RUNNING TITLE: Genetic basis of local adaptation for indigenous chickens

Genomic insights into local adaptation of indigenous chickens

Jing-wei Yuan[1, #, *], Xiao-xu Jia[2, #], Xiao-long Huang[1], Yun-lei Li[1], Ai-xin Ni[1], Guo-hong Chen[3], Nen-zhu Zheng[4], Sheng He[5], Yan-yan Sun[1], Guo-qiang Yi[6], Wei Han[2, *], Ji-lan Chen[1, *]

[#] These authors contributed equally to this work.

*Corresponding author: Ji-lan Chen (chen.jilan@163.com); Jing-wei Yuan (yuanjingwei@caas.cn); Wei Han (hanwei830@163.com)

[1]State Key Laboratory of Animal Biotech Breeding, Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

[2]Jiangsu Institute of Poultry Science, Yangzhou 225125, China

[3]Joint International Research Laboratory of Agriculture and Agri-Product Safety of the Ministry of Education of China, College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China

[4]Institute of Animal Husbandry and Veterinary Medicine, Fujian Academy of Agricultural Sciences, Fuzhou 350013, China

[5]Nujiang Promotion Station of Livestock Technology, Nujiang 673199, China

[6]Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China

Science Press ZR Zoological Research

ABSTRACT

Indigenous chickens are an essential part of biodiversity and a vital protein resource to humans, yet global warming and environmental changes pose serious threats to their survival and productivity. Therefore, assessing population adaptive capacity under shifting environments is crucial for breeding resilient animals, and guiding conservation strategies. Here, we integrated ecological and whole-genome resequencing data from 1 022 chickens from 44 Chinese indigenous populations to reveal genomic signatures of local adaptation. From 87 agroclimatic variables, we identified eight dominant environmental factors including solar radiation, precipitation, diurnal temperature range, and five landcover variables (cropland areas, water areas, trees coverage, bare ground and shrubs coverage) that shape ecological niches of indigenous chickens. Landscape and comparative genomics analyses revealed both known and novel candidate genes, such as *UNC80*, *PTPRO*, *NCOR2*, *CSF2RB*, *NXT2* and *PALLD* for the solar radiation, precipitation, diurnal temperature range, cropland areas, trees coverage and bare ground, respectively. Particularly, adaptive non-coding variants harbored in these genes exhibited spatial allelic changes across populations and acted as regulatory elements via chromatin accessibility and DNA methylation, influencing adaptation in a tissue-specific manner. Our findings underscore the rich genetic diversity of Chinese indigenous chickens and provide new insights into genomic mechanisms of local adaptation, offering valuable references for domestic animal breeding, conservation, and climate resilience.

Keywords: Local adaptation; Ecological niche modeling; Landscape genomics;

45 Population genetics; Indigenous chicken

Science Press
ZR Zoological Research

INTRODUCTION

Chicken (*Gallus gallus domesticus*) is found almost everywhere providing high-quality and affordable protein, entertainment, cultural heritage, and scientific research value (Eda, 2021). Indigenous chicken populations inhabit diverse ecosystems and serve as valuable biomedical models. Their gene reservoirs are very critical for conservation biology and for long-term genetic improvement in response to changing breeding goals, such as adaptation to extreme climate conditions, disease resistance, and stress tolerance. However, the rapid expansion of commercial breeds has severely reduced the number of indigenous chickens, increasing the risk of extinction and the loss of unique genotypes and adaptive traits (Padhi, 2016). This concern is amplified by the challenges of environment changing, underscoring the need to harness the genetic diversity of indigenous populations for climate-resilient agriculture (Meek et al., 2023).

Adaptive responses that enable survival and reproduction in extreme environments are often beyond the range of phenotypic plasticity. Locally adapted populations are therefore expected to experience strong selection and genomic modifications, resulting in phenotypic changes in metabolism, physiology, and behavior (Rocha et al., 2021). Earlier comparative studies focused on single-gene or few-gene polymorphisms underlying adaptation, such as the heat shock proteins (HSPs) and heat shock factors (HSFs) related to the heat stress in animals (Gourdine et al., 2021; Hariyono & Prihandini, 2022; McManus et al., 2022; Balakrishnan et al., 2023). However, differential expression of candidate genes may evolve neutrally and not necessarily

reflect adaptation. Meanwhile, causal variants may reside in distant regulatory elements or trans-acting factors, making it difficult to directly link genetic variation with expression data. Moreover, measuring and collecting fitness-related phenotypes is challenging due to long generation intervals, experimental constraints, and population-specific limitations (Savolainen & Lascoux, 2013). Advances in genome sequencing now allow affordable population-scale analyses, providing a powerful complementary strategy to study local adaptation across a species' distribution range (Ellegren, 2014). Furthermore, new approaches including genotype-environment association analyses, landscape genomics, and multi-omics integration are increasingly applied for conserving species in wildlife (Meek et al., 2023; Wang et al., 2025) and for breeding in crops (Campbell et al., 2025) and animals (Leng et al., 2024; Lai et al., 2025).

China's vast geography spans wide gradients in latitude, altitude and topography, creating remarkable climatic and environmental diversity. This mosaic of environments provides a natural laboratory for studying ecological adaptation (Mi et al., 2021). Archaeological and genomic evidence indicates that domestic chickens originated from the red junglefowl subspecies (*Gallus gallus spadiceus*) in southwestern China ~10000 years ago (Wang et al., 2020). Since then, more than 100 indigenous chicken breeds distributed across nearly all ecosystems have been developed across the mainland (Zhang et al., 2023). These populations represent an exceptional resource for investigating the molecular mechanisms of environmental adaptation. While prior studies examined chicken adaptation to high-altitude (Wang et al., 2015; Yuan et al., 2022), tropical and frigid (Tian et al., 2020; Shi et al., 2023), and desert or arid

environment (Gu et al., 2020). However, the specific ecological environment is driven by many contributing factors, which were rarely investigated in the region-limited or single-variable studies. How geography and environmental complexity shape genetic variations in chickens is still fascinating.

To address this, we sequenced and retrieved genomes of over 1 000 chickens from 44 indigenous populations across China, for which we characterized the functional relationship between genomic variants and agroclimatic variables. By integrating ecological and multi-layer omics analysis, we aim to 1) characterize the spatial patterns of genetic diversity and structure, 2) identify key environmental drivers of ecological niche partitioning, 3) dissect genomic signatures of local environment adaptation and allele frequency shifts across geographical ranges, and 4) elucidate potential regulatory mechanism linking adaptive variants to candidate genes.

MATERIALS AND METHODS

Ethics Statement

All animal procedures were approved by the Animal Care and Use Committee of Institute of Animal Science, Chinese Academy of Agricultural Sciences (No. IAS2021-46) and performed in accordance with the relevant guidelines and regulations set by Ministry of Agriculture and Rural Affairs of the People's Republic of China.

Chicken samples

A total of 1 029 individuals from 44 domesticated chicken populations were used in our study. The 44 chicken populations were distributed over three-quarters of China's administrative divisions (26/34), representing most agro-ecological zones in China

(**Figure 1a and Supplementary Table S1**). These areas encompassed diverse environmental conditions, including variations in temperature, solar radiation, precipitation, land cover and soil properties. Among these populations, a total of 735 blood samples from 20~30 female chickens per population were collected for 28 populations, while another 294 individual sequencing data from 16 chicken populations were retrieved from previous studies (Gu et al., 2020; Wang et al., 2020; Huang et al., 2020; Shi et al., 2023; Zhi et al., 2023; Li et al., 2023; Xu et al., 2023). Genomic DNA was extracted from the blood of the samples using Qiagen DNeasy Blood & Tissue kit.

Genome sequencing, variant calling and quality control

DNA integrity was verified prior to the library construction. Paired-end libraries with insert sizes of 350 bp were constructed and sequenced on the BGISEQ-500 platform (BGI, Shenzhen, China) to generate 150 bp paired-end reads. Low quality reads were filtered using fastp (v.0.20.1) (Chen, 2023), and the clean reads were aligned to the chicken reference genome (GRCg7b) using BWA (v.0.7.17) (Li & Durbin, 2010) with the default parameters. The Genome Analysis Toolkit (GATK v.4.6.0.0) 'MarkDuplicatesSpark' module was subsequently used to filter out potential PCR duplicate reads (McKenna et al., 2010). The resulting alignments were reordered using Picard and then indexed using SAMtools (v.1.13) (Li et al., 2009). SNP calling was processed according to the best practices for GATK by invoking 'HaplotypeCaller', 'CombineGVCFs' and 'GenotypeGVCFs' modules. We set a minimum quality score of 30 for both bases and mapped reads to call variants, and then the common SNP dataset was subjected to filter with the following rigorous criteria using the GATK

'VariantFiltration' module: (a) quality by depth < 5.0, (b) mapping quality score < 40.0, (c) strand odds ratio > 4.0, (d) Fisher strand > 60.0, (e) MQRankSum < -12.5, (f) ReadPosRankSum < -8.0, and (g) any three SNPs clustered in a 10 bp window. The high-quality SNP sets of 1 029 birds was filtered using PLINK v.1.90 software (Chang et al., 2015) with the following parameters: sample call rate > 90%, SNP call rate > 90%, minor allele frequency (MAF) > 0.05 and Hardy–Weinberg equilibrium *p*-value < 1e-5. After these steps, a total of 7 276 708 SNPs distributed across 39 autosomes in 1 022 chickens (seven chickens were filtered out due to low sample call rate) were obtained for downstream analyses. Genotypes were subsequently phased and imputed using Beagle (v.4.1) (Browning et al., 2021) with default parameters to obtain genotyping rate and accuracy of 100%.

Population genomic analysis

The imputed genome data were first pruned with the "indep-pairwise 25 5 0.2" using PLINK v.1.90, yielding 1 415 317 independent SNPs for population structure analysis. Principal components (PC) analysis was carried out with the pruned SNPs using the "pca" in PLINK v.1.90, and the first two PCs were extracted and visualized using R software. We used ADMIXTURE v.1.3.0 (Alexander et al., 2009) with default parameters to investigate population structure across all samples, with the number of clusters (K) being set from 2 to 45. By using pruned SNPs. We also constructed the Neighbor-Joining (NJ) phylogenetic tree based on the distance matrix using FastME 2.0 software (Lefort et al., 2015), and the visualization of tree was performed using Interaction Tree Of Life (iTOL, https://itol.embl.de). Linkage disequilibrium (LD)

decay was assessed using PopLDdecay (Zhang et al., 2019). For the isolation-by-distance (IBD) analysis, VCFtools (v.0.1.15) (Danecek et al., 2011) was applied to calculate the population differentiation coefficient (F_{ST}) using 50 kb sliding windows with 10 kb increments at each step. The matrix of $F_{ST}/(1-F_{ST})$ and the matrix of geographic distance (km) among populations were used for performing the Mantel tests, implemented in the vegan R package (Oksanen et al. 2018) with the significance being determined based on 9 999 permutations.

Environmental data acquisition and analysis

For each of the 44 chicken populations, the central coordinates of their original habitats were recorded, based on conservation farms or marketplaces. To better capture the local environmental conditions, nine additional geographic coordinates were selected within 2 km grids surrounding each population site using Google Earth Pro v.7.3. This resulted in 440 coordinate points from all populations, which were considered as "occurrence points" to predict the potential geographic distribution for chicken populations. Based on the occurrence points, we obtained 87 agroclimatic variables using the raster R package, including 67 bioclimatic, eight landcover, one topographical and 11 pedologic variables (**Supplementary Table S2**) from the WorldClim (Fick & Hijmans, 2017), WorldCover (Zanaga et al., 2022) and ISRIC (https://www.isric.org/) databases at a spatial resolution of 30 arc-seconds. These variables reflect ecological factors relevant to chicken survival, such as climate and elevation (physiological tolerance), soil characteristics (food resources), and land cover (food availability and predation risk). After removing two variables lacking records

within the occurrence ranges, the remaining 85 environmental variables were fed to R package ENMeval for model evaluation (Kass et al., 2021). Models were tuned with the *maxnet* algorithm and *randomkfold* partitions, testing regularization multipliers (RM) from 0.5 to 6.0 (increments of 0.5) and six feature class (FC) combinations: L = linear, Q = quadratic, H = hinge, P = product, T = threshold. The best FC and RM selected based on ΔAICc were applied in the Maxent software (Fourcade et al., 2014) to estimate the contribution of each variable. To improve the model's robustness and biological interpretability, variables with contributions > 1% were retained for further analysis. To reduce collinearity, pairwise correlations among variables were calculated, and the variable with lower contribution was removed for correlated pairs ($r > 0.6$) (Gheyas et al., 2021). Variables with variance inflation factor (VIF) < 5 calculated using the usdm R package, were retained. The relative contribution and permutation importance of remained variables were re-evaluated by Maxent with best model parameters output from ENMeval.

