a block-relaxation algorithm. To identify the best K-value, the cross-validation error was tested for each K-value from 2 to 9. The principal component analysis (PCA) was conducted using the GCTA program [30]. The pattern of LD for these interest regions was computed using the PopLDdecay v3.40 [31].

Selective sweeps during pig domestication and breeding

SNPs with minor allele frequency below 1% were removed from this analysis (Fig. S1). Considering that the vast majority of LD block sizes fall within the 0–50 kb range (98.48%) in our data (Fig. S2), a sliding-window approach (50-kb windows sliding in 10-kb steps) was applied to calculate the average SNP F_{ST} values (the Fixation Index, a measure in population genetics used to quantify genetic differentiation or genetic distance between populations) [32] and polymorphism levels ($\theta\pi$, pairwise nucleotide variation as a measure of variability) [33, 34]. Windows containing ≥ 100 SNPs were used to detect signatures of selection sweeps. The F_{ST} was calculated using VCFtools v1.17 [35] with parameter "-weirfst-pop group1 -weirfst-pop group2 -fst-window-size 50,000 -fst-window-step 10,000". The $\theta\pi$ ratios were calculated using VCFtools v1.17 with parameters as follows: "-keep group1/group2 -window-pi 50,000 -window-pi-step 10,000". The XP-EHH (Cross Population Extended Haplotype Homozygosity) [36] was performed for every SNP using the default settings by selscan v2.0.0 [37], and genotypes were phased using Beagle 5.2 software [38] with default parameters. The test statistic was the average normalized XP-EHH score in each 50-kb region.

Those windows within the top 5% quantile of the two statistics were considered candidate selection targets and annotated using the genomic database search engine BioMart [39].

Enrichment analysis

Previous studies have been mostly focused on traits of high value in animal production, like meat and carcass quality in pigs, leading to their overrepresentation in the QTL database. To counter this bias, we used a bootstrap simulation via the R package GALLO v1.4 [40] for QTL enrichment analyses against the pig QTL Database [41], accepting only P -values under 0.05 from multiple tests. Further, employing GALLO v1.4, we pinpointed genes in these regions against the Sscrofa11.1 reference genome assembly, extracting positional candidate genes for a more nuanced understanding of the genetic basis of diverse traits in pigs.

Gene Ontology (GO) enrichment analysis of swept genes was implemented with the R package clusterProfiler 4.0 [42]. We considered GO terms with corrected P -value < 0.05 to be significantly enriched.

We downloaded 15 chromatin states, including promoters (TssA, TssAHet, and TssBiv), TSS-proximal transcribed regions (TxFlnk, TxFlnkWk, and TxFlnkHet), enhancers (EnhA, EnhAMe, EnhAWk, EnhAHet, and EnhPois), repressed regions (Repr and ReprWk), quiescent regions (Qui), and accessible but did not coincide with any other measured epigenetic marks (ATAC islands) for 14 pig tissues (Adipose, Cecum, Cerebellum, Colon, Cortex, Duodenum, Hypothalamus, Ileum, Jejunum, Liver, Lung, Muscle, Spleen, and Stomach) from publicly available datasets [43]. We calculated the significance of enrichment based on Fisher's exact test using the R package LOLA v1.32.0 [44].

Identification of putative functional SNPs

To identify putative functional SNPs, we first calculated F_{ST} by site and the top 1% sites located in specifically selected regions with high absolute allele frequency difference ($\Delta AF > 0.7$) considered as candidate SNPs. The pCADD scores (a model based on the CADD methodology to create scores for the prioritisation of SNVs with respect to their putative deleteriousness in the genomes of wild and domesticated pigs) [45] were retrieved from public databases to prioritize coding variants. We downloaded the gene expression matrices of different pig breeds from publicly available datasets, PIGOME (<http://pig123456789.pigome.com/>). Motif analysis based on the JASPAR database [46] using the HOMER [47] for non-coding candidate SNPs located in the promoter or enhancer regions of candidate genes.

Cell transfection and dual-luciferase reporter assays

A region of the *ABCA1* gene, encompassing 401 bp including the SNP (G-A, chr1: 246,175,129), was amplified and cloned into the pGL4.23-basic Luciferase Reporter Vector (GeneCreate, China). HEK293T cells

(sourced from our laboratory's cell bank), resuspended in DMEM supplemented with 10% fetal bovine serum, were seeded into 12-well plates and cultured overnight at 37 °C in a 5% CO₂ incubator. Upon reaching approximately 60% confluency, the cells were transfected with plasmids. The luciferase vector containing the SNP site (experimental vector; 1 µg) and the control reporter vector pRL-TK (20 ng) were co-transfected into the HEK293T cells at a 50:1 ratio using Attractene Transfection Reagent (Qiagen, Germany) according to the manufacturer's protocol. After 24 h of transfection, the transfected cells were collected and lysed. Luciferase activity was then measured using the Dual-Luciferase Reporter Assay System (Promega, USA), and the results were normalized to Renilla luciferase activity. Two-tailed Student's *t*-test was employed to determine the significance of differences, setting the significance threshold at $P < 0.05$.

Results

Genomic diversity, phylogenetic relationships and population structure

We collected data from 82 individuals of 6 Chinese native pig breeds, namely, Ding'an pigs (DA), Tunchang pigs

(TUC), Wuzhishan pigs (WZS), Min pigs (MZ), Hetao pigs (HT), and Tibetan pigs (TP), which are spread across three classical geographical regions, i.e., tropical, high-altitude, and frigid environments (Fig. 1A and Table S1). All 82 individuals were sequenced at depths greater than 10×. After applying stringent quality control criteria, we identified a total of 25,602,818 SNPs. By comparing the SNP set with the pig dbSNP database, we found that more than 13.5% of the variants (3,466,300 SNPs) were novel, which substantially expanded the catalog of porcine genetic variants (Fig. S3A). Further functional annotation revealed that most SNPs (64.12%) were located in intronic regions, followed by intergenic regions (22.82%). In addition, 1.21% of the SNPs were identified in coding regions, 102,024 of which were nonsynonymous variants (100,954 missense, 874 stop-gain, and 196 stop-loss mutants) (Fig. S3B).

LD generally decreased as the distance between loci increased, and the strength of LD varied widely between populations. The physical distance between SNPs, measured as half of the maximal value, was 34.1 kb ($r^2 = 0.34$) for DA and 1.8–5.7 kb ($r^2 = 0.23$ –0.31) for the remaining five pig breeds. At longer marker distances, the LD value

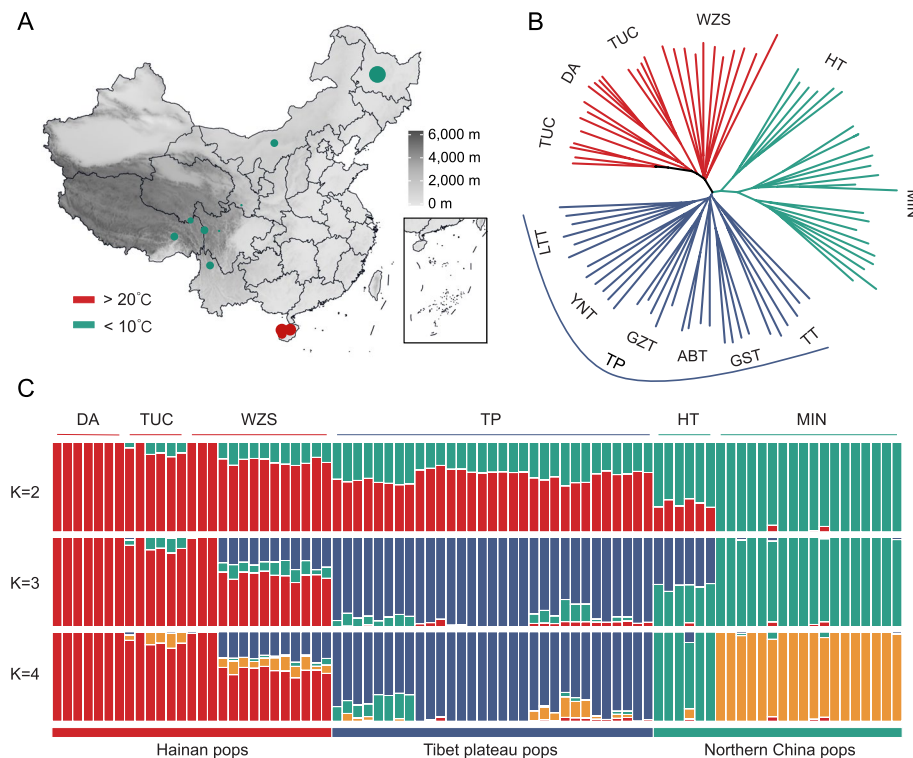


Fig. 1 Geographic distribution and population genetics analyses of 6 domestic pig breeds. DA, Ding'an pigs; TUC, Tunchang pigs; WZS, Wuzhishan pigs; MZ, Min pigs; HT, Hetao pigs; TP, Tibetan pigs (YNT, TP in Yunnan; LTT, TP in Litang; GZT, TP in Ganzhi; GST, TP in Gansu; ABT, TP in A'ba; TT, TP in Tibet). **A** Sampling locations of the 82 individuals in this study. The point size indicates the population size (4–18); the dot color indicates the annual average temperature. The elevation is shown on the map with a gradient. **B** The neighbor-joining tree was constructed based on whole-genome SNPs. **C** Population genetic structure of the 82 individuals. The number of assumed genetic clusters *K* ranged from 2 to 4 are shown

was highest for the DA but lowest for the Tibetan pigs (Fig. S4 and Table S3).