In the Maxent, 75% of occurrence points were used for the model training and the rest 25% were used as model testing data according to Guisan and Thuiller (Guisan & Thuiller, 2005). The remained environmental variables were used for predicting the suitability of chicken populations. The model was replicated 10 times according to the subsample method to verify the prediction accuracy, and the area under the receiver operating characteristic curve (AUC) was used to evaluate the model's performance, where an AUC value of 0.9~1.0 indicates a perfect prediction and value < 0.5 represent a random prediction (Elith et al., 2011). The Jackknife test of each environmental

variable was used to evaluate its importance. Finally, the environmental variables with contribution rate larger than 4% were considered as important factors for exploring the distribution of indigenous chickens (Gheyas et al., 2021). The continuous probability of distribution was divided into four groups using the reclassify tool in ArcGIS (v.10.2) according to the previous studies: 0-0.05 (unsuitable due to severe environmental limitations), 0.05~0.33 (moderate environmental limitations), 0.33~0.66 (suitable without environmental limitations), 0.66~1.00 (optimal environmental conditions for local chicken populations) (Yu et al., 2022).

The environmental values for each population were calculated by the raster R package. By ranking the value of each environmental variable, populations ranking in the top three or bottom three for each variable were designated as “extremely high” and “extremely low” groups, respectively. For some variables, all populations with zero values of the variable were divided into “extremely low” group. Then, the niche differentiation analyses were performed between two groups using the ENMTools (v1.0.4) R package (Warren et al., 2021). Niche overlap was quantified by Schoener’s D and Hellinger’s I, where a value of 0 suggests no overlap and 1 indicates complete overlap.

Genotype-environment association analysis

In this study, we applied two different methods to identify genomic variants associated with environmental variables. Firstly, we conducted redundancy analysis (RDA) to identify genomic variants showing significant relationships with multiple environmental axes. RDA has been shown to be one of the best genotype-environment

association analysis (GEA) approaches for large SNP sets and exhibits low false positive rates (Capblancq & Forester, 2021). Herein, the partial RDA analyses was implemented using the R package vegan, with environmental variables specified as explanatory factors and genotypes as response variables, while controlling for latitude, longitude, and the first three principal components (PCs) derived from PCA. The significance of model constraints was assessed by using ANOVA like permutation test (anova.cca). Then, we selected SNPs that distributed in the two-tails of normal distribution of SNP loadings using a standard deviation cutoff of 3 along each significant RDA axis as significant environment-associated variants, which were further screened out based on largest third-quantile value among all variables. Second, we employed a univariate latent-factor linear mixed model (LFMM) to detect associations between whole genomic variants and the eight environmental variables. The method is good at controlling type I errors by estimating residual structure. LFMM analysis was implemented in the R package LEA (v.3.3.2) (Caye et al., 2019). By setting the number of ancestry clusters inferred from ADMIXTURE as latent factors, we computed regularized least squares estimates for LFMM to account for population structure using a ridge penalty. Then, significance values for the associations were calibrated by using the genomic control method. The calibrated *P*-values of each variant were adjusted for multiple tests using a false discovery rate (FDR) correction, with FDR < 0.01 considered significant.

The associated SNPs that detected by both RDA and LFMM for each variable were united as the adaptive variants, which were used to investigate the role of geography

and environment in shaping spatial genetic variations, compared to neutral (the 1 415 317 LD-pruned SNPs) variants according to Sang et al (Sang et al., 2022). Mantel tests were used to test for correlations among $F_{ST}(F_{ST}/1-F_{ST})$, geographic distance (isolation-by-distance) and environmental distance (isolation-by-environment), in which environmental distance was represented by Euclidean distance of all scaled environmental variables. The significance of correlations was determined using 9 999 permutations performed in the R package vegan.

Selection signature analysis for chicken populations from extremely high and low environmental conditions

Selection signature analysis (SSA) was performed by using two complementary approaches, in which fixation index (F_{ST}) analysis compared the frequency of differential alleles between population groups in extremely high versus extremely low environmental conditions (Sang et al., 2022), and cross population extended haplotype homozygosity (XP-EHH) was designed to identify genomic regions under selection by comparing haplotype homozygosity between two groups (Sabeti et al., 2007). F_{ST} value was calculated in VCFtools (v0.1.15) using the Weir and Cockerham method (Weir & Cockerham, 1984), with a 50 kb sliding window and a 10 kb step size. The weighted F_{ST} were then standardized to zF_{ST}, and genomic regions within top 1% zF_{ST} values were retained for further analysis. Prior to the XP-EHH analysis, genotypes of high- and low-environmental value groups were phased using beagle (v4.1) software, respectively. XP-EHH analyses were carried out using selscan (v2.0.3) software (Szpiech & Hernandez, 2014) to obtain the XP-EHH score for each SNP, and then the

standardized XP-EHH (XPEHH_std) and its absolute value of XPEHH_std (|s_XPEHH|) were calculated. In order to integrate the results of the two methods, |s_XPEHH| were averaged across the same windows used in the F_{ST} analysis, and genomic regions within top 1% |s_XPEHH| values were considered as selected regions for further analysis. Based on the zF_{ST} and |s_XPEHH|, the correlations between F_{ST} and XP-EHH were calculated using Pearson correlation method. To mitigate potential confounding effects of population structure and demographic history on SSA outcomes, we combined three chicken populations each in high and low-environment value group (Gheyas et al., 2021). Furthermore, the SHRUBS and WATER were excluded from SSA due to that the environmental values between two groups were very similar, and the number of populations with extreme values exceeded three per group.

Function annotation for candidate genes

For the genotype-environment association analysis (GEA), the adaptive SNPs were annotated using SnpEff (v.4.3) (Cingolani, 2022). Genomic regions located in the 100 kb upstream and downstream of adaptive SNPs were annotated using Bedtools (v2.30.0) with the GRCg7b genome annotation. For SSA, genomic regions identified by F_{ST} and XP-EHH analysis were annotated using Bedtools (v2.30.0) with GRCg7b genome annotation. Finally, the common genes identified by both GEA and SSA were considered as candidate genes that associated with environmental adaptation. Function annotation and enrichment analysis of genes were performed with Gene Ontology (GO) knowledge base and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database using clusterProfiler package (Wu et al., 2021) implemented in R. False

discovery rate (FDR) method was used to adjust the *P*-values from the hypergeometric test. GO terms and KEGG pathways with FDR < 0.05 were significantly enriched.

Identification of functional SNPs

To further characterize adaptive genomic regions, we used VCFtools (v0.1.15) to calculate nucleotide differences (θπ), genetic diversity (Tajima's D) and fixation index (F_{ST}) statistics. The basic parameters were set to a 15 kb sliding window with a 15 kb step for calculating θπ. Tajima's D was calculated in a 15 kb window and F_{ST} was calculated for every single SNP harbored in candidate genomic region. Linkage disequilibrium (LD) and haplotype block structure were evaluated with r^2 statistic and visualized with LDBlockShow (Dong et al., 2021) for SNPs and R package locuszoomr for genomic regions (Lewis & Wang, 2025), respectively. The haplotype network analysis was performed by using R package geneHapR by removing individuals with heterozygous sites in the region (Zhang et al., 2023).

To investigate the tissue-specific regulatory pattern of selected variants, we integrated 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) which covered 23 tissues including adipose, burse, bone marrow, cecum, cerebellum, colon, cortex, duodenum, gizzard, hypothalamus, ileum, jejunum, liver, lung, muscle, proventricular, spleen, eggshell gland, testis, trachea and thymus from publicly available database (Pan et al., 2023). The selected

SNPs were also annotated with signal peaks including four histone modifications (H3K4me3, H3K4me1, H3K27ac and H3K27me3), CTCF, ATAC-seq and DNase-seq. Meanwhile, gene expression profiles across tissues were retrieved from ChickenGTEx database (Hou et al., 2025). To ensure consistency with our genomic dataset, the genomic coordinates of regulatory elements and functional features were converted from GRCg6a reference genome to GRCg7b reference genome using LiftOver with custom chain files, which were generated through whole-genome alignment using the transanno toolkit (https://github.com/informationsea/transanno) with minimap2chain module.

RESULTS

Population structure and genetic diversity of local chicken populations

In the present study, a total of 1 029 chickens from 44 populations were investigated across China, with sampling sites categorized into six agroclimatic areas including the North (N), Northwest (NW), Middle (M), Southeast (SE), Southwest (SW) and Tibet (T) (**Figure 1a**). Whole-genome resequencing data of these chickens were generated using next generation sequencing (NGS) techniques on BGI and Illumina platform. By aligned onto the chicken reference genome (GRCg7b), average > 99% of the clean reads were mapped with an average depth of 14.40 × and coverage of 96.23% (**Supplementary Table S3**). Among 35 739 816 high-quality SNPs across the 1 029 samples, the majority of substitutions were A/G (37.20%) and C/T (37.19%) transitions (**Supplementary Figure S1a**). Rare SNPs that with MAF < 5% accounted for 60.41% of total variants, with most ranging between 0.01 and 0.05 (**Supplementary Figure**

S1b), suggesting population-specific pattern of SNPs distribution. we performed principal component analysis (PCA) for the 1 022 qualified samples based on 1 415 317 pruned autosomal SNPs. Along the first PC, the clusters were changed from negative to positive values with geographic regions ordered as Southwest (SW), Southeast (SE), Middle (M), North (N), Northwest (NW) and Xizang (T). The PC2 largely reflected the difference between chicken populations in North and South China (**Figure 1b**). Through deep inspection, we found chicken populations in SE could be separately clustered, whereas several chicken populations were overlapped with each other in other regions, such as the LDC, DGC and BSC in N; HDC and NYC in NW; LSC, GSC and MCL in M; NDC, HTC, GXS, HXC and WCC in SE; XZJ and LZC in Xizang (**Supplementary Figure S2**). The neighbor-joining (NJ) tree revealed similar results (**Supplementary Figure S3**).

To investigate population structure, we inferred ancestry proportions for all samples using ADMIXTURE. The number of ancestors that ADMIXTURE converged with minimum cross-validation (CV) error was 27. We found that BJC, BYC and JYC (N), LHC and XHC (M), PDC, LYC, XSC, JHW and HTC (SE), LSF, DWS and GZC (SW), and RKC (T) showed high genetic purity, whereas all NW chicken populations exhibited high admixture. Interestingly, populations in central China (LSC, GSC, MCL, HJC, DAC, and HLC) also showed similar admixture patterns, suggesting that extensive gene exchange existed throughout the history of these breeds (**Figure 1c**). We also assumed six agroclimatic areas as the six ancestries for all chickens and found that LHC (M), WBD (M), XSC (SE), DWS (SW) and RKC (T) were genetically

354 independent. The remaining ancestry was showing significant different patter between
355 chicken populations in Southwest China and that in other part of China
356 (**Supplementary Figure S4**). Moreover, the substantial admixture increased the
357 genetic diversity of chicken populations, among which the physical distance that
358 linkage disequilibrium (LD) decreased to half value (half-life) for pure populations (254
359 nt) were higher than that of populations with high admixture (194 nt) (**Supplementary**
360 **Figure S5**).