To infer the genetic and evolutionary relationships among pig breeds adapted to different environments, we first constructed a phylogenetic tree of 82 individuals using the NJ algorithm (Fig. 1B). The phylogenetic tree revealed distinct groupings of individuals from different regions. Specifically, the genetic relationships among the 6 breeds were strongly associated with their habitats, with three pig breeds from Hainan (DA, TUC, and WZS) exhibiting closer genetic relationships. Interestingly, within the Tibetan pig breed, the internal genetic relationships also showed significant geographical partitioning. The results of PCA were consistent with those of the phylogenetic analysis, with the first principal component (PC1 = 6.54%) and the second principal component (PC2 = 5.10%) able to separate the 6 breeds by geographical region (Fig. S5).

Population structure analysis revealed that the optimal number of clusters was three, at which point the cross-validation error was lowest, and the results were considered most reliable (Fig. 1C). When $K=2$, the three pig breeds from Hainan (DA, TUC, and WZS), TP, and HT shared more ancestral components; when $K=3$, we observed that pig breeds clustered by geographical region, consistent with the previous phylogenetic analysis and PCA results; when $K=4$, HT were separated, consistent with the fact that their distribution area does not overlap with that of other pig breeds. The ADMIXTURE results further confirmed that the genetic relationships among pig breeds were closely related to their geographical distribution.

Selection signatures on autosomes and functional annotation

To better leverage the diversity of our dataset, we partitioned the three local population samples into four groups based on the putative population structure: the high-temperature group (HP, 27 individuals consisting of DA, TUC, and WZS), the low-temperature group (NTP, 55 individuals consisting of HT, MZ, and TP), the high-altitude group (TP, 31 individuals consisting of Tibetan pigs), and the low-altitude group (HNP, 51 individuals consisting of DA, TUC, WZS, HT, and MZ). To elucidate the selective pattern of pigs in tropical environments, we conducted a comparative analysis between HP and NTP to detect selection signals. By applying the top 5% of the cutoffs for both F_{ST} and XP-EHH, we identified 55.50 Mb and 20.47 Mb selective sweep regions in HP (Table S4) and NTP, respectively. A similar approach was employed to compare TP and HNP to investigate their adaptive mechanisms in high-altitude environments, revealing 43.60 Mb selective sweep regions in TP (Table S5). To

focus on the unique regions associated with adaptation to frigid environments, we excluded the overlapping regions between the NTP selection and TP selection from the NTP selection and ultimately identified 17.72 Mb regions (Table S6). The top 25 sweep regions with the highest F_{ST} and XP-EHH scores within the candidate genomic areas were considered highly relevant regions.

We first performed an in-depth exploration of the selective patterns associated with adaptation to tropical environments (Fig. 2A). Most genes located in highly relevant regions are functional candidates for adaptation to tropical environments according to their annotations in previous studies. These genes included *VPS13A*, *GNAI4*, and *NR6A1*, which are involved in blood coagulation and circulation [48]; *STIMATE* and *NR5A1*, which participate in the temperature stress response [49–51]; *AGMO*, which affects human inflammation and energy homeostasis [52]; *LMTK2*, which is associated with cell apoptosis [53]; and *CFAP299*, which may affect the hair phenotype of yaks [54]. Additionally, we noticed the *ABCA1* gene, located in the 54th position among 2,254 candidate regions (Table S4), has been shown to reduce arteriosclerosis risk when up-regulated [55–57]. We explored the potential biological function of the detected signals with publicly available QTL and GO enrichment analyses. QTL enrichment analyses revealed that health, meat, and carcass traits were significantly enriched. We noticed significant enrichment for "Cholesterol level" and "Mean platelet volume" (Fig. S6A and Table S7). The positively selected genes (PSGs) involved in adaptation to tropical environments were mainly associated with blood circulation, protein degradation, and inflammation, including "blood vessel diameter maintenance" ($P=0.001046$), "NIK/NF-kappaB signaling" ($P=0.010334$), and "proteasomal ubiquitin-independent protein catabolic process" ($P=0.004777$) (Fig. S6B and Table S8). Improving blood flow to the surrounding skin can mitigate the effects of heat stress in tropical environments [58]. Previous research has demonstrated that the NF- κ B pathway can stimulate HSP activation in immune cells [59], contributing to reducing heat stress, and that the proteasomal ubiquitin-independent protein catabolic process can degrade misconfigured proteins caused by heat stress, reducing damage [60].

We next explored the mechanisms of hypoxia tolerance in Tibetan pigs (Fig. 2B). In the highly relevant regions of high-altitude adaptation-specific selection, two genes affecting the cardiovascular system were identified: *SOX18*, which is associated with the regulation of blood vessel development [61], and *TNNI3K*, which affects heart function [62, 63]. Additionally, in line with previous research [64], the candidate gene *EPAS1* in the hypoxia-inducible factor pathway [65] showed

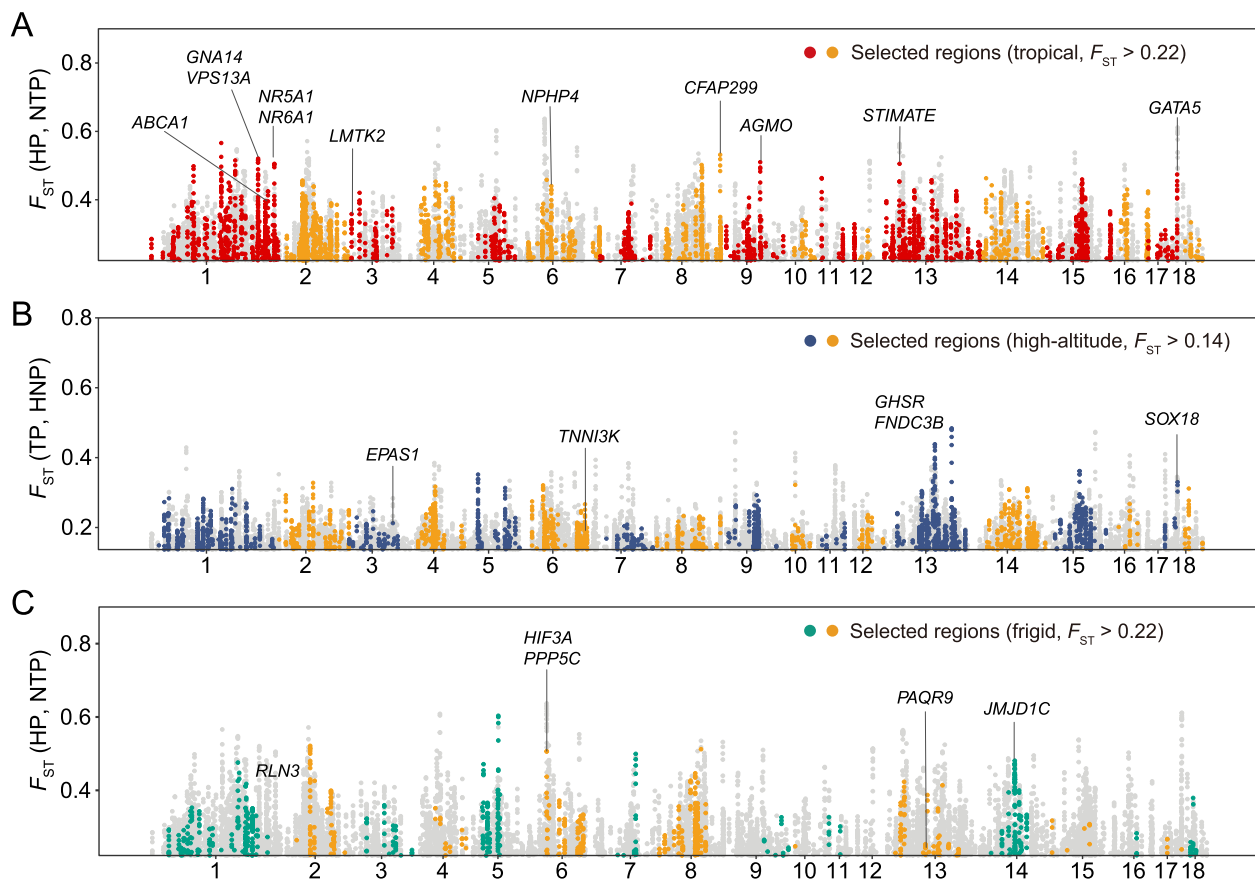


Fig. 2 Genome-wide selective signals across environmental adaptations based on F_{ST} and XP-EHH. Candidate selection regions detected by two statistics (F_{ST} and XP-EHH) are plotted across the genome. All dots represent regions identified as outliers with the F_{ST} method. Red, blue, green, and yellow dots represent the regions identified as outliers in both methods. Genes in the selected regions were marked. **A** Selected regions of adaptation to tropical environments. **B** Selected regions of adaptation to high-altitude environments. **C** Selected regions of adaptation to frigid environments