361 The genetic differentiation among the 44 populations was assessed using weighted
362 F_{ST}. The F_{ST} value ranging from 0.01 to 0.25 (mean = 0.09) indicated a moderate
363 genetic differentiation among populations. As showing in the **Figure 1d**, the highest
364 F_{ST} (0.25) was found between LHC and WBD. We detected that populations with high
365 purity were significantly divergent with other populations (average F_{ST} = 0.07-0.16),
366 and the divergence among admixed populations showed area-specific differentiation.
367 For example, the LDC and DGC (0.04), the NYC and XWC (0.04), the LSC, GSC,
368 MCL, HJC, DAC and HLC (0.01-0.06), the HXC, GXS and WCC (0.01-0.03), the PDG
369 and CSL (0.04) and the XZJ and LZC (0.03) were slightly divergent (F_{ST} = 0.01-0.06)
370 within each area compared to that among areas (F_{ST} = 0.01-0.19). This was consistent
371 with the clusters identified in the PCA. Notably, although CLZ exhibited an admixed
372 genetic structure, it was highly divergent with all other populations with F_{ST} ranging
373 from 0.09 to 0.22, exceeding divergence seen in some purebreds like JYC, LYC, JHW,
374 HTC, LSF and RKC from other populations. This suggested that CLZ represented an
375 independent genetic lineage shaped by human migration and unique environmental

conditions in the edge of the Qinghai–Xizang Plateau. Furthermore, the genome-wide estimates of nucleotide diversity (π) revealed that the π of CLZ was lower than almost all populations except for WBD and XSC (**Supplementary Figure S6**). Combined with high half-life LD decay of CLZ, it is highly recommended that CLZ is a potentially new genetic resource undergone a long-term natural selection and domestication.

Environmental variables contribute to ecological niche of chicken populations

Ecological niche model (ENM) was applied to predict the distribution of each chicken population, following the conceptual model of environmental factors governing species distributions (Soberón & Osorio-Olvera, 2023). Firstly, we built the models with different tuning parameters and evaluated their performance and detected that the ΔAICc reached zero, when the feature class (FC) was set as linear, quadratic and hinge (LQH) with RM = 1 (**Supplementary Figure S7**). Then, based on the contribution rate for each variable, we selected a total of 22 environmental variables with contribution rate larger than 1% for the further analysis (**Figure 2a**). After removing variables highly correlated ($r > 0.6$) and VIF > 5, nine variables including BARE (areas with minimal vegetation), BIO_2 (mean diurnal range), CLAY (proportion of clay particles), CROPLAND (land used for crops), SRAD_11 (solar radiation in November), PREC_08 (precipitation in August), TREES (land covered by tree foliage), SHRUBS (woody plants) and WATER (areas covered by lakes and rivers) were fed to the ENMeval for automated tuning and evaluations of ENM. The optimal settings for ENM included five FCs (linear, quadratic, hinge, product and threshold) and a regularization multiplier value of 0.5 (**Supplementary Figure S8**), suggesting

that the optimized model had a better fit and transferability to reduce overfitting of the chicken distribution model compared to the default parameters (FC =LQHPT, RM =1). After running 10 replicates under the optimized parameter settings, the average AUC values for Chinese local chicken reached 0.95, and average omission rate was closed to the predicted omission (**Supplementary Figure S9**), indicating the high accuracy and effectiveness of the optimized MaxEnt model in predicting the suitability of local chickens. The Jackknife test revealed that BARE, BIO_2, CROPLAND, PREC_08 and SRAD_11 that with test gain > 0.4 were the primary variables influencing the distribution of local chickens, whereas the CLAY contributed little in both regularized gain and AUC (**Supplementary Figure S10**). The relative contributions of each variable to the Maxent model showed that SRAD_11 contributed largest to the ecological niche of chicken populations (28%), followed by the CROPLAND (27%), WATER (12.5%), PREC_08 (9.1%), BIO_2 (6.5%), TREES (5.2%), BARE (4.7%), SHRUBS (4.2%), and the CLAY possessed the smallest contribution (2.9%) and permutation importance (1.7%) (**Figure 2b)**.

Based on the nine environmental variables, we found the highly suitable regions for chicken populations were located in the Southwest and Southeast China (**Figure 2c**), where rich biodiversity and a long history of traditional poultry farming were previously recognized and inferred for survival and formation of diverse local breeds (Zhuang et al., 2023). Moreover, we explored the suitability map for each chicken population by excluding the CLAY due to its negligible effect. The AUC values of all distribution models exceeded 0.9 in the suitability analysis, and we found that the eight

remained environmental variables exerted different impacts on the chicken distributions (**Supplementary Figure S11**). For example, the ecological niche of BJC (N) and XSC (SE) was mainly affected by BIO_2, which make them adapted to the north China (high diurnal range) and the east coast of China (low diurnal range), respectively (**Figure 2d and 2g**). CLZ (NW), LHC (M), DWS (SW) and RKC (T group) adapted to mountain areas, north China plain, south humid-hot areas and Xizang environment, respectively, driven by dominant environment factors (**Figure 2e, 2f, 2h and 2i**).

By focusing on the eight remained variables, we selected three populations each with extremely high and low value, respectively (**Supplementary Figure S12**), between which niche differentiation were tested. For the BIO_2, the H group included NYC, XWC and SNC with average value of 15.59, and the GXS, PDC and WCC with average value of 7.07 were selected as L group. Similarly, we extracted H and L group for the BARE (0.68 vs. 0.01), CROPLAND (0.62 vs. 0.003), PREC_08 (292.87 vs. 3.33), SRAD_11 (14389.97 vs. 8199.07), TREES (0.92 vs. 0), SHRUBS (0.04 vs. 0) and WATER (0.23 vs. 0), respectively. For all variables, the empirical values of Hellinger's I and Schoener's D were significantly lower than those expected from 100 pseudo-replicated simulations (**Figure 2j**), suggested a clear niche differentiation affected by selected environmental variables. However, observed I and D of SHRUBS and WATER were larger than that of other variables, indicating smaller differentiation. Overall, ENM revealed the environmental diversity of Chinese indigenous chicken habitats and identified eight important environmental drivers of local adaptation. The significant niche difference was in consistent to the genomic divergence between

chicken populations with extremely high and low environmental variables.

Genotype-environment associations analysis

By integrating eight environmental variables with genomic data from 1 022 individuals, we performed genotype-environment association (GEA) analyses to detect the environment-associated genetic variants using RDA and LFMM methods. In the RDA, the overall model was significant ($P < 0.01$) under 500 permutations, explaining 3.04% of the total genetic variance (**Supplementary note 1**). This result was consistent with previous findings in Ethiopia indigenous chickens (Gheyas et al., 2021), supporting the polygenic inheritance of adaptive traits since only a small proportion of genomic variants are expected to be associated with several environmental predictors (Forester et al., 2018). The eight RDA axes in the model were significant with permutation P-value < 0.05, explaining all the variance captured by the RDA model (**Supplementary Figure S13, Supplementary note 1**). The PCoA showed that first two RDAs explained 36.75% (RDA1, 19.12%; RDA2, 17.63%) of the variation of environmental variables. The RDA1 was positively correlated with most variables, including BARE, CROPLAND, PREC_08, SHRUBS, SRAD_11 and WATER, whereas RDA2 was correlated with and TREES. BIO_2 and CROPLAND were pointed to opposite direction, in which BJC was positively correlated with high BIO_2, whereas PDC was positively correlated with CROPLAND. BARE and TREES were pointed to opposite direction, in which TGC was positively correlated with BARE, whereas CLZ was positively correlated with TREES. XHC was found in the areas with highly covered with water land (**Figure 3a**). We identified 35 688 outlier SNPs (1 862~6 549 SNPs per

environmental variable), with correlations ranging from 0.03 to 0.41 (**Supplementary Figure S13, Supplementary Table S4**). Using r > 0.2 as threshold, we obtained 307, 812, 510, 613, 113, 883, 1 693 and 1 436 SNPs for BARE, BIO_2, CROPLAND, PREC_08, SHRUBS, SRAD_11, TREES and WATER, respectively. LFMM was a complementary univariate landscape genomic method to test associations between genomic variants and environment variables, while accounting for background population structure (Frichot et al., 2013). We identified a total of 205~19 126 SNPs were significantly (FDR < 0.01) associated with one or more environmental factors. The largest number of associated SNPs was found in BIO_2, whereas the BARE and SHRUBS were associated with only ~ 200 each. Moreover, a total of 1 190 SNPs affected two or more environmental factors (**Supplementary Figure S14, Supplementary Table S5**).

By overlapping SNPs detected by both RDA and LFMM, we identified 13, 276, 66, 69, 39, 19, 307 and 500 SNPs, corresponding to 90, 1 860, 546, 578, 315, 110, 2 167 and 2 865 genes for BARE, BIO_2, CROPLAND, PREC_08, SHRUBS, SRAD_11, TREES and WATER, respectively (**Figure 3b, c, Supplementary Table S6**). Gene enrichment analysis showed that BARE-associated genes were overrepresented in cytokine-cytokine receptor interaction and metabolic pathways including amino acid metabolism (xenobiotics by cytochrome P450, tyrosine, retinol and pyruvate), glycolysis and fatty acid degradation. CROPLAND-associated genes were overrepresented in biological processes, such as cellular response to chemical stimulus, response to bacterium and chemotaxis. Interestingly, WATER-associated genes were

overrepresented in four same biological processes that identified in the CROPLAND, involving same genes including AvBD (avian β-defensin) gene family (1, 6, 8, 9, 10, 11, 13), *DEFB4A* and *SBDS*. (**Figure 3d, Supplementary Table S7**). The AvBD represented a key family of antimicrobial host defense peptides in birds and might also influence nutrient metabolism under varying nutrient availability, as indicated by CROPLAND (Fu et al., 2023).

We used the 1 283 variable-associated SNPs that were regarded as adaptive variants to explore patterns of isolation-by-distance (IBD) and isolation-by-environment (IBE) compared to neutral variants (excluding adaptive SNPs). We found that genetic divergence calculated by both adaptive and neutral SNPs were not correlated with geographic distance (**Figure 3e**). In contrast, adaptive SNPs showed a significant IBE indicated by partial Mantel tests (**Figure 3f**), whereas the neutral SNPs was not significant, indicating that genetic variation of the adaptive variants was mainly influenced by the environment. Our results suggested that the adaptive SNPs were relatively robust to the confounding effects of population structure and highlight the central role of ecological factors in shaping genomic variation across landscapes.

Selection signatures associated with environmental extremes in indigenous chickens

The niche differential analysis for BARE, BIO_2, CROPLAND, PREC_08, SRAD_11 and TREES suggested that the significant environmental divergence prompted the genomic changes between chicken populations with extremely high and low environmental variables. Given that smaller niche differentiation and more than

three populations in the extreme group, the SHRUBS and WATER were discarded before the selection signature analysis and candidate gene analysis. For the six retained drivers, we applied two methods to detected the potential genomic regions and variants that associated with the environmental changes. First, we performed F_{ST} analysis and extracted top 1% selected regions (N =935) with z*F*st threshold larger than 3.34, 3.44, 3.29, 3.34, 3.37 and 3.42 for BARE, BIO_2, CROPLAND, PREC_08, SRAD_11 and TREES, respectively. These regions were annotated with 548-679 Ensembl genes per variable, the gene sets of which were independent with few overlaps (**Supplementary Figure S15a, Supplementary Table S8**). Function annotation revealed that about half (41.97%~49.30%) of genes lacked with gene symbols. GO and KEGG enrichment showed that MAPK signaling pathway, calcium signaling pathway were significantly (FDR < 0.05) enriched by BARE-associated genes (**Supplementary Figure S15b**). In the XP-EHH analysis, the signatures of individual SNP were merged into 935 windows, in which top 1% selected regions were obtained with |s_XPEHH| threshold exceeding 2.82, 2.89, 2.86, 2.79, 2.80 and 2.83 for the six variables, respectively. These genomic regions were annotated with 425-510 Ensembl genes for 6 variables (**Supplementary Figure S15c, Supplementary Table S9**), in which the BARE-associated genes and the CROPLAND-associated genes were significantly (FDR < 0.05) overrepresented in eight and four KEGG pathways, respectively (**Supplementary Figure S15d**). Among these pathways, alpha-linoleic acid metabolism, linoleic acid metabolism and GnRH signaling pathway were shared by two variables. Similarly, the gene sets of the six variables from XP-EHH were independent with few overlaps, and 41.39%~53.18%

genes lacked gene symbol annotation.

Notably, we observed a modest but significant positive correlation ranging from 0.23 to 0.39 between the F_{ST} and the XP-EHH results (P < 2.2e-16) across the six environmental variables (**Supplementary Figure S16**), consistent with previous studies and confirming the robustness of our analysis (Ma et al., 2015). However, the overlap between the top 1% signatures detected by both methods was low, with significant correlations observed only for BIO_2 and PREC_08 (**Figure 4a**). This highlighted the complementary roles of two methods in detecting significant selection signatures. We further screened top genes defined as those annotated to the top 0.1% selection signatures in either method or top 1% selected genes shared by two methods (**Supplementary Table S10**). This yielded 168, 194, 107, 214, 126 and 184 top genes for BARE, BIO_2, CROPLAND, PREC_08, SRAD_11 and TREES, respectively (**Figure 4b, Supplementary Table S11**). Unlike to the distribution pattern in F_{ST} and XP-EHH, we found that top genes were common to two or more variables, with overlap proportion ranging from 46.26% to 71.73% (**Figure 4b-c**). Spearman correlation analysis indicated that this overlap was not owing to the number of shared chicken populations used in the SSA (**Figure 4d**). Functional enrichment revealed that calcium signaling pathway was significantly enriched by BARE-associated genes including *CHRM2*, *CAMK2G*, *CAMK2D*, *CYSLTR2*, *CACNA1B* and *FLT1*. (**Figure 4e**). Moreover, we mapped these genes to the chicken QTL database and revealed associations with 66 traits, covering growth, health, feeding, fatness, exterior, digestion, egg and meat quality traits, physiology and reproduction traits (**Figure 4f,**

552 **Supplementary Table S12**). Notably, only a small proportion of genes (9-39) were
553 overlapped with known QTLs, suggested that the candidate genes were more potential
554 to affect traits, such as adaptation, since the QTLs mapping was generally conducted
555 for economic traits and those associated with environment changes were rarely reported
556 in chickens (Hu et al., 2022). Few environment-associated genes were overlapped with
557 QTLs, together with few enriched pathways by environment-associated genes
558 confirmed polygenic nature of adaptation in local chicken populations, which were a
559 complex combination of multiple minor gene effects (Sang et al., 2022).

560 **Geographic distribution of genomic variants in the candidate genes**
561 **underlying adaptive responses to climate changes**

562 According to previous analysis results, SRAD_11, PREC_08 and BIO_2 are three
563 most important climatic factors contributing to the ecological niche of local chicken
564 population (**Figure 2b**). SRAD_11, representing total solar radiation in November, was
565 strongly correlated with annual solar radiation (r = 0.59, p = 2.6e-5; **Supplementary**
566 **Figure S17a**), making it a robust indicator of adaptation to ultraviolet and heat stress.
567 We identified *PRMT3*, *EPC2* and *TPK1* associated with SRAD_11 were involved in the
568 regulation of DNA damage and repair, which are crucial in how organisms respond to
569 the solar radiation (Jessulat et al., 2021; Angrand, 2022; Brobbey et al., 2022) (**Figure**
570 **5a**). The common genomic regions identified by SSA and GEA included *MAML2*,
571 *C20H20ORF112*, *POU6F2* and one gene cluster (**Supplementary Table S13**). The
572 gene *MAML2* encoded the Mastermind-like proteins that acted as a vital transcriptional
573 coactivator in the Notch signaling, which was responded to high-dose radiation

574 (Banerjee et al., 2020). Within *MALM2*, we identified 11 SNPs with F_{ST} > 0.5 in high
575 LD (**Supplementary Figure S18a, b)**. We selected an index SNP *rs15536665* (F_{ST} =
576 0.65) to show the alleles frequency pattern and found that chicken populations
577 originated from high SRAD_11 in Xizang and Southwest areas possessed more G allele
578 (**Supplementary Figure S18c**). In the gene cluster, we firstly identified index SNP
579 *rs316099750* that was divergently selected between high and low group (grey shadow).
580 Given that the negative Tajima's D value suggested an excess of rare genetic variants in
581 a population compared to that under neutral evolution, we further focused on two
582 genomic regions with -0.57 and -0.36 of Tajima'D in the low SRAD_11 group in the
583 gene cluster, among which the genomic region ranging from 2 745 001 bp to 2 760 000
584 bp (red shadow in **Figure 5b**) contained significant divergent values across three
585 statistics. The region was located in *UNC80* gene and harbored an index intronic SNP
586 *rs316174149* (F_{ST} = 0.56). The allele frequency patterns showed that chicken
587 populations originated from areas with high SRAD_11 possessed more A allele,
588 whereas the T allele was dominant in chicken populations originated from areas with
589 low SRAD_11 (**Figure 5c, Supplementary Figure S19**). Notably, the Sichuan basin
590 was exposed to less solar radiation, as shown in **Figure 5c**. We then investigated the
591 gene frequency of *rs316174149* for chicken breeds in Sichuan basin including Miyi
592 chicken (MYC), Pengxian Yellow chicken (PXC), Emei Black chicken (EMB), Jiuyuan
593 Black chicken (JYB) and Jinyang silky chicken (JYS) ((Li et al., 2019), frequency of
594 which was consistent with pattern that T allele was dominant in these chicken
595 populations in contrast to that A allele was dominant in the Xizang fowl (TF) and Red

596 jungle fowl (RJF) in tropical areas (**Figure 5d**). Given that most non-coding variants
597 were located in regulatory elements, we further annotated the SNP based on the atlas of
598 chicken regulatory elements (Pan et al., 2023). Five regulatory elements including
599 EnhAWk, ATAC, Repr, ReprWk and Qui were annotated for *rs316174149* across 23
600 tissues (**Figure 5e**). Correspondingly, we also found significant H3K27me3 signals for
601 hypothalamus, duodenum, and eggshell gland, as well as ATAC-seq signals for kidney
602 (**Figure 5f**). Moreover, the expression data from Chicken GTEx (Guan et al., 2025)
603 showed that the *UNC80* was highly expressed in the hypothalamus (**Figure 5g**). *UNC80*
604 encodes a key component of the NALCN sodium channel complex that regulates
605 neuronal excitability and is actively expressed in the nervous systems (Chua et al., 2020;
606 Wie et al., 2020). These suggested that the *rs316174149* might act as a repressor to alter
607 gene expression within the hypothalamus, mediating the long-term adaptation to solar
608 radiation (Vrettos et al., 2025).

609 PREC_08, an indicator of monthly precipitation, was strongly correlated with
610 annual precipitation (r = 0.75, p = 5.8e-9; **Supplementary Figure S17b**), making it a
611 robust marker of adaptation to humid versus arid environments. We identified SSA
612 genes including *ABCB10*, *ANK2*, *KCNA1*, *PRMT1*, *SLC22A15* and *TRAPPC11*
613 associated with PREC_08 were involved in the matter transportations to enhanced
614 utilization of metabolic water in arid-dwelling animals (Rocha et al., 2021) (**Figure 6a**).
615 The common genomic regions identified by SSA and GEA were annotated with *PTPRO*,
616 *SOX5*, *SPOCK1* and four gene clusters (**Supplementary Table S14**). In *PTPRO*,
617 Tajima'D and F_{ST} analyses identified a region (62 835 001 bp ~ 62 840 000 bp) with

618 negative Tajima'D (-0.18) for the high PREC_08 group, contrasting with a positive
619 Tajima's D (1.64) in low PREC_08 group (**Figure 6b**), indicated that an excess of low
620 frequency SNPs was under putative selective sweep in the chickens originated from
621 areas with high precipitation. In this region, 10 SNPs with F_{ST} > 0.25 were detected.
622 The most divergent SNP was an intronic SNP *rs15292135*, which was in a LD block
623 with nearby 6 SNPs (**Figure 6c**). The allele frequency of *rs15292135* showed that C
624 allele was dominant in the chicken populations adapted to the arid environment,
625 whereas T allele was dominant in the chicken populations adapted to the humidity
626 environment (**Figure 6d**). The chromatin state analyses showed that *rs15292135* lied
627 in short regions for most tissues annotated by ATAC, Repr, ReprWk and Qui (**Figure**
628 **6e**). Notably, the key roles of kidney and trachea were confirmed by the significant
629 signals identified in the ATAC-seq data (**Figure 6f**), suggesting that *rs15292135*
630 affected the chromatin accessibility, altering gene regulation and contributed to the
631 adaptation of humid or arid environment. Among four gene clusters, we scanned
632 genomic regions with large difference among θπ, Tajima's D and F_{ST} value on the
633 chromosome 1. In the region, the index SNP *rs3384544359* that possessed highest F_{ST}
634 (0.46) was located in the region ranging from 64 545 001 bp to 64 560 000 bp. However,
635 the MAF of *rs3384544359* was low (0.13), resulting in almost fixed A allele in many
636 chicken populations (**Supplementary Figure S20**). Therefore, we focused on another
637 region ranging from 64 590 001 bp to 64 605 000 bp (red shadow in **Supplementary**
638 **Figure S21a**), in which we identified a set of SNPs with high F_{ST} were under putative
639 selection sweeps. The index SNP *rs312280416* with a MAF of 0.4 was in high LD with

surrounding SNPs, locating in the *AEBP2* (**Supplementary Figure S21b**). The allele frequency pattern showed that the frequency of G allele was growing up when the precipitation increased across chicken populations, suggested that the function of SNP and candidate gene *AEBP2* were involved in the adaptation of humidity (**Supplementary Figure S21c**).

For the BIO_2, we also identified a number of genes that previously reported to be associated with temperature adaptation (**Supplementary Figure S22a**). For example, the gene *DIO3* identified for the BIO_2 was a deiodinase enzyme that converts the inactive thyroid hormone T4 to its active form T3. Cold exposure decreases hepatic *DIO3* activity in chicks, leading to less T3 degradation. The increased T3 levels contribute to enhanced metabolic rate and heat production, which are essential for thermoregulation in cold environments (Amaz and Mishra, 2024). Using same pipeline, we obtained 2 intersected genes (*EVL*, *NCOR2*), 1 lncRNA (*ENSGALG00010004931*) and 6 gene clusters (**Supplementary Table S15**). We identified a 30 kb genomic region in the *NCOR2* was differentially selected between high and low groups across three indices, with Tajima'D =-0.71 in low BIO_2 group (**Supplementary Figure S22b)**. Accompanied with F_{ST} value of single SNP, we detected that an intronic SNP *rs316512683* possessed the highest F_{ST} (0.31) and was in a medium LD (0.46~0.66) with surrounding SNPs with $F_{ST} > 0.25$ (**Supplementary Figure S22c**). The A allele of *rs316512683* was mainly distributed in the north region that is characterized by large diurnal temperature difference, whereas the G allele was almost fixed in other chicken populations (**Supplementary Figure S22d, e**). Moreover, we detected SNP

rs316512683 as a repressor in the heart via H3K27me3 modification, contributing to the thermal adaptation (**Supplementary Figure S22f, g**). Among gene clusters, we identified candidate SNP *rs14836021* was harbored in the intron of *EPS8*. As a signaling adaptor, it was involved in the EGFR pathway that promote cell survival and tissue repair under temperature-induced damage (Tito et al., 2025). The A allele of *rs14836021* was almost fixed in the region that is characterized by small diurnal temperature difference, while the G allele increased in the north region (**Supplementary Figure S23**). Another SNP *rs317631326* was an index SNP located in the *DSEL* and was in high LD with surrounding SNPs. The A allele of *rs317631326* was dominant in NYC, XWC, SNC and BJC with extremely high BIO_2 (**Supplementary Figure S24**). *DSEL* encoded enzymes for dermatan sulfate biosynthesis, a key extracellular matrix component influencing cell behavior, metabolism, and homeostasis (Mizumoto & Yamada, 2022).