substantial selection, but the gene *EGLN1* in the same pathway did not. By annotating specific selected regions, we discovered significant enrichment for blood index-associated traits, such as "Hemoglobin" and "Platelet-crit" (Fig. 3A and Table S9) Gene Ontology analysis revealed an overrepresentation of genes involved in biological processes that contribute to maintaining typical vital signs in high-altitude environments. (Fig. 3B and Table S10). PSGs detected in Tibetan pigs have particularly enriched in hypoxia adaptation-related processes, including "cardiac cell development" ($P=6.99E-05$), "coronary vascular development" ($P=0.001084$), "platelet-derived growth factor receptor signaling pathway" ($P=0.008778$), "response to hypoxia" ($P=0.013765$), and "respiratory tube development" ($P=0.000769$). Additionally, we observed enrichment for the nucleotide metabolism process, including "pyrimidine-containing compound metabolic process" ($P=0.001629$), "deoxyribose phosphate metabolic process" ($P=0.018889$), and

"GTP metabolic process" ($P=0.002461$), which provided the basis for DNA repair and help to maintain genome stability by repairing UV-induced errors during DNA replication [66].

For adaptation to frigid environments, we detected four genes in the highly relevant regions, including *HIF3A*, which is associated with adiposity [67]; *JMJD1C*, which affects de novo lipogenesis [68]; *RLN3*, which is associated with food intake [69, 70]; and *PAQR9*, which activates brown adipocyte thermogenesis [71] (Fig. 2C). The function of these genes is critical for adaptation to frigid environments. The signal for adaptation to frigid environments was mainly associated with meat and carcass traits, especially "Fat area percentage in carcass", which may contribute to heat retention (Fig. S7A and Table S11). Biological process enrichment analysis revealed the process involved in thermogenesis "regulation of fibroblast growth factor receptor signaling pathway" ($P=0.004554$). Elevated levels of FGF21 (fibroblast



Fig. 3 Annotation of the regions and genes under the high-altitude adaptation-specific selection based on the Animal QTLdb and the Gene Ontology Resource, respectively. **A** Significantly enriched QTL terms for high-altitude specific selection. The richness factor was obtained by calculating the ratio of the number of QTLs annotated in the candidate regions and the total number of each QTL. **B** Significantly enriched GO terms (Biological process, top 10) for selected genes in high-altitude specific selection

growth factor 21) promote beige adipose tissue and enhance energy expenditure [72, 73]. We also found that domestic pigs also exhibit cold-induced vasodilation (CIVD), as indicated by the enrichment of "vasodilation" ($P=0.008992$). CIVD is a dramatic increase in peripheral blood flow observed during cold exposure. It supposedly protects against cold injuries [74]. Meanwhile, we noticed the enrichment for "cell cycle" ($P=0.002365$), in line with the fact that the mammalian cell cycle is temperature sensitive [75] (Fig. S7B and Table S12).

Annotation of variants candidate to be under selection for environmental adaptation

To better interpret the genetic basis of domestic selection, we annotated the SNPs within selected regions. We identified 10 nonsynonymous variants with high pCADD values ($pCADD > 10$) in regions specifically selected for adaptation to tropical environments (Table S13). For example, we identified the mutation p.Val244Gly in *VPS13A*, which was previously reported as a mutation that may impact the secretion and aggregation of blood platelets and reduce the risk of thrombosis in southern Chinese pigs from hot environments [15] (Fig. S8A). Additionally, in *NPHP4*, which has been verified by F_{ST} , XP-EHH, $\theta\pi$ ratio, and genotype patterns as a positively selected gene for adaptation to tropical environments (Fig. 4A–C), we found a nonsynonymous variant (p.Ala897Thr) that showed a large difference in allele frequency between HP and NTP; this was predicted to be a functional-altering variant and was found to be highly conserved across multiple

vertebrate species (Fig. 4D). *NPHP4* is a cilia-associated protein that negatively regulates the mammalian Hippo signaling pathway and is linked to severe degenerative renal disease, nephronophthisis and blindness in humans [76, 77]. We propose that the missense mutation in *NPHP4* may enhance water reabsorption in the kidney to mitigate the effects of heat stress on pigs.

To evaluate whether these putative promising variants identified in this study are originated from novel mutations or standing variants from the early stage of domestication, we referenced resequencing data from project PRJNA754250, which includes 33 Asian wild boars (AW) and 33 European wild boars (EW), covering diverse subpopulations (Table S2). Our allele frequency analysis identified a missense mutation (p.Asp2,905Asn, $pCADD=19.41$) in *VPS13B* prevalent in tropical populations, which was also found in wild populations (Fig. S8B–C and Table S13). This suggested it may be a standing variant that became dominant as pigs were domesticated in hot areas, potentially playing a role in tropical adaptation, given the *VPS13B* is closely related to *VPS13A* within the same gene family.

Chromatin state analysis enhanced the biological interpretations of adaptive evolution

Tissue-specific gene regulation plays a crucial role in the process of adaptive evolution [43]. Thus, we performed chromosome state enrichment analysis for the genomic regions under selective pressure within tropical, high-altitude, and frigid environments (Fig. 5A and Tables S14–S16). The results showed a

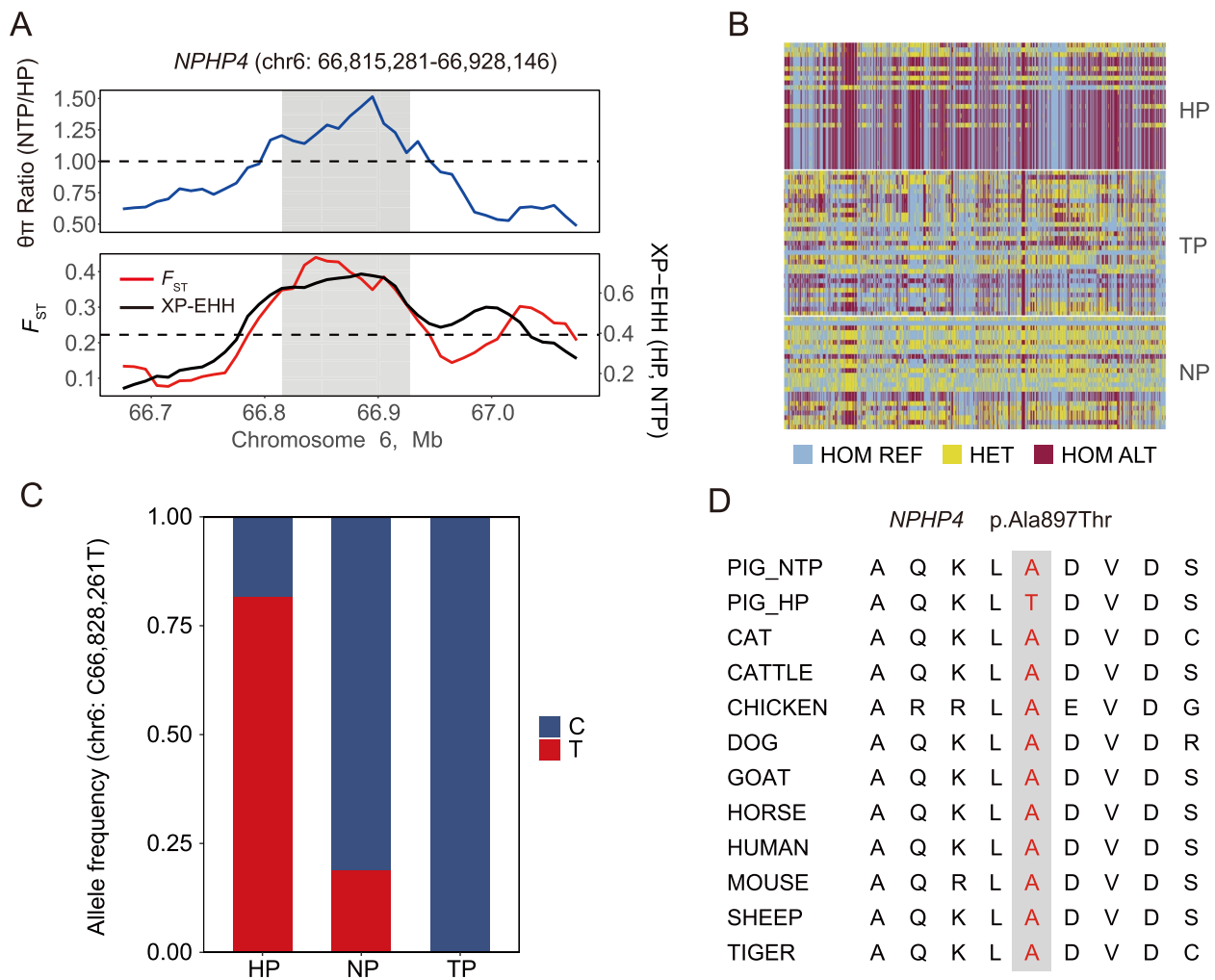


Fig. 4 Candidate variant (p.Ala897Thr) in *NPHP4* for tropical adaptation. **A** θ_T ratios (50-kb windows, 10-kb steps), F_{ST} values, and XP-EHH values around the *NPHP4* gene locus. The blue line represents θ_T ratios. The red and black lines represent F_{ST} and XP-EHH values, respectively. **B** Haplotype pattern in the genomic region of *NPHP4* among HP, TP, and NP. **C** Allele frequency of the mutation site. **D** Multispecies alignment of the protein sequences around the variant. HP, consisting of Ding'an pigs, Tunchang pigs, and Wuzhishan pigs; NP, consisting of Hetao pigs and Min pigs; TP, Tibetan pigs

high consistency: TssA and TSS-proximal transcribed regions were most enriched, followed by enhancers. Then we examined the tissue-specific promoters (TssA) (Fig. 5B and Tables S17–S19). Using a common promoter as a reference, our analysis revealed that lung-specific and ileum-specific promoters were preferentially enriched in all the three types of adaptation. Interestingly, most tissue-specific promoters in this study were preferentially enriched in adaptation to tropical environments. This indicated that heat stress affects a wide range of tissues. Several tissue-specific promoters were found to be preferentially enriched in one or more types of selection, including liver-specific promoters involved in tropical/high-altitude

adaptation and cortex-specific promoters involved in frigid/high-altitude adaptation. Additionally, spleen-specific and stomach-specific promoters were not preferentially enriched in any of the three types of adaptation.

Variation within the *cis*-regulatory regions involving tropical adaptation

The *ABCA1* gene was positively correlated in adaptation to tropical environments according to the F_{ST} , XP-EHH, θ_T ratio, and genotype patterns (Fig. S9A–B). To investigate whether the expression of the *ABCA1* gene was associated with the environmental origins of pig breeds, we analyzed gene expression data (Table S20). RNA-seq

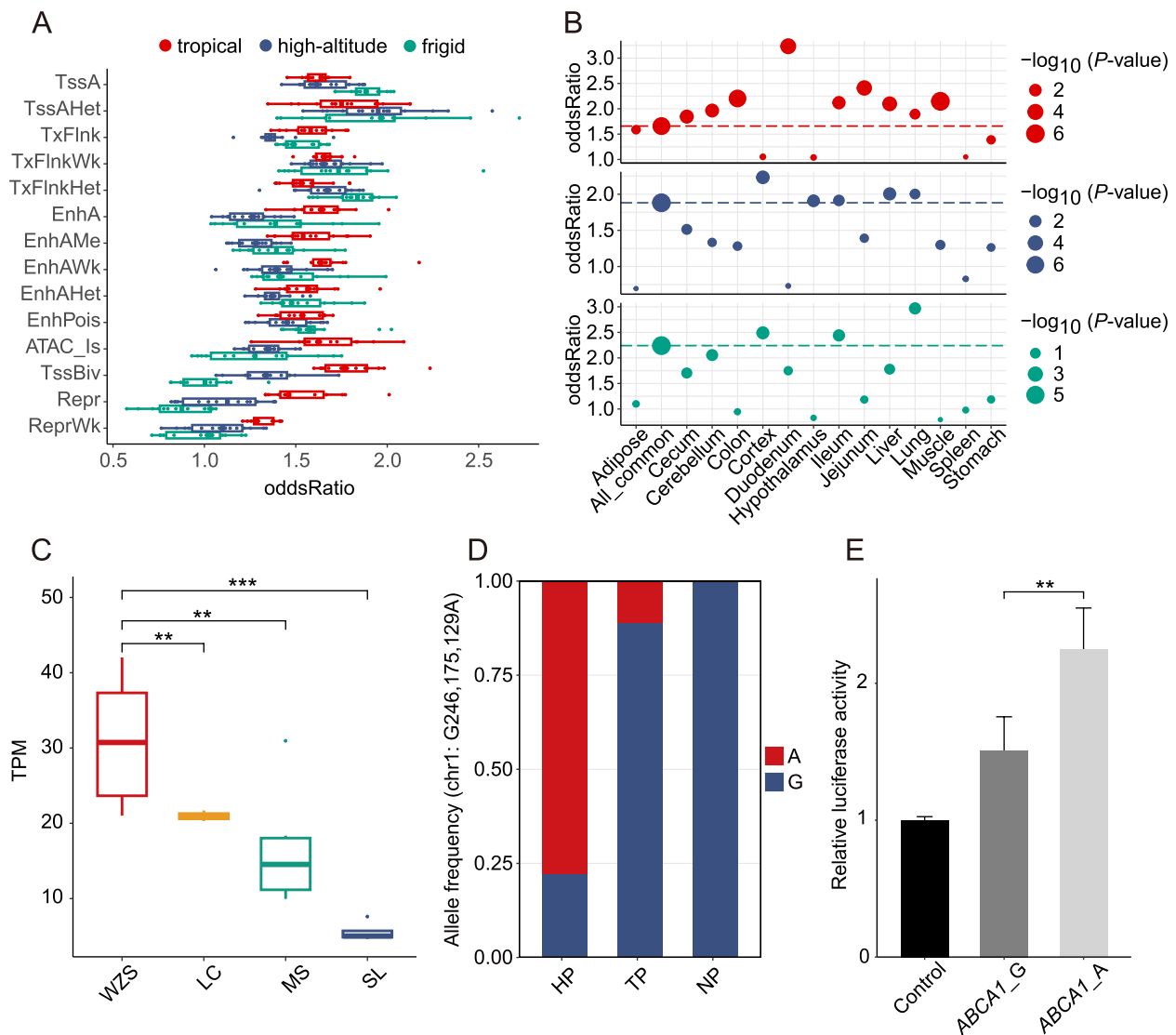


Fig. 5 Chromatin state plays an important role in the environmental adaptations of Chinese domestic pigs. **A** Enrichment of 14 chromatin states in candidate genomic regions of adaptation to tropical, high-altitude, and frigid environments. Each point represents a tissue, and oddsRatio displays the enrichment intensity of each chromatin state. **B** Enrichment of tissue-specific promoters (TssA) candidate genomic regions of adaptation to tropical, high-altitude, and frigid environments. Using the enrichment intensity of All_common promoter as a reference to divide the enrichment intensity of each tissue-specific promoter. **C** Expression of *ABCA1* gene in liver of different pig breeds. From left to right, the average annual temperature of pig breeding areas is decreasing. WZS, Wuzhishan pigs; LC, Luchuan pigs; MS, Meishan pigs; SL, Songliao pigs. Climate information can be found in Table S1. The significance of the genotype difference is tested. ** $P < 0.05$ (t -test), *** $P < 0.001$ (t -test). **D** Allele frequency of the mutation site in the intron of *ABCA1* (chr1: G246,175,129A). HP, consisting of Ding'an pigs, Tunchang pigs, and Wuzhishan pigs; NP, consisting of Hetao pigs and Min pigs; TP, Tibetan pigs. **E** Luciferase reporter assays in HEK-293T cells to compare enhancer activity between the two alleles. ** $P < 0.05$ (t -test)

data showed that *ABCA1* expression was elevated in pig liver (Fig. S9C) and positively correlated with the annual mean temperature of the breed's origin (Fig. 5C). This trend was specific to liver tissue and was not observed in other tissues (Fig. S9D–E). Based on the chromatin state data, we found a variant (chr1: 246,175,129, G-to-A) in the intron of *ABCA1* that may regulate expression, as it is

located in the enhancer region. The SNP showed a higher allele frequency in HP (77.8%) than in TP (11.3%) and NP (0). HOMER analysis revealed that the G-to-A mutation might alter the transcription factor binding motif at this position (Fig. S10 and Table S21). Subsequent luciferase reporter assays revealed higher enhancer activity of the *ABCA1* gene segment that including the A allele (Fig. 5E).