Geographic distribution of genomic variants underlying local landcover adaptation

For the landcover, we identified three physical materials at the land surface including CROPLAND, TREES and BARE non-negligibly affect the animal local adaptations (**Figure 2b**). For the CROPLAND, a few top genes were annotated in the selection signatures analysis (**Figure 7a, Supplementary Table S11)**, which were rarely reported to be associated with cropland coverage and availability. However, genes including *TPK1*, *COA3*, *G6PC1*and *LEPROT* were demonstrated to play important role in metabolism, digestion and energy homeostasis. For example, *G6PC1*

encodes glucose-6-phosphatase, crucial for maintaining blood glucose levels between meals (Claxton et al., 2022). After overlapping with associated genes identified in landscape genomic analysis, we detected a continuous genomic regions ranging from 51 548 866 bp to 51 626 768 bp on chromosome 1 **(Supplementary Table S16)**. The θπ and Tajima's D indicated that chicken populations located in the areas with high CROPLAND were under intensive selection (**Figure 7b**). We further screened a 10 kb region with most significant value across three indices. Within this region, an index SNP *rs794437951* with highest F_{ST} (0.65) was in high LD with eight surrounding SNPs with F_{ST} > 0.5. The allele frequency pattern showed that the index SNP was almost fixed to T allele for most chicken populations (**Supplementary Figure S25**), which probably due to the low minor allele frequency (MAF). Therefore, we profiled the MAF of nine SNPs, among which an intronic SNP *rs732137070* harbored in *ENSGALG00010014829* carried highest MAF (0.26), and the surrounding SNPs were in high LD with *rs732137070* (r^2 = 0.53~0.74) (**Figure 7c, d**). The chicken populations originated from areas with high CROPLAND possessed more G allele of *rs732137070*, while the A allele of *rs732137070* was dominant in chicken populations originated from mountain areas with few croplands (**Figure 7e, f**). The chromatin state annotation revealed that *rs732137070* lied in the EnhAWk of hypothalamus and testis, suggested transcriptional activation at this site in two organs (**Supplementary Figure S26**). However, no any regulatory peak was detected for the SNP.

For the TREES, we identified *ITGB2* that belonged to integrin-beta gene family (**Figure 8a, Supplementary Table S11)**, in which *ITGB1* was previously reported to

706 associated with land use pattern by RDA analysis in chickens (Gheyas et al., 2021),
707 suggesting that the cell-to-cell and cell-to-extracellular matrix interactions regulated by
708 integrin-beta were essential for individuals adapted to the specific environment. Using
709 same pipeline, we identified six genes and four gene clusters distributing on the
710 chromosome 1, 4, 10 and 23 **(Supplementary Table S17**). Some genes were involved
711 in the neural development and signaling transduction (*FOXP2*, *TRPA1*, *PRKCE*, and
712 *NRXN3*), which contribute to the ability in sensor, communicating, mating and seeking
713 in the dense tree areas. For example, *TRPA1* was demonstrated to be participates in
714 sensory processes, such as cold sensation (Zhang et al., 2022) and hearing (Vélez-
715 Ortega et al., 2023), which were essential for living in the forest. In the *TRPA1*, we
716 identified SNP *rs317006675* in the intron was under selection (**Supplementary Figure**
717 **S27a**). The C allele of *rs317006675* was significantly dominant in chicken populations
718 from high TREES group and significantly decreased in chicken populations from areas
719 with few trees (**Supplementary Figure S27b**). We then refined the four gene clusters,
720 in which we identified most divergently selected regions on the chromosome 4
721 (**Supplementary Figure S28**). By focusing on genomic region with highest F_{ST} (red
722 shadow in **Figure 8b**), we found a large divergence ($\Delta\theta\pi$ = 0.0023 and ΔTajima's D =
723 2.81) between high and low TREES group. The region harbored 17 SNPs with F_{ST} >
724 0.25. The index SNP *rs314594662*, a synonymous variant in the *NXT2*, was in high LD
725 with surrounding eight SNPs. Chicken populations originated from high trees coverage
726 possessed more G allele, while the A allele was dominant in chicken populations
727 originated from northwest areas with few trees (**Figure 8c, d**). Given that the nine SNPs

located in the gene *NXT2* were in high LD, we further analyzed the haplotypes between high and low TREES group. The 9 SNPs formed 5 unique haplotypes across 62 samples. Among these haplotypes, the haplotype *CCGTGAGGA* (H001) was dominant in terms of total amount and frequency in high TREES group, whereas the SNP sites in low TREES group were heterozygous and haplotypes were sporadic (**Figure 8e**), confirmed that *NXT2* was under intensive selection in chickens to adapted to the environment with dense coverage of trees. Moreover, the chromatin state and regulatory annotation revealed that candidate SNPs might be active enhancers within chromatin accessibility and DNA methylation region in bone marrow, trachea, kidney, liver, heart, gizzard, ileum, colon and cecum (**Supplementary Figure S29**).

For the BARE, we identified a bunch of top genes like *CACNA1B*, *COL11A1*, *MGAT4C*, *SLC4A10, SLC25A20* was associated with starvation and dehydration that usually happened in BARE environment (Rocha et al., 2021; Rocha et al., 2023) in the SSA (**Supplementary Figure S30a, Supplementary Table S11**). However, no common genes were screened across four methods, whereas a total of 31 genes were identified by any three of four methods (**Supplementary Figure S30b**). Notably, the common genes identified in the GEA were rarely under intensive selection by F_{ST} (n=6) and XPEHH (n=1), indicated polygenic inheritance of chicken adaption to the BARE environment. Nevertheless, the genomic regions under intensive selection were from natural and artificial interventions. We further focused on common SSA genes that overlapped with RDA or LFMM, from which we obtained three genes and one gene cluster in the RDA and one gene, one lncRNA and two gene clusters in the LFMM,

respectively (**Supplementary Table S18**). By scanning individual SNP among these candidates, we found significant selection signatures in *PALLD* across three indices in RDA. The intronic SNP *rs313027723* harbored in the *PALLD* was divergently selected between extremely high and low BARE group (**Supplementary Figure S30c**). The SNP was in a completed LD block with surrounding SNPs (F_{ST} > 0.25), and the haplotype analysis showed that 2 haplotypes were existed in the populations. The haplotype CCAATAG was abundant and dominant in low BARE group, whereas the haplotype TGGGCGA was less belonging to the high group (**Supplementary Figure S30d, e**). The G allele of *rs313027723* was increasingly important in the chicken located in the areas with extremely high bare ground (**Supplementary Figure S30f**). For LFMM, the candidate region was mapped in the chromosome 3. An intronic SNP *rs316823963* located in the *GPR63* was screened out and in high LD (r2 = 0.61) with an index SNP *rs15388384* that located in the 3'UTR (**Supplementary Figure S31a**). Functionally important, we further showed the allele frequency pattern of *rs15388384*, where the C allele was increased with increased bare ground (**Supplementary Figure S31b**). Furthermore, the two SNP *rs313027723* and *rs15388384* were annotated as regulatory SNPs in the brain and heart, respectively (**Supplementary Figure S32**), suggested the distinct tissue-specific regulatory patterns.

DISCUSSION

Coping with environmental changes is common to all species. Unlike studies that focus on intraspecific adaptive variation to mitigate extinction risks in wildlife (Sang et al., 2022), we leveraged the abundant genetic diversity and adaptive variations to

772 cultivate environment-resilient livestock and poultry for sustainable global food supply
773 (Boettcher et al., 2014). Compared to commercial chicken strains used in the intensive
774 farming, indigenous populations exhibit superior adaptive traits, such as environmental
775 resilience, foraging ability, welfare living (free range and organic product) and disease
776 resistance, providing valuable resources for new breeding goals that met changes in
777 climate, husbandry environment and customer demand (Gheyas et al., 2021). In this
778 study, we collected and generated genomic variants data for over 1 000 Chinese
779 indigenous chickens, representing the largest population dataset spanning all
780 agroclimatic zones of China to date. Based on the large-scale population genomic data,
781 we systematically analyzed the genetic diversity and structure of 44 chicken
782 populations, which confirmed the purity of some populations like BYC (Wang et al.,
783 2023), JYC (Yue et al., 2025), LHC (Wang et al., 2023), as well as the admixed
784 genomic component of populations like CSL (Xu et al., 2023), WCC (Qi et al., 2024)
785 previously reported by using whole genome resequencing data. The large-scale
786 population genomic analysis also enlarged the admixture of NDC (Huang et al., 2020),
787 CHC (Zhong et al., 2022). Regionally, Xizang populations (SNC and RKC) formed a
788 genetically distinct cluster, likely shaped by prolonged geographic isolation and
789 independent domestication history (Li et al., 2023). In addition, chicken populations
790 from central China (LSC and GSC in Henan province, MCL and HJC in Hubei province,
791 DAC and HLC in Hunan province) displayed similar genomic structure, consistent with
792 findings in local pig populations (Du et al., 2024), indicated that local pig and chicken
793 populations experienced similar demographic histories and human interventions, such

as selective breeding, migration and ecologically diverse habitats (Wiener & Wilkinson, 2011).

Animal adaptation was a long-term systematic process to cope with local stressors like heat, drought, disease and nutritional challenges, through a combination of genetic selection and physiological changes over generations. Traditionally, it is difficult to dissect the genetic basis of environmental adaptation due to the complexity of environment posing as selection pressures. With aid of ecological methods, we firstly disentangle and identify the important environmental factors of adaptation from a large array of environmental variables. Then, by integrating large-scale genomic data, we focused on the important factors to identify the associated genomic regions using multiple complementary genomic approaches. These were effectively and robustly applied in plant and animal, avoiding false positive errors and offering more reliable and explainable results (Gheyas et al., 2021; Wu et al., 2024; Teng et al., 2025). Notably, our analysis refined solar radiation (SRAD_11), cropland use (CROPLAND), rainfall (PREC_08), temperature difference (BIO_2), trees coverage (TREES) and bare ground (BARE) as most important environmental drivers, which was more representative and significant than other similar studies based on only climate variables (Niu et al., 2024; Liu et al., 2025; Liu et al., 2025; Zhao et al., 2025) or a smaller geographical region (Gheyas et al., 2021). These variables were further demonstrated to be key factors for comparative genomic analysis by niche difference test, which was a result of genetic selection over centuries by environmental conditions and human purposes within livestock species (Velado-Alonso et al., 2022). Moreover, the genetic divergence

calculated by adaptive SNPs showed a significant association with environmental distances based on these variables, reinforcing the important and independent selection of environmental variables in shaping adaptive differentiation, as also reported in livestock like sheep (Lv et al., 2014) and pig (Liu et al., 2025). Among a bunch of adaptive genes, the function annotation confirmed their roles in environmental adaptation, such as BARE-associated genes that were overrepresented in energy metabolism (amino acid, fatty acid and glycolysis) pathways that help animals manage energy stores to survive periods of food and water scarcity (Rocha et al., 2021; Cheng et al., 2023; Rocha et al., 2023).

Environmental adaptation within species generally accompanied by micro-evolution happened in the genome (Bomblies & Peichel, 2022). By leveraging genomic data using multiple methods, we presented robust evidence for adaptive genes including well-known gene *PTPRO* and *NCOR2* for precipitation and diurnal temperature difference, respectively, as well as novel gene *UNC80*, *CSF2RB*, *NXT2*, and *PALLD* for solar radiation, cropland use, trees coverage and bare ground, respectively. *NCOR2* was a nuclear receptor corepressor that played a crucial role in regulating thermogenesis in brown adipose tissue and maintaining body temperature functionally associated with BIO_2 (Richter et al., 2022). *PTPRO* modulates the glomerular pressure and filtration rate in the kidney, affecting conservation and loss of water to adapted the humid and dry environments (Wharram et al., 2000; Ozaltin et al., 2011). *CSF2RB* encoded receptors that bound to ligands of colony stimulating factor (CSF), which was conserved in the vertebrates (Mun et al., 2020). The knockout of CSF receptors

838 significantly caused fatty changes in the liver in the zebrafish (Meng et al., 2021),
839 suggested that the alteration of gene modified the metabolism of chickens to adapted
840 different crops availability. In addition, the pleiotropic genes associated with BIO2,
841 SRAD_11 and BARE indicated that interaction of climatic and landcover factors at
842 genomic level shaped the local adaptation, which need more investigation by
843 considering human-mediated management. The variants suffered from intensive
844 selection in the candidate genes were mostly located in non-coding region, which
845 implied that the regulatory elements played dominant role allowing more subtle and
846 fine-tuned evolutionary changes with fewer negative side effects (Haygood et al., 2010;
847 Bomblies & Peichel, 2022). By integrating Chicken GTEx data, we found these SNPs
848 possessed different chromatin state across tissues, for example, acting as repressor in
849 hypothalamus through H3K27me3 modification to respond the solar radiation;
850 adaptation to the humidity/arid through altering chromatin accessibility (ATAC-seq) in
851 kidney and trachea. These indicated that the non-coding SNPs generally regulated gene
852 function in a tissue-specific manner (Albert & Kruglyak, 2015; Khurana et al., 2016),
853 on the other hand, the rapid adaptation controlled by these variants were prevalent
854 across local chicken populations (Agrawal & Hastings, 2023; Henschen et al., 2023;
855 Durkin et al., 2024; Nair et al., 2024). Therefore, the corresponding interactions
856 between key genes and genomic variants explained their potential regulatory
857 mechanism for environmental adaptation.

858 Specifically, our findings provide genomic resources to further understand the
859 genetic basis and regulatory mechanisms of adaptations in domestic animals and lays a

foundation for future research with potential application in genetic resource conservation and animal breeding under environmental change, as well as may accelerate comparative studies on adaptation in other species. Nevertheless, several limitations should be acknowledged. First, we found substantial adaptive genes lacked gene symbols, which limited the function interpretation. A more well-assembled genome and completed genome annotation were needed for the function annotations. Second, the relatively low sequencing depth and the presence of modest breed admixture may reduce the statistical power, potential leading to the omission of some causal variants. Third, our study lacked more comprehensive experiment, such as CRISPR-based functional validation and population-specific regulatory omics data for investigating SNP-specific regulatory mechanisms. In addition, we should recognize that adaptive signals are generally not stable and shifting rapidly due to the climate change and modern genomic breeding practices. Future studies incorporating higher-depth sequencing, larger sample size and field experiment areas, complementary multi-omics technologies, and validation experiment could strengthen the applicability of the conclusions, and enhance the translational relevance for agricultural and biomedical research.

In conclusion, we dissected the agroclimatic factors and identified key environment factors driving the ecological niche partitioning of Chinese indigenous chickens. By incorporating landscape genomic analysis, comparative genomic analysis and atlas of chicken regulatory elements, we proposed candidate genes and SNPs that affected the environmental adaptation in a tissue-specific manner. Furthermore, we compiled a

comprehensive list of potential variant-gene associations to serve as a valuable resource for future functional studies and breeding programs.

ACKNOWLEDGEMENTS

The authors thank Yong Zhang for insightful comments for the manuscript. The authors were grateful to the participants and all staffs who supported this study and the anonymous reviewers for their critical review of our manuscript. This work was supported by the National Key Research and Development Program of China (2021YFD1200302), the China Agriculture Research System of MOF and MARA (CARS-40) and the Agricultural Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (ASTIP-IAS16).

COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ETHICS STATEMENT

All animal procedures were approved by the Animal Care and Use Committee of Institute of Animal Science, Chinese Academy of Agricultural Sciences (No. IAS2021-49) and performed in accordance with the relevant guidelines and regulations set by Ministry of Agriculture and Rural Affairs of the People's Republic of China.

AUTHORS' CONTRIBUTIONS

J. Y. and X. J. contributed equally to this work. J. Y., W. H. and J. C.: Conceptualization, Supervision, Project administration and Writing – review & editing. J. Y., X. J. and X. H.: Formal analysis, Methodology, Validation, Visualization, Software, Data curation, Writing – original draft. J. Y., X. J., Y. Li, A. Ni, G. C., W. H., N. Z., S. H., Y. S., G. Y. and J. C.: Investigation, Materials collection and preparation.

DATA AVAILABILITY

The genomic data of indigenous chickens has been deposited in National Genomics Data Center (NGDC, https://www.cncb.ac.cn/) at PRJCA046667.

REFERENCES

Agrawal AA, Hastings AP. 2023. Tissue-specific plant toxins and adaptation in a specialist root herbivore. *Proceedings of the National Academy of Sciences of the United States of America*. 120:e1992716176.

Amaz Al S, Mishra B. 2024. Embryonic thermal manipulation: a potential strategy to mitigate heat stress in broiler chickens for sustainable poultry production. *Journal of Animal Science and Biotechnology,* 15:75.

Albert FW, Kruglyak L. 2015. The role of regulatory variation in complex traits and disease. *Nature Reviews Genetics,* 16:197-212.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research,* 19:1655-1664.

Angrand P. 2022. Structure and function of the Polycomb repressive complexes PRC1 and PRC2. *International of Journal of Molecular Science,* 23 (11):5971.

Balakrishnan KN, Ramiah SK, Zulkifli I. 2023. Heat shock protein response to stress in poultry: a review, *Animals,* 13: 2.

Banerjee D, Barton SM, Grabham PW, et al. 2020. High-Dose Radiation Increases Notch1 in Tumor Vasculature. *International Journal of Radiation Oncology-Biology-Physics,* 106:857-866.

Boettcher PJ, Hoffmann I, Baumung R, et al. 2014. Genetic resources and genomics for adaptation of livestock to climate change. *Frontiers in Genetics* 5:461.

Bomblies K, Peichel CL. 2022. Genetics of adaptation. *Proceedings of the National Academy of Sciences of the United States of America*. 119:e2122152119.

933 Brobbey C, Liu L, Yin S, et al. 2022. The role of protein arginine methyltransferases in
934 DNA damage response. *International Journal of Molecular Sciences,* 23(17):9780.
935 Browning BL, Tian X, Zhou Y, et al. 2021. Fast two-stage phasing of large-scale
936 sequence data. *American Journal of Human Genetics,* 108:1880-1890.
937 Campbell Q, Bedford JA, Yu Y, et al. 2025. Agricultural landscape genomics to
938 increase crop resilience. *Plant Communications,* 6:101260.
939 Capblancq T, Forester BR. 2021. Redundancy analysis: A Swiss Army Knife for
940 landscape genomics. *Methods in Ecology and Evolution,* 12:2298-2309.
941 Caye K, Jumentier B, Lepeule J, et al. 2019. LFMM 2: Fast and accurate inference of
942 gene-environment associations in genome-wide studies. *Molecular Biology and*
943 *Evolution,* 36:852-860.
944 Chang CC, Chow CC, Tellier LC, et al. 2015. Second-generation PLINK: rising to the
945 challenge of larger and richer datasets. *Gigascience.* 4:7.
946 Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and
947 deduplication using fastp. *iMeta.* 2:e107.
948 Cheng J, Peng X, Li H, et al. 2023. Similar adaptative mechanism but divergent
949 demographic history of four sympatric desert rodents in Eurasian inland.
950 *Communications Biology* 6:33.
951 Chua HC, Wulf M, Weidling C, et al. 2020. The NALCN channel complex is voltage
952 sensitive and directly modulated by extracellular calcium. *Science Advances,*
953 6:eaaz3154.
954 Cingolani P. 2022. Variant annotation and functional prediction: SnpEff. *Methods in*

Molecular Biology 2493:289-314.

Claxton DP, Overway EM, Oeser JK, et al. 2022. Biophysical and functional properties of purified glucose-6-phosphatase catalytic subunit 1. *Journal of Biological Chemistry,* 298:101520.

Danecek P, Auton A, Abecasis G, et al. 2011. The variant call format and VCFtools. *Bioinformatics.* 27:2156-2158.

Dong S, He W, Ji J, et al. 2021. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Briefings in Bioinformatics,* 22:bbaa227.

Du H, Zhou L, Liu Z, et al. 2024. The 1000 Chinese Indigenous Pig Genomes Project provides insights into the genomic architecture of pigs. *Nature Communications,* 15:10137.

Durkin SM, Ballinger MA, Nachman MW. 2024. Tissue-specific and cis-regulatory changes underlie parallel, adaptive gene expression evolution in house mice. *PLoS Genetics,* 20:e1010892.

Eda M. 2021. Origin of the domestic chicken from modern biological and zooarchaeological approaches. *Animal Frontiers* 11:52-61.

Elith J, Phillips SJ, Hastie T, et al. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions.* 17:43-57.

Ellegren H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution,* 29:51-63.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces

977 for global land areas. *International Journal of Climatology.* 37:4302-4315.
978 Forester BR, Lasky JR, Wagner HH, et al. 2018. Comparing methods for detecting
979 multilocus adaptation with multivariate genotype-environment associations.
980 *Molecular Ecology.* 27:2215-2233.
981 Fourcade Y, Engler JO, Rödder D, et al. 2014. Mapping species distributions with
982 MAXENT using a geographically biased sample of presence data: a performance
983 assessment of methods for correcting sampling bias. *PLoS One.* 9:e97122.
984 Frichot E, Schoville SD, Bouchard G, et al. 2013. Testing for associations between loci
985 and environmental gradients using latent factor mixed models. *Molecular Biology*
986 *and Evolution.* 30:1687-1699.
987 Fu J, Zong X, Jin M, et al. 2023. Mechanisms and regulation of defensins in host
988 defense. *Signal Transduction and Targeted Therapy,* 8:300.
989 Gheyas AA, Vallejo-Trujillo A, Kebede A, et al. 2021. Integrated environmental and
990 genomic analysis reveals the drivers of local adaptation in African indigenous
991 chickens. *Molecular Biology and Evolution* 38:4268-4285.
992 Gourdine J, Rauw WM, Gilbert H, et al. 2021. The genetics of thermoregulation in pigs:
993 a review. *Frontiers in Veterinary Science,* 8:770480.
994 Gu J, Liang Q, Liu C, et al. 2020. Genomic analyses reveal adaptation to hot arid and
995 harsh environments in native chickens of China. *Frontiers in Genetics* 11:582355.
996 Guan D, Bai Z, Zhu X, et al. 2025. Genetic regulation of gene expression across
997 multiple tissues in chickens. *Nature Genetics* 57:1298-1308.
998 Guisan A, Thuiller W. 2005. Predicting species distribution: offering more than simple

habitat models. *Ecology Letters* 8:993-1009.

Hariyono DNH, Prihandini PW. 2022. Association of selected gene polymorphisms with thermotolerance traits in cattle - A review. *Animal Bioscience* 35:1635-1648.

Haygood R, Babbitt CC, Fedrigo O, Wray GA. 2010. Contrasts between adaptive coding and noncoding changes during human evolution. *Proceedings of the National Academy of Sciences of the United States of America.* 107:7853-7857.

Henschen AE, Vinkler M, Langager MM, Rowley AA, Dalloul RA, Hawley DM, Adelman JS. 2023. Rapid adaptation to a novel pathogen through disease tolerance in a wild songbird. *PLoS Pathogens.* 19:e1011408.

Holsinger KE, Weir BS. 2009. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nature Reviews Genetics,* 10:639-650.

Hou Y, Zou D, Chu Q, et al. 2025. The ChickenGTEx portal: a pan-tissue catalogue of regulatory variants shaping transcriptomic and phenotypic diversity. *Nucleic Acids Research,* gkaf731, doi: 10.1093/nar/gkaf731.

Hu Z, Park CA, Reecy JM. 2022. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Research,* 50:D956-D961.

Huang X, Otecko NO, Peng M, et al. 2020. Genome-wide genetic structure and selection signatures for color in 10 traditional Chinese yellow-feathered chicken breeds. *BMC Genomics,* 21:316.

Oksanen J, Blanchet FG, Friendly M, et al. 2018. vegan: Community Ecology Package. R package version 2.5, Available from: https://CRAN.R-

project.org/package=vegan.

Jessulat M, Amin S, Hooshyar M, et al. 2021. The conserved Tpk1 regulates non-homologous end joining double-strand break repair by phosphorylation of Nej1, a homolog of the human XLF. *Nucleic Acids Research* 49:8145-8160.

Kass JM, Muscarella R, Galante PJ, et al. 2021. ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution,* 12:1602-1608.

Khurana E, Fu Y, Chakravarty D, et al. 2016. Role of non-coding sequence variants in cancer. *Nature Reviews Genetics,* 17:93-108.

Lai D, Qu X, Wang Z, et al. 2025. Integrative genomic, transcriptomic and epigenomic analysis reveals cis-regulatory contributions to high-altitude adaptation in Tibetan pigs. *Molecular Biology and Evolution,* 42 (7):msaf169.

Lefort V, Richard D, Gascuel O. 2015. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution,* 32:2798-2800.

Lewis MJ, Wang S. 2025. locuszoomr: an R package for visualizing publication-ready regional gene locus plots. *Bioinformatic Advances,* 5:vbaf006.