Therefore, we inferred that the mutation enhanced the expression of *ABCA1*, which may contribute to adaptation to tropical environments.

Discussion

Domestic pigs are important agricultural animals, serving as a substantial source of animal protein globally. Intense artificial selection and crossbreeding have increased the productivity of modern commercial pig breeds but reduced their adaptive potential [78, 79]. With the changing global climate, studying the genetic adaptations of local breeds to diverse environments is crucial. In this study, we conducted a comprehensive investigation of the environmental adaptability of Chinese domestic pigs using whole-genome sequencing data and multiple omics datasets. Our findings underscore the importance of understanding the adaptive potential of domestic pigs to environmental challenges and have significant implications for the breeding of highly adaptable pig breeds.

Population genetic analysis

Genomic analyses revealed differences among pig breeds from distinct geographic regions, with local Chinese breeds likely originating from ancient Yellow River basin domestication centers [80–83]. Compared to modern breeds, Chinese local pigs, particularly Tibetan pigs, exhibit faster linkage disequilibrium decay, though slower than wild boars [84–87]. This suggested that after domestication, local pig populations spread with human migration to diverse agricultural zones and were shaped by combined artificial and natural selection or gene flow. The genomic diversity of local Chinese pig breeds was closely associated with breeding practices, such as the free-range breeding of Tibetan pigs by Tibetans, which may have increased gene flow with local wild boars and resulted in faster linkage disequilibrium decay.

Functional annotation

Annotation against a single database alone cannot fully reveal the primary roles of genes within regions under selection in an organism. To explore the functions of candidate genes, we employed a multifaceted approach combining GO enrichment and QTL enrichment analyses. Our analysis confirmed a significant enrichment of traits related to blood circulation in both tropical and high-altitude adaptation in pigs, consistent with prior investigations of blood biochemical indicators in heat-stressed pigs and Tibetan pigs under normal conditions [88–90]. However, upon further analysis of the specific enriched QTLs ("LDL cholesterol" and "Cholesterol level" for adaptation to tropical environments; "Red blood cell count", "Red cell distribution width", and "Hemoglobin" for high-altitude adaptation) and specific enriched GO terms, we found that

the overlap was coincidental. Pigs adapted to the tropics need to increase blood circulation for heat dissipation and to reduce the risk of thrombosis [91, 92]. In contrast, pigs adapted to high altitudes need to reduce blood viscosity caused by high hemoglobin levels and thus reduce cardiac burden [93, 94]. This parallel selection of traits warrants consideration in the breeding of pigs with broad environmental adaptability. In contrast to the parallel selection of traits, specific selection for adaptation to distinct environments is more prevalent. Our results support this notion, as evidenced by the selection for fat coverage in pig breeds from cold regions, granulocyte activity and blood lipid content in tropical pig breeds, and hemoglobin content in pig breeds from high-altitude hypoxic environments. These specifically selected traits may facilitate the breeding of pig breeds adapted to particular environments.

Chromatin state annotation and *cis*-regulatory mutations

Gene regulation plays a crucial role in speciation and adaptive diversification [95–97]. *cis*-Regulatory mutations can alter the expression of proximal genes and have long been considered important targets for adaptive phenotypic evolution, as they may have fewer deleterious effects than changes in protein-coding sequences [98–100]. While protein-coding mutations may affect protein products throughout tissues and developmental stages, *cis*-regulatory mutations can influence gene expression in a spatially and temporally specific manner. Several studies have identified the importance of noncoding region mutations in local adaptation [101]. Previous research on the adaptive evolution of domestic pigs has focused primarily on protein-coding regions of the genome, annotating candidate gene functions to elucidate environmental adaptation, with little systematic exploration of regulatory regions.

To investigate the mechanisms by which variation in regulatory regions affects the environmental adaptability of pigs, we analyzed tissue-specific chromatin states in candidate regions. By examining the enrichment of tissue-specific regulatory factors, we pinpointed the lung and ileum as common functional tissues for adaptation to tropical, high-altitude, and frigid environments. The role of the lung, the undertaker of respiration, in the process of adapting to various environments has been well studied [102–104]. The ileum is the exclusive site for vitamin B₁₂ absorption, crucial for maintaining energy production and stress response under environmental changes [105]. As the main intestinal segment where Peyer's patches are distributed, ileum plays a significant role in immune homeostasis by regulating gut microbiota through interactions involving IgA and epithelial barriers, enhancing disease resistance and environmental adaptability [106–109]. Additionally, the interactions between the host and the gut microbiome influences host metabolic processes

and adaptability to diverse climates [110–116], underscoring the pivotal role of ileum in physiological resilience. In contrast, the spleen and stomach did not appear to have a specific role in the adaptation of domestic pigs to tropical, high-altitude, or frigid environments.

Our investigation demonstrated a possible association between the *ABCA1* gene and tropical adaptation in domestic pigs. The identified SNP (chr1: 246,175,129, G-to-A), situated within the *cis*-regulatory region of *ABCA1*, exhibits genotype frequencies correlated with warmer climates. However, the precise mechanisms through which this SNP influences *ABCA1* expression—and thereby contributes to tropical adaptation—remain unclear. To substantiate this hypothesis, further studies investigating the three-dimensional interactions between the SNP and the *ABCA1* gene, alongside comprehensive animal experiments, are essential for a more definitive understanding of the gene's role in environmental adaptability.

Through the elucidation of the positive selection phenomena occurring within regulatory regions during environmental adaptation, our study extended our understanding of pig environmental adaptability to specific tissues, providing a framework for incorporating single-cell data into future adaptability research while also highlighting the crucial role played by *cis*-regulatory mutations in enabling pig adaptation to tropical environments.

Conclusion

This genome-wide study revealed the genetic mechanisms behind the adaptation of Chinese pig breeds to tropical, high-altitude, and frigid environments. Our findings revealed significant population stratification driven by habitat isolation and pinpointed genomic regions under selection linked to key adaptive traits. Notably, we identified candidate functional tissues for adaptation include the lung and ileum and uncovered a mutation (chr1: G246,175,129A) in the *cis*-regulatory region of *ABCA1* gene as a promising variant for tropical adaptation. Overall, the genetic diversity observed, encompassing both coding and non-coding regions, underscored the complex genetic basis of environmental tolerance and survival. Our results provided profound insights into the genetic architecture of environmental adaptation, offering valuable implications for pig breeding programs and serving as a reference for understanding human diseases related to environmental stress.

Abbreviations

AF	Allele frequency
CIVD	Cold-induced vasodilation
EnhA	Strong active enhancer
EnhAHet	Active enhancer no ATAC
EnhAMe	Medium enhancer with ATAC
EnhAWk	Weak active enhancer
EnhPois	Poised enhancer

GO	Gene Ontology
LD	Linkage disequilibrium
MAF	Minor allele frequency
NJ	Neighbor-Joining
PCA	Principal component analysis
pCADD	Combined annotation dependent depletion for pig
PCR	Polymerase chain reaction
PSG	Positively selected gene
QTL	Quantitative trait locus
Qui	Quiescent regions
Repr	Repressed polycomb
ReprWk	Weak repressed polycomb
SNP	Single nucleotide polymorphism
TssA	Strongly active promoters/transcripts
TssAhet	Flanking active TSS without ATAC
TssBiv	Bivalent/Poised TSS
TxFlnk	Transcribed at gene
TxFlnkHet	Transcribed region without ATAC
TxFlnkWk	Weak transcribed at gene

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40104-024-01053-0>.

Additional file 1: Fig. S1. Genome-wide distribution of SNPs. Out of 226,375 windows of 50 kb in length sliding in 10 kb steps across the domestic pig genome, 5,413 windows contain < 100 SNPs (red bars) and cover 2.39% of the genome (dashed lines). 220,096 windows contain \geq 100 SNPs (blue bars) and cover 97.61% of the genome, and these were used to detect signatures of selective sweeps. The cumulative % in whole genome length (black line) is also charted. **Fig. S2.** The block size distribution for each autosome in our data. **Fig. S3.** Genetic variants of 82 genomes from 6 local Chinese pig breeds (Ding'an pigs; Hetao pigs; Min pigs; Tibetan pigs; Tunchang pigs; Wuzhishan pigs). **A** Venn diagrams for novel variants detected in domestic pigs. **B** Annotation of 25,602,818 SNPs retrieved from domestic pigs. **Fig. S4.** Decay of Linkage disequilibrium (LD) for 6 breeds, with one line per breed. (DA, Ding'an pigs; HT, Hetao pigs; MZ, Min pigs; TP, Tibetan pigs; TUC, Tunchang pigs; WZS, Wuzhishan pigs). **Fig. S5.** A principal component plot of the 82 individuals based on SNP information. The color represents the location of the pig breeds. DA, Ding'an pigs; HT, Hetao pigs; MZ, Min pigs; TP, Tibetan pigs; TUC, Tunchang pigs; WZS, Wuzhishan pigs. **Fig. S6.** Annotation of the regions and genes under tropical adaptation-specific selection based on the Animal QTLdb and the Gene Ontology Resource, respectively. **A** Significantly enriched QTL terms for tropical-specific selection. **B** Significantly enriched GO terms (Biological process, top 10) for tropical-specific selection. **Fig. S7.** Annotation of the regions and genes under frigid adaptation-specific selection based on the pig QTLdb and the Gene Ontology Resource, respectively. **A** Significantly enriched QTL terms for frigid-specific selection. **B** Significantly enriched GO terms (Biological process, top 10) for frigid-specific selection. **Fig. S8.** Multispecies regional alignment of the *VPS13A* and *VPS13B* protein sequences around candidate variants for adaptation to tropical environments. **A** p.Val244Gly in *VPS13A* gene. **B** p.Asp2,905Asn in *VPS13B* gene. **C** Allele frequency of the putative promising variant in *VPS13B* among different breeds. AW, Asian wild pigs; EW, European wild boars; AWNC, Northern Chinese wild boars; AWSC, Southern Chinese wild boars. **Fig. S9.** Distinct genomic landscape around *ABCA1* gene. **A** $\theta\pi$ ratios (50-kb windows, 10-kb steps), F_{ST} values, and XP-EHH values around the *ABCA1* gene locus. The blue line represents $\theta\pi$ ratios. The red and black lines represent F_{ST} and XP-EHH values, respectively. **B** Haplotype pattern in the genomic region of *ABCA1* among HP, TP, and NP. HP, consisting of Ding'an pigs, Tunchang pigs, and Wuzhishan pigs; TP, Tibetan pigs; NP, consisting of Hetao pigs and Min pigs. **C** The *ABCA1* gene shows high expression in pig liver. **D** and **E** Expression of *ABCA1* gene in adipose and muscle tissues of different pig breeds. From left to right, the average annual temperature of pig breeding areas is decreasing. The significance of the genotype difference is tested. ** $P < 0.05$ (*t*-test), *** $P < 0.001$ (*t*-test). **Fig. S10.** Sequence logos of *NFIA* and *MEIS2*. Site 7 (G allele) of *NFIA* and site 4 (A allele) of *MEIS2* are the binding sites of chr1: 246,175,129.

Additional file 2: Table S1. Samples, origin, and environmental variables of Chinese pigs in this study. **Table S2.** The sequencing quality and read alignment statistic of whole-genome sequencing data in this study. **Table S3.** The average value of LD in different distance regions among 6 domestic pig breeds. **Table S4.** List of regions under selection for adaptation to tropical environments and enumeration of the genes within such regions. **Table S5.** List of regions under selection for adaptation to high-altitude environments and enumeration of the genes within such regions. **Table S6.** List of regions under selection for adaptation to frigid environments and enumeration of the genes within such regions. **Table S7.** Fold enrichment of tropical adaptation-specific selection signatures for pig QTL terms. **Table S8.** Significantly enriched GO terms of the genes in tropical adaptation-specific selection regions. **Table S9.** Fold enrichment of high-altitude adaptation-specific selection signatures for pig QTL terms. **Table S10.** Significantly enriched GO terms of the genes in high-altitude adaptation-specific selection regions. **Table S11.** Fold enrichment of frigid adaptation-specific selection signatures for pig QTL terms. **Table S12.** Significantly enriched GO terms of the genes in frigid adaptation-specific selection regions. **Table S13.** Candidate nonsynonymous variants ($\Delta AF > 0.7$) of tropical adaptation-specific selection with high pCADD score. **Table S14.** Fold enrichment of tropical adaptation-specific selection signatures for chromatin states. **Table S15.** Fold enrichment of high-altitude adaptation-specific selection signatures for chromatin states. **Table S16.** Fold enrichment of frigid adaptation-specific selection signatures for chromatin states. **Table S17.** Fold enrichment of tropical adaptation-specific selection signatures for tissue-specific promoters. **Table S18.** Fold enrichment of high-altitude adaptation-specific selection signatures for tissue-specific promoters. **Table S19.** Fold enrichment of frigid adaptation-specific selection signatures for tissue-specific promoters. **Table S20.** *ABCA1* expression levels across various tissues in different pig breeds (Download from <http://pig123456789.pigome.com/>). **Table S21.** Prediction of the transcription factors binding around the SNP (chr1:246,175,129).

Acknowledgements

We are grateful to Hongwei Yin for his suggestions for performing selective signal detection and to Guohao Han for collecting RNA-seq data.

Authors' contributions

GY and ZW conceived and designed the study. ZW collected the data, performed bioinformatics analysis, and visualized the results. BS and JY conducted the experiments. ZW wrote the original draft. GY, ZT, ZW, BS, XL and YZ reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFF1000600), the National Natural Science Foundation of China (32002150 and U23A20229), the Basic and Applied Basic Research Foundation of Guangdong Province (2020B1515120053), the Shenzhen Science and Technology Innovation Commission (JCYJ20190813114401691), the Central Government Guiding Funds for Local Science and Technology Development of China (He-Ke ZY220603), and the Open Project of Hainan Provincial Key Laboratory of Tropical Animal Reproduction & Breeding and Epidemic Disease Research (HKL2020101).

Availability of data and materials

The whole genome resequencing data that were analyzed during the current study are available in the NCBI primary data archive (PDA) with accession number PRJNA754250. Chromatin state data used in this study can be found in: http://farm.cse.ucdavis.edu/~zhypan/Nature_Communications_2021. The RNA-seq data is available in the publicly available datasets, PIGOME (<http://pig123456789.pigome.com/>).

Declarations

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Livestock and Poultry Multi-Omics of MARA, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China. ²Kunpeng Institute of Modern Agriculture at Foshan, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Foshan 528226, China. ³School of Life Sciences, Henan University, Kaifeng 475004, China. ⁴Shenzhen Research Institute of Henan University, Shenzhen 518000, China. ⁵State Key Laboratory of Pharmaceutical Biotechnology, The University of Hong Kong, Hong Kong SAR, China. ⁶Key Laboratory of Tropical Animal Breeding and Disease Research, Institute of Animal Science and Veterinary Medicine, Hainan Academy of Agricultural Sciences, Haikou 571100, China. ⁷Bama Yao Autonomous County Rural Revitalization Research Institute, Bama 547500, China.

Received: 7 February 2024 Accepted: 20 May 2024

Published online: 10 July 2024

References

- Miraglia M, Marvin HJP, Kleter GA, Battilani P, Brera C, Coni E, et al. Climate change and food safety: an emerging issue with special focus on Europe. *Food Chem Toxicol.* 2009;47:1009–21.
- Nevo E. Evolution of genome–phenome diversity under environmental stress. *Proc Natl Acad Sci.* 2001;98:6233–40.
- Barrett RDH, Schluter D. Adaptation from standing genetic variation. *Trends Ecol Evol.* 2008;23:38–44.
- Burke MK. How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc R Soc B Biol Sci.* 2012;279:5029–38.
- Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, et al. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci.* 2007;104:15276–81.
- Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, Robins J, et al. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science.* 2005;307:1618–21.
- Rischkowsky B, Pilling D. *The State of the World's Animal Genetic Resources for Food and Agriculture.* Rome: Commission on Genetic Resources for Food and Agriculture, Food & Agriculture Organization of the United Nations; 2007.
- Schook LB, Collares TV, Darfour-Oduro KA, De AK, Rund LA, Schachtschneider KM, et al. Unraveling the swine genome: implications for human health. *Annu Rev Anim Biosci.* 2015;3:219–44.
- Wang M-S, Li Y, Peng M-S, Zhong L, Wang Z-J, Li Q-Y, et al. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Mol Biol Evol.* 2015;32:1880–9.
- Liu X, Zhang Y, Li Y, Pan J, Wang D, Chen W, et al. EPAS1 gain-of-function mutation contributes to high-altitude adaptation in Tibetan horses. *Mol Biol Evol.* 2019;36:2591–603.
- Kim E-S, Elbeltagy AR, Aboul-Naga AM, Rischkowsky B, Sayre B, Mwacharo JM, et al. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity.* 2016;116:255–64.
- Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, et al. The yak genome and adaptation to life at high altitude. *Nat Genet.* 2012;44:946–9.
- Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, et al. Camelid genomes reveal evolution and adaptation to desert environments. *Nat Commun.* 2014;5:5188.
- Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, et al. Genome sequences of wild and domestic bactrian camels. *Nat Commun.* 2012;3:1202.
- Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, et al. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat Genet.* 2015;47:217–25.
- Cesconeto RJ, Joost S, McManus CM, Paiva SR, Cobuci JA, Braccini J. Landscape genomic approach to detect selection signatures in locally adapted Brazilian swine genetic groups. *Ecol Evol.* 2017;7:9544–56.