Leng D, Zeng B, Wang T, et al. 2024. Single nucleus/cell RNA-seq of the chicken hypothalamic-pituitary-ovarian axis offers new insights into the molecular regulatory mechanisms of ovarian development. *Zoological Research*, 45(5): 1088-1107. DOI: 10.24272/j.issn.2095-8137.2024.037

Li D, Li Y, Li M, et al. 2019. Population genomics identifies patterns of genetic

diversity and selection in chicken. *BMC Genomics*. 20:263.

Li H, Handsaker B, Wysoker A, et al. 2009. The sequence alignment/Map format and SAMtools. *Bioinformatics*. 25:2078-2079.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 26:589-595.

Li S, Zhang X, Dong X, et al. 2023. Genetic structure and characteristics of Tibetan chickens. *Poultry Science,* 102:102767.

Liu Y, Xu Y, Li G, et al. 2025. Local climate adaptation in Chinese indigenous pig genomes. *Animals*. 15:16.

Liu Z, Zhang N, Wen Z, et al. 2025. Genomic insights into the origin, high fecundity and environmental adaptation of Hu sheep. *Advanced Science,* 12:e06492.

Lv F, Agha S, Kantanen J, et al. 2014. Adaptations to climate-mediated selective pressures in sheep. *Molecular Biology and Evolution,* 31:3324-3343.

Ma Y, Ding X, Qanbari S, et al. 2015. Properties of different selection signature statistics and a new strategy for combining them. *Heredity*. 115:426-436.

McKenna A, Hanna M, Banks E, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research,* 20:1297-1303.

McManus CM, Lucci CM, Maranhão AQ, et al. 2022. Response to heat stress for small ruminants: Physiological and genetic aspects. *Livestock Science* 263:105028.

Meek MH, Beever EA, Barbosa S, et al. 2023. Understanding local adaptation to prepare populations for climate change. *BioScience,* 2025:36-47.

Meng H, Shang Y, Cheng Y, et al. 2021. Knockout of zebrafish colony-stimulating factor 1 receptor by CRISPR/Cas9 affects metabolism and locomotion capacity. *Biochemical and Biophysical Research Communications,* 551:93-99.

Mi X, Feng G, Hu Y, et al. 2021. The global significance of biodiversity science in China: an overview. *National Science Review,* 8:nwab032.

Mizumoto S, Yamada S. 2022. The specific role of dermatan sulfate as an instructive glycosaminoglycan in tissue development. *International Journal of Molecular Sciences,* 23.

Mun SH, Park PSU, Park-Min K. 2020. The M-CSF receptor in osteoclasts and beyond. *Experimental and Molecular Medicine,* 52:1239-1254.

Nair VD, Pincas H, Smith GR, et al. 2024. Molecular adaptations in response to exercise training are associated with tissue-specific transcriptomic and epigenomic signatures. *Cell Genomics,* 4:100421.

Niu Y, Li Y, Zhao Y, et al. 2024. Whole-genome sequencing identifies functional genes for environmental adaptation in Chinese sheep. *Journal of Genetics and Genomics,* 51:1278-1285.

Ozaltin F, Ibsirlioglu T, Taskiran EZ, et al. 2011. Disruption of PTPRO causes childhood-onset nephrotic syndrome. *American Journal Human Genetics,* 89:139-147.

Padhi MK. 2016. Importance of indigenous breeds of chicken for rural economy and their improvements for higher production performance. *Scientifica.* 2016:2604685.

Pan Z, Wang Y, Wang M, et al. 2023. An atlas of regulatory elements in chicken: A

resource for chicken genetics and genomics. *Science Advances,* 9:eade1204.

Qi L, Xiao L, Fu R, et al. 2024. Genetic characteristics and selection signatures between Southern Chinese local and commercial chickens. *Poultry Science* 103:103863.

Richter HJ, Hauck AK, Batmanov K, et al. 2022. Balanced control of thermogenesis by nuclear receptor corepressors in brown adipose tissue. *Proceedings of the National Academy of Sciences of the United States of America.* 119:e2089691177.

Rocha JL, Godinho R, Brito JC, et al. 2021. Life in deserts: the genetic basis of mammalian desert adaptation. *Trends in Ecology & Evolution,* 36:637-650.

Rocha JL, Silva P, Santos N, et al. 2023. North African fox genomes show signatures of repeated introgression and adaptation to life in deserts. *Nature Ecology & Evolution,* 7:1267-1286.

Sabeti PC, Varilly P, Fry B, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature.* 449:913-918.

Sang Y, Long Z, Dan X, et al. 2022. Genomic insights into local adaptation and future climate-induced vulnerability of a keystone forest tree in East Asia. *Nature Communications* 13:6541.

Savolainen O, Lascoux M, Merilä J. 2013. Ecological genomics of local adaptation. *Nature Reviews Genetics,* 14:807-820.

Shi S, Shao D, Yang L, et al. 2023. Whole genome analyses reveal novel genes associated with chicken adaptation to tropical and frigid environments. *Journal of Advanced Research,* 47:13-25.

Soberón J, Osorio-Olvera L. 2023. A dynamic theory of the area of distribution. *Journal*

of Biogeography 50:1037-1048.

Szpiech ZA, Hernandez RD. 2014. selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution,* 31:2824-2827.

Teng Y, Li W, Wang X, et al. 2025. Integrating population genomics and environmental data to predict adaptation to climate change in post-bottleneck Tibetan macaques. *Science Advances,* 11:eadw0562.

Tian S, Zhou X, Phuntsok T, et al. 2020. Genomic analyses reveal genetic adaptations to tropical climates in chickens. *iScience,* 23:101644.

Tito C, Masciarelli S, Colotti G, et al. 2025. EGF receptor in organ development, tissue homeostasis and regeneration. *Journal of Biomedical Science.* 32:24.

Velado-Alonso E, Morales-Castilla I, Gómez-Sal A. 2022. The landscapes of livestock diversity: grazing local breeds as a proxy for domesticated species adaptation to the environment. *Landscape ecology.* 37:1035-1048.

Vélez-Ortega AC, Stepanyan R, Edelmann SE, et al. 2023. TRPA1 activation in non-sensory supporting cells contributes to regulation of cochlear sensitivity after acoustic trauma. *Nature Communications* 14:3871.

Vrettos M, Reynolds C, Amar A. 2025. No support for solar radiation as a major evolutionary driver of malar stripes in falcons. *Journal of Avian Biology,* 2025:e03322.

Wang H, Zhao X, Wen J, et al. 2023. Comparative population genomics analysis uncovers genomic footprints and genes influencing body weight trait in Chinese

indigenous chicken. *Poultry Science* 102:103031.

Wang J, Lei Q, Cao D, et al. 2023. Whole genome SNPs among 8 chicken breeds enable identification of genetic signatures that underlie breed features. *Journal of Integrative Agriculture,* 22:2200-2212.

Wang M, Thakur M, Peng M, et al. 2020. 863 genomes reveal the origin and domestication of chicken. *Cell Research* 30:693-701.

Wang M, Li Y, Peng M, et al. 2015. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Molecular Biology and Evolution,* 32:1880-1889.

Wang Q, Li D, Guo A, et al. 2020. Whole-genome resequencing of Dulong Chicken reveal signatures of selection. *British Poultry Science,* 61:624-631.

Wang T, Yang M, Shi X, et al. 2025. Multiomics analysis provides insights into musk secretion in muskrat and musk deer. *GigaScience*, 14: giaf006, https://doi.org/10.1093/gigascience/giaf006

Warren DL, Matzke NJ, Cardillo M, et al. 2021. ENMTools 1.0: an R package for comparative ecological biogeography. *Ecography.* 44:504-511.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution.* 38:1358-1370.

Wharram BL, Goyal M, Gillespie PJ, et al. 2000. Altered podocyte structure in GLEPP1 (Ptpro)-deficient mice associated with hypertension and low glomerular filtration rate. *Journal of Clinical Investigation,* 106:1281-1290.

Wie J, Bharthur A, Wolfgang M, et al. 2020. Intellectual disability-associated UNC80

mutations reveal inter-subunit interaction and dendritic function of the NALCN channel complex. *Nature Communications,* 11:3351.

Wiener P, Wilkinson S. 2011. Deciphering the genetic basis of animal domestication. *Proceedings of the Royal Society B: Biological Sciences,* 278:3161-3170.

Wu T, Hu E, Xu S, et al. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation,* 2:100141.

Wu Z, Chapman MA, Liu J, et al. 2024. Genomic variation, environmental adaptation, and feralization in ramie, an ancient fiber crop. *Plant Communications,* 5:100942.

Xu D, Zhu W, Wu Y, et al. 2023. Whole-genome sequencing revealed genetic diversity, structure and patterns of selection in Guizhou indigenous chickens. *BMC Genomics,* 24:570.

Yu X, Tao X, Liao J, et al. 2022. Predicting potential cultivation region and paddy area for ratoon rice production in China using Maxent model. *Field Crops Research* 275:108372.

Yuan J, Li S, Sheng Z, et al. 2022. Genome-wide run of homozygosity analysis reveals candidate genomic regions associated with environmental adaptations of Tibetan native chickens. *BMC Genomics.* 23:91.

Yue Y, Zhang X, Zhai B, et al. 2025. Whole-genome resequencing revealed genetic structure and specific breed identification loci of Ningxia Jingyuan chicken breed (Gallus gallus). *BMC Genomics.* 26:772.

Zanaga D, Van De Kerchove R, Daems D, et al. 2022. 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221

Zhang C, Dong S, Xu J, et al. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics,* 35:1786-1788.

Zhang H, Wang C, Zhang K, et al. 2022. The role of TRPA1 channels in thermosensation. *Cell Insight,* 1:100059.

Zhang M, Wang S, Xu R, et al. 2023. Managing genomic diversity in conservation programs of Chinese domestic chickens. *Genetics Selection Evolution,* 55:92.

Zhang R, Jia G, Diao X. 2023. geneHapR: an R package for gene haplotypic statistics and visualization. *BMC Bioinformatics,* 24:199.

Zhao X, Zhang J, Wen J, et al. 2025. Genome-environment association analysis reveals climate-driven adaptation of chickens. *Genetics Selection Evolution,* 57:43.

Zhi Y, Wang D, Zhang K, et al. 2023. Genome-Wide genetic structure of Henan indigenous chicken breeds. *Animals.* 13:4.

Zhong H, Kong X, Zhang Y, et al. 2022. Microevolutionary mechanism of high-altitude adaptation in Tibetan chicken populations from an elevation gradient. *Evolutionary Applications,* 15:2100-2112.

Zhuang Z, Zhao L, Zong W, et al. 2023. Genetic diversity and breed identification of Chinese and Vietnamese local chicken breeds based on microsatellite analysis. *Journal of Animal Science,* 101:skad182.

.