17. Zhong ZQ, Li R, Wang Z, Tian SS, Xie XF, Wang ZY, et al. Genome-wide scans for selection signatures in indigenous pigs revealed candidate genes relating to heat tolerance. *Animal*. 2023;17:100882.
18. Liu P, Liang Y, Li L, Lv X, He Z, Gu Y. Identification of selection signatures and candidate genes related to environmental adaptation and economic traits in Tibetan pigs. *Animals*. 2024;14:654.
19. Yi G, Liu L, Yao Y, Liu Y, Li J, Yang Y, et al. Multi-omics analysis reveals signatures of selection and loci associated with complex traits in pigs. *BioRxiv*. 2023. <https://doi.org/10.1101/2023.09.19.558553>.
20. Zhao Y, Liao J, Zhang Q, Chen J, Gong X, Shi Y, et al. Development of China ground climate value dataset from 1991 to 2020. *Chin J Atmospheric Sci*. 2022;48(2):555–71.
21. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience*. 2018;7(1):gix120.
22. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics*. 2018;34:i884–90.
23. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997v2 [q-bio.GN]*. 2013.
24. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf Engl*. 2009;25:2078–9.
25. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–303.
26. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. 2012;6:80–92.
27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analysis. *Am J Hum Genet*. 2007;81:559–75.
28. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49:W293–6.
29. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–64.
30. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.
31. Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinforma Oxf Engl*. 2019;35:1786–8.
32. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*. 1965;19:395–420.
33. Kaplan NL, Hudson RR, Langley CH. The “hitchhiking effect” revisited. *Genetics*. 1989;123:887–99.
34. Wiehe TH, Stephan W. Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from *Drosophila melanogaster*. *Mol Biol Evol*. 1993;10:842–54.
35. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–8.
36. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 2007;449:913–8.
37. Szpiech ZA, Hernandez RD. selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. *Mol Biol Evol*. 2014;31:2824–7.
38. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007;81:1084–97.
39. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart – biological queries made easy. *BMC Genomics*. 2009;10:22.
40. Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *GigaScience*. 2020;9:giaa149.
41. Hu Z-L, Park CA, Reecy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrdB. *Nucleic Acids Res*. 2019;47:D701–10.
42. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141.
43. Pan Z, Yao Y, Yin H, Cai Z, Wang Y, Bai L, et al. Pig genome functional annotation enhances the biological interpretation of complex traits and human disease. *Nat Commun*. 2021;12:5848.
44. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinforma Oxf Engl*. 2016;32:587–9.
45. Groß C, Derks M, Megens H-J, Bosse M, Groenen MAM, Reinders M, et al. pCADD: SNV prioritisation in *Sus scrofa*. *Genet Sel Evol*. 2020;52:4.
46. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviute I, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2022;50:D165–73.
47. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38:576–89.
48. Wang Y, Zhang Y, Dai X, Liu Z, Yin P, Wang N, et al. NR6A1 couples with cAMP response element binding protein and regulates vascular smooth muscle cell migration. *Int J Biochem Cell Biol*. 2015;69:225–32.
49. Song J, Zheng S, Nguyen N, Wang Y, Zhou Y, Lin K. Integrated pipeline for inferring the evolutionary history of a gene family embedded in the species tree: a case study on the STIMATE gene family. *BMC Bioinformatics*. 2017;18:439.
50. Hooper R, Samakai E, Kedra J, Soboloff J. Multifaceted roles of STIM proteins. *Pflugers Arch - Eur J Physiol*. 2013;465:1383–96.
51. Görres A, Ponsuksili S, Wimmers K, Muráni E. Genetic variation of the porcine NR5A1 is associated with meat color. *J Appl Genet*. 2016;57:81–9.
52. Sailer S, Keller MA, Werner ER, Watschinger K. The emerging physiological role of AGMO 10 years after its gene identification. *Life*. 2021;11:88.
53. Cruz DF, Farinha CM, Swiatecka-Urban A. Unraveling the function of lemur tyrosine kinase 2 network. *Front Pharmacol*. 2019;10:24.
54. Bao Q, Ma X, Jia C, Wu X, Wu Y, Meng G, et al. Resequencing and signatures of selective scans point to candidate genetic variants for hair length traits in long-haired and normal-haired Tianzhu White Yak. *Front Genet*. 2022;13:798076.
55. Zhao Z-W, Zhang M, Wang G, Zou J, Gao J-H, Zhou L, et al. Astragalin retards atherosclerosis by promoting cholesterol efflux and inhibiting the inflammatory response via upregulating ABCA1 and ABCG1 expression in macrophages. *J Cardiovasc Pharmacol*. 2021;77:217.
56. Yu X-H, Chen J-J, Deng W-Y, Xu X-D, Liu Q-X, Shi M-W, et al. Biochanin A mitigates atherosclerosis by inhibiting lipid accumulation and inflammatory response. *Oxid Med Cell Longev*. 2020;2020:8965047.
57. Wang G, Gao J-H, He L-H, Yu X-H, Zhao Z-W, Zou J, et al. Fargesin alleviates atherosclerosis by promoting reverse cholesterol transport and reducing inflammatory response. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids*. 2020;1865:158633.
58. Shilja S, Sejian V, Bagath M, Mech A, David CG, Kurien EK, et al. Adaptive capability as indicated by behavioral and physiological responses, plasma HSP70 level, and PBMC HSP70 mRNA expression in Osmanabadi goats subjected to combined (heat and nutritional) stressors. *Int J Biometeorol*. 2016;60:1311–23.
59. Basu S, Binder RJ, Suto R, Anderson KM, Srivastava PK. Necrotic but not apoptotic cell death releases heat shock proteins, which deliver a partial maturation signal to dendritic cells and activate the NF- κ B pathway. *Int Immunol*. 2000;12:1539–46.
60. McLoughlin F, Kim M, Marshall RS, Vierstra RD, Vierling E. HSP101 interacts with the proteasome and promotes the clearance of ubiquitylated protein aggregates. *Plant Physiol*. 2019;180:1829–47.
61. Downes M, Koopman P. SOX18 and the transcriptional regulation of blood vessel development. *Trends Cardiovasc Med*. 2001;11:318–24.
62. Wang L, Wang H, Ye J, Xu R, Song L, Shi N, et al. Adenovirus-mediated overexpression of cardiac troponin I-interacting kinase promotes cardiomyocyte hypertrophy. *Clin Exp Pharmacol Physiol*. 2011;38:278–84.
63. Wang X, Wang J, Su M, Wang C, Chen J, Wang H, et al. TNNI3K, a cardiac-specific kinase, promotes physiological cardiac hypertrophy in transgenic mice. *PLoS One*. 2013;8:e58570.
64. Ai H, Yang B, Li J, Xie X, Chen H, Ren J. Population history and genomic signatures for high-altitude adaptation in Tibetan pigs. *BMC Genomics*. 2014;15:834.

65. Simonson TS, McClain DA, Jorde LB, Prchal JT. Genetic determinants of Tibetan high-altitude adaptation. *Hum Genet.* 2012;131:527–33.
66. Liu R, Li J, Shao J, Lee J-H, Qiu X, Xiao Y, et al. Innate immune response orchestrates phosphoribosyl pyrophosphate synthetases to support DNA repair. *Cell Metab.* 2021;33:2076–2089.e9.
67. Pan H, Lin X, Wu Y, Chen L, Teh AL, Soh SE, et al. HIF3A association with adiposity: the story begins before birth. *Epigenomics.* 2015;7:937–50.
68. Viscarra JA, Wang Y, Nguyen HP, Choi YG, Sul HS. Histone demethylase JMJD1C is phosphorylated by mTOR to activate de novo lipogenesis. *Nat Commun.* 2020;11:796.
69. Calvez J, de Ávila C, Timofeeva E. Sex-specific effects of relaxin-3 on food intake and body weight gain. *Br J Pharmacol.* 2017;174:1049–60.
70. Kania A, Szlaga A, Sambak P, Gugula A, Blasiak E, Bonaventura MVMD, et al. RLN3/RXFP3 Signaling in the PVN inhibits magnocellular neurons via M-like current activation and contributes to binge eating behavior. *J Neurosci.* 2020;40:5362–75.
71. Ling LIU, Si-jia LIU, Yang LIU, Qi-qun TANG. Effects of PAQR9 on brown fat thermogenesis and UCP1 expression. *Fudan Univ J Med Sci.* 2020;47:816–24.
72. Fisher FM, Kleiner S, Douris N, Fox EC, Mepani RJ, Verdeguer F, et al. FGF21 regulates PGC-1 α and browning of white adipose tissues in adaptive thermogenesis. *Genes Dev.* 2012;26:271–81.
73. Emanuelli B, Vienberg SG, Smyth G, Cheng C, Stanford KI, Arumugam M, et al. Interplay between FGF21 and insulin action in the liver regulates metabolism. *J Clin Invest.* 2014;124:515–27.
74. Daanen HA, Van Der Struijs NR. Resistance index of frostbite as a predictor of cold injury in arctic operations. *Aviat Space Environ Med.* 2005;76(12):1119–22.
75. Rieder CL, Cole RW. Cold-shock and the mammalian cell cycle. *Cell Cycle Georget Tex.* 2002;1:169–75.
76. Habbig S, Bartram MP, Müller RU, Schwarz R, Andriopoulos N, Chen S, et al. NPHP4, a cilia-associated protein, negatively regulates the Hippo pathway. *J Cell Biol.* 2011;193:633–42.
77. Awata J, Takada S, Standley C, Lechtreck KF, Bellvé KD, Pazour GJ, et al. NPHP4 controls ciliary trafficking of membrane proteins and large soluble proteins at the transition zone. *J Cell Sci.* 2014;127(21):4714–27.
78. Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, et al. Are cattle, sheep, and goats endangered species? *Mol Ecol.* 2008;17:275–84.
79. Leroy G, Baumung R, Boettcher P, Scherf B, Hoffmann I. Review: Sustainability of crossbreeding in developing countries; definitely not like crossing a meadow. . . . *Animal.* 2016;10:262–73.
80. Xiang H, Gao J, Cai D, Luo Y, Yu B, Liu L, et al. Origin and dispersal of early domestic pigs in northern China. *Sci Rep.* 2017;7:5602.
81. Larson G, Liu R, Zhao X, Yuan J, Fuller D, Barton L, et al. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc Natl Acad Sci.* 2010;107:7686–91.
82. Wang Y, Sun Y, Royle TCA, Zhang X, Zheng Y, Tang Z, et al. Ancient DNA investigation of the domestication history of pigs in Northeast China. *J Archaeol Sci.* 2022;141:105590.
83. Ma Y-F, Han X-M, Huang C-P, Zhong L, Adeola AC, Irwin DM, et al. Population genomics analysis revealed origin and high-altitude adaptation of Tibetan pigs. *Sci Rep.* 2019;9:11463.
84. Herrero-Medrano JM, Megens H-J, Groenen MA, Ramis G, Bosse M, Pérez-Enciso M, et al. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genet.* 2013;14:106.
85. Ai H, Huang L, Ren J. Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers. *PLoS ONE.* 2013;8:e56001.
86. Wang L, Sørensen P, Janss L, Ostensen T, Edwards D. Genome-wide and local pattern of linkage disequilibrium and persistence of phase for 3 Danish pig breeds. *BMC Genet.* 2013;14:115.
87. Li D, Huang M, Zhuang Z, Ding R, Gu T, Hong L, et al. Genomic analyses revealed the genetic difference and potential selection genes of growth traits in two Duroc lines. *Front Vet Sci.* 2021;8:725367.
88. Waltz X, Baillet M, Connes P, Gourdière J-L, Philibert L, Beltan E, et al. Effect of heat stress on blood rheology in different pig breeds. *Clin Hemorheol Microcirc.* 2014;58:395–402.
89. Podder M, Bera S, Naskar S, Sahu D, Mukherjee J, Patra AK. Physiological, blood-biochemical and behavioural changes of Ghongroo pigs in seasonal heat stress of a hot-humid tropical environment. *Int J Biometeorol.* 2022;66:1349–64.
90. Kong XY, Gou X, Ma T, Leng J, Mao HM, Yang SL. Study on blood physiological indicators of adaptation to hypoxia in Tibet pig (in Chinese). *J Yunnan Agric Univ.* 2014;29(2):297–300.
91. Jansky L, Hart JS. Cardiac output and organ blood flow in warm- and cold-acclimated rats exposed to cold. *Can J Physiol Pharmacol.* 1968;46:653–9.
92. Keatinge WR, Coleshaw SRK, Easton JC, Cotter F, Mattock MB, Chelliah R. Increased platelet and red cell counts, blood viscosity, and plasma cholesterol levels during heat stress, and mortality from coronary and cerebral thrombosis. *Am J Med.* 1986;81:795–800.
93. Tang F, Feng L, Li R, Wang W, Liu H, Yang Q, et al. Inhibition of suicidal erythrocyte death by chronic hypoxia. *High Alt Med Biol.* 2019;20:112–9.
94. Yang Y, Lu F, Zhuang L, Yang S, Kong Y, Tan W, et al. Combined preconditioning with hypoxia and GYKI-52466 protects rats from cerebral ischemic injury by HIF-1 α /eNOS pathway. *Am J Transl Res.* 2017;9(12):5308–19.
95. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature.* 2012;484:55–61.
96. Mack KL, Ballinger MA, Phifer-Rixey M, Nachman MW. Gene regulation underlies environmental adaptation in house mice. *Genome Res.* 2018;28:1636–45.
97. Liu Y, Fu Y, Yang Y, Yi G, Lian J, Xie B, et al. Integration of multi-omics data reveals cis-regulatory variants that are associated with phenotypic differentiation of eastern from western pigs. *Genet Sel Evol.* 2022;54:62.
98. Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet.* 2007;8:206–16.
99. Stern DL, Orgogozo V. The loci of evolution: how predictable is genetic evolution? *Evolution.* 2008;62:2155–77.
100. Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet.* 2012;13:59–69.
101. Fraser HB. Gene expression drives local adaptation in humans. *Genome Res.* 2013;23:1089–96.
102. Frisanzo AR. Functional adaptation to high altitude hypoxia. *Science.* 1975;187:313–9.
103. Wong HR, Wispe JR. The stress response and the lung. *Am J Physiol-Lung Cell Mol Physiol.* 1997;273:L1–9.
104. Liu J, Wu J, Qiao C, He Y, Xia S, Zheng Y, et al. Impact of chronic cold exposure on lung inflammation, pyroptosis and oxidative stress in mice. *Int Immunopharmacol.* 2023;115:109590.
105. Shaw S, Jayatilake E, Meyers S, Colman N, Herzlich B, Herbert V. The ileum is the major site of absorption of vitamin B12 analogues. *Am J Gastroenterol (Springer Nature).* 1989;84(1):22–6.
106. Fransen F, Zagato E, Mazzini E, Fosso B, Manzari C, El Aidi S, et al. BALB/c and C57BL/6 mice differ in polyreactive IgA abundance, which impacts the generation of antigen-specific IgA and microbiota diversity. *Immunity.* 2015;43:527–40.
107. Moon C, Baldridge MT, Wallace MA, Burnham CA, Virgin HW, et al. Vertically transmitted faecal IgA levels determine extra-chromosomal phenotypic variation. *Nature.* 2015;521:90–3.
108. Reboldi A, Cyster JG. Peyer's patches: organizing B-cell responses at the intestinal frontier. *Immunol Rev.* 2016;271:230–45.
109. Lycke NY, Bemark M. The regulation of gut mucosal IgA B-cell responses: recent developments. *Mucosal Immunol.* 2017;10:1361–74.
110. Suzuki TA, Ley RE. The role of the microbiota in human genetic adaptation. *Science.* 2020;370:eaa26827.
111. Liu G, Li C, Liu Y, Zheng C-M, Ning Y, Yang H-G, et al. Highland adaptation of birds on the Qinghai-Tibet Plateau via gut microbiota. *Appl Microbiol Biotechnol.* 2022;106:6701–11.
112. Lv J, Qi P, Bai L-H, Yan X-D, Zhang L. Review of the relationship and underlying mechanisms between the Qinghai-Tibet plateau and host intestinal flora. *Front Microbiol.* 2022;13:1055632.

113. Ma Y, Ma S, Chang L, Wang H, Ga Q, Ma L, et al. Gut microbiota adaptation to high altitude in indigenous animals. *Biochem Biophys Res Commun.* 2019;516:120–6.
114. Zhang Y, Sun L, Zhu R, Zhang S, Liu S, Wang Y, et al. Porcine gut microbiota in mediating host metabolic adaptation to cold stress. *Npj Biofilms Microbiomes.* 2022;8:18.
115. Chevalier C, Stojanović O, Colin DJ, Suarez-Zamorano N, Tarallo V, Veyrat-Durebex C, et al. Gut microbiota orchestrates energy homeostasis during cold. *Cell.* 2015;163:1360–74.
116. Liang Z-L, Jin Y-Y, Guo Y, Qiu S-J, Zhao Y, Zhao Z-H, et al. Heat stress affects duodenal microbial community of indigenous yellow-feather broilers as determined by 16S rRNA sequencing. *Ital J Anim Sci.* 2021;20:1222–31.