Figure Legend

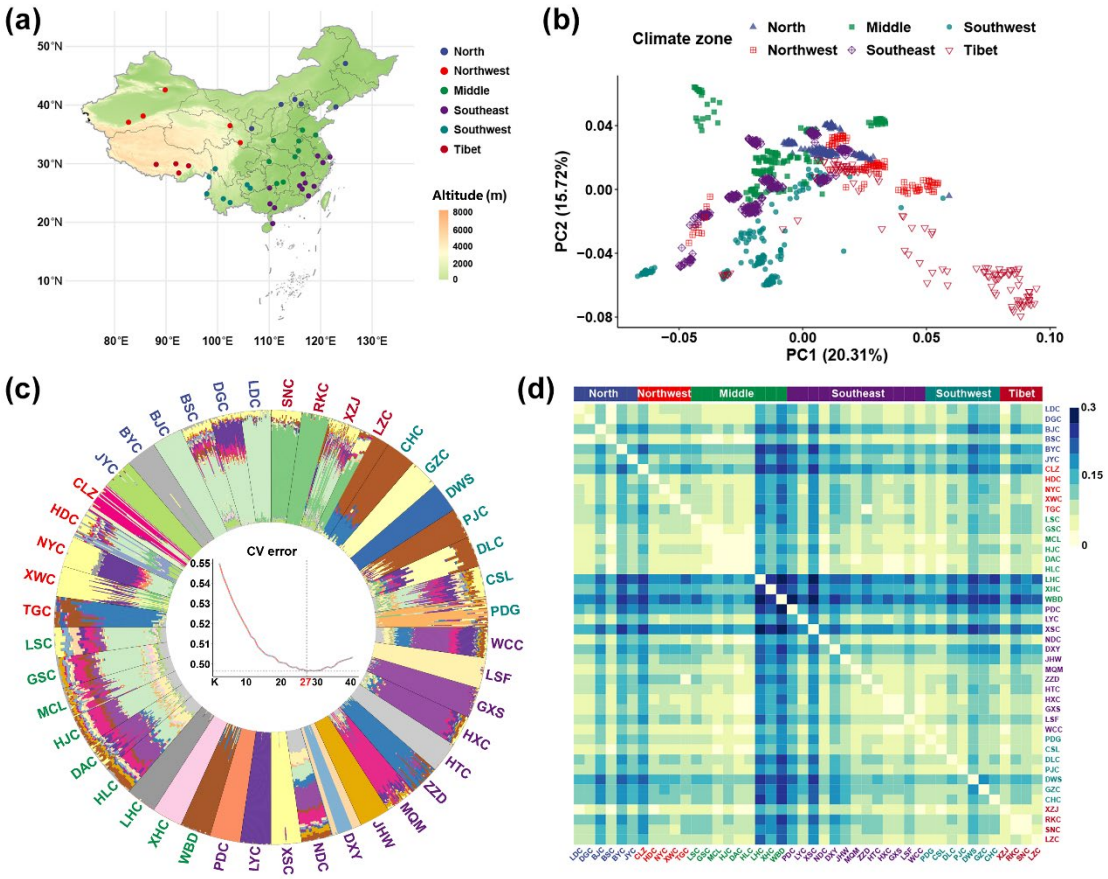

Figure 1. Geographic location and genetic structure of Chinese indigenous chicken populations. (a) Geographic origin of 44 chicken populations. Populations are grouped into six agroclimatic zones: North, Northwest, Middle, Southeast, Southwest, and Tibet. **(b)** Principal component analysis (PCA) of the 44 chicken populations based on 1 415 317 pruned SNPs. **(c)** Ancestry composition of 44 chicken populations inferred by ADMIXTURE analysis (K = 2-45), with the optimal K as 27. **(d)** Genetic divergence among populations measured by pairwise F_{ST}. The chicken populations are ordered by agroclimatic zone.

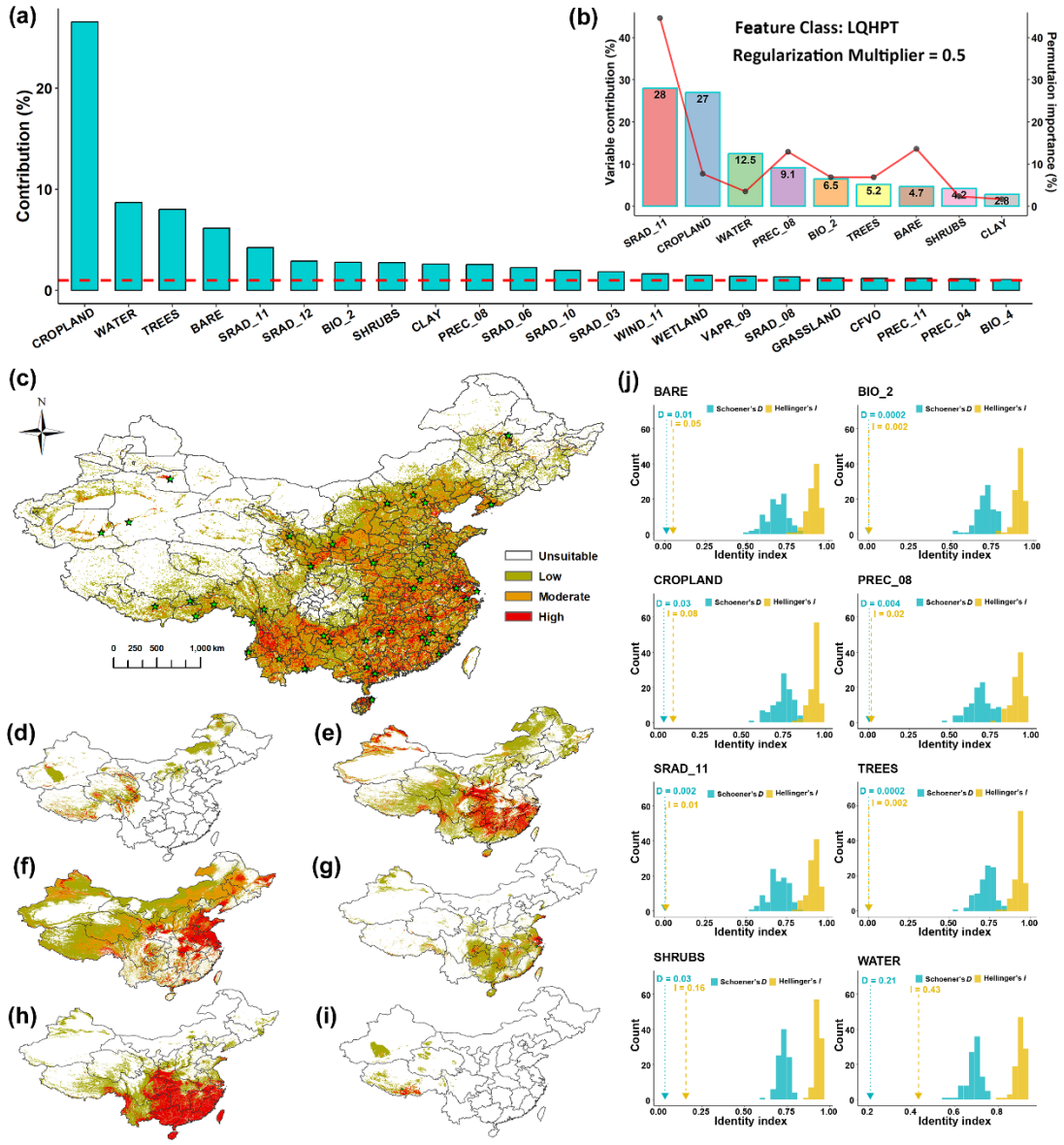

Figure 2. Ecological niche modeling (ENM) for Chinese indigenous chicken populations. (a). Twenty-two environmental variables that contribute more than 1% to the distribution of indigenous chickens predicted by ENM using all 87 environmental variables. **(b)**. Relative contribution and permutation importance of nine environmental variables selected based on correlation ($r \leqslant 0.6$) and variance inflation factor (VIF $\leqslant$ 5). **(c)**. The suitability maps of indigenous chicken populations in China generated by ENM using the nine selected environmental variables. **(d-i).** Suitability maps for representative chicken population: BJC **(d)** in the North, CLZ **(e)** in the Northwest,

LHC **(f)** in the Middle, XSC **(g)** in Southeast, DWS **(h)** in the Southeast and RKC **(i)** in Tibet, predicted from eight key variables (SRAD_11, CROPLAND, WATER, PREC_08, BIO_2, TREES, BARE and SHRUBS). The details of the variables, refer to Supplemental Table 2. Green star denotes sampling location. **(j)**. Niche identity tests between groups of populations from extremes (high vs. low) for each environmental variable. In each group, three or more chicken populations are included based on the value of environmental variable. The arrows indicate the observed niche equivalency, and the histograms represent the simulated (expected) equivalency. All differences between the observed index and the expected index rejected the hypothesis that environmental niches between regions were identical ($p < 0.01$).

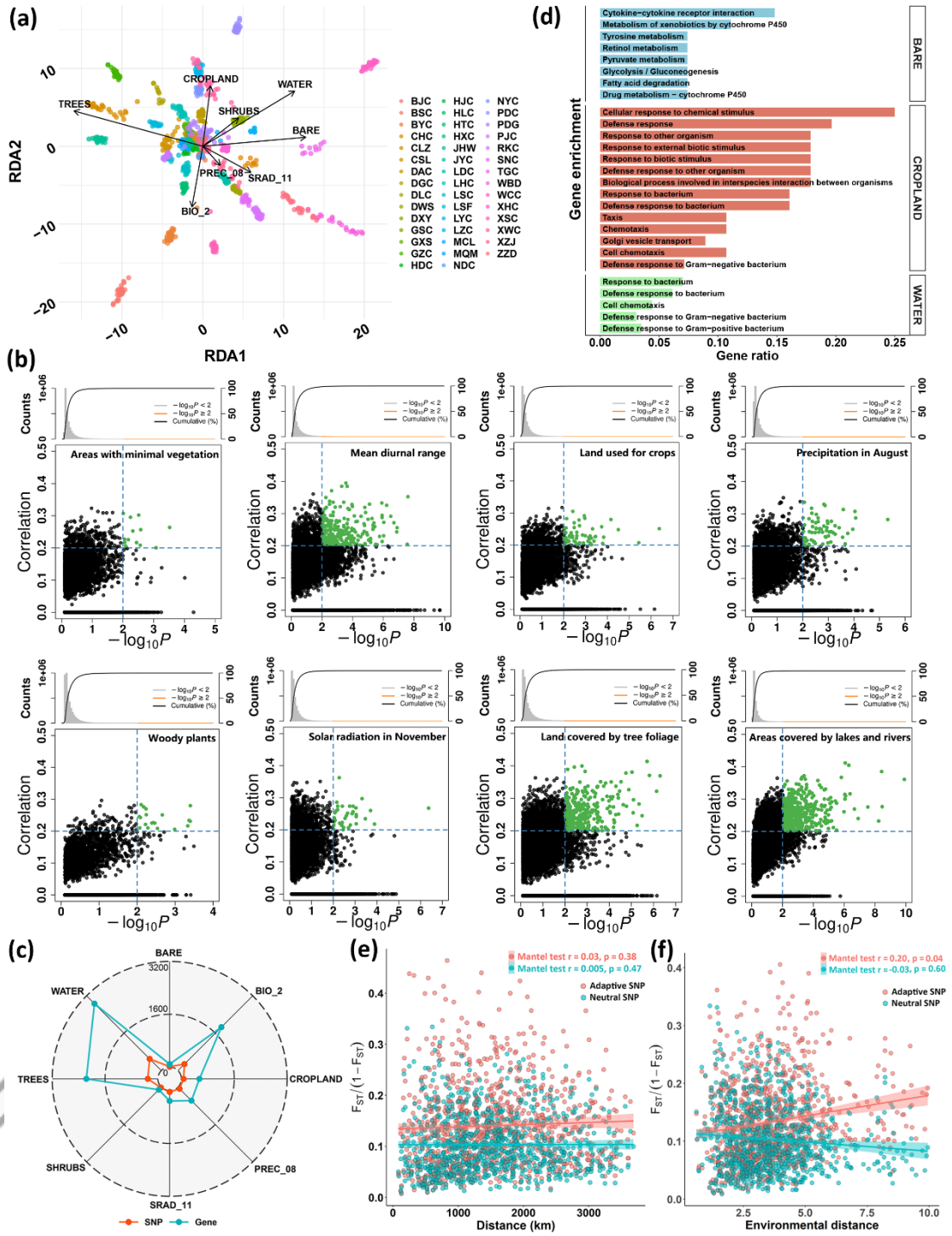

Figure 3. Genotype-environment associations analysis for eight important environmental variables. (a). Principal coordinate analysis (PCoA) for eight important environmental variables based on the RDA1 and RDA2. Arrow length denotes the relative contribution of each environmental factor to the principal components. **(b).** Scatter plot for SNP statistics of RDA correlation coefficients (y-axis)

and LFMM *P*-value (x-axis) for each environmental variable. Dashed lines denote the threshold (r ⩾0.2 and FDR ⩽0.01). Green points indicate candidate SNPs identified by both RDA and LFMM. **(c).** The number of environment-associated SNPs and the number of annotated genes for each environmental variable. **(d).** Significantly enriched GO biological processes and KEGG pathways for genes associated with each environmental variable. **(e).** Isolation-by-distance analyses for local chicken populations based on neutral and adaptive variants, respectively. The shadow of linear regression denotes the 95% confidence interval. **(f).** Isolation-by-environment analyses for local chicken populations based on neutral and adaptive variants, respectively. The shadow of linear regression denotes the 95% confidence interval.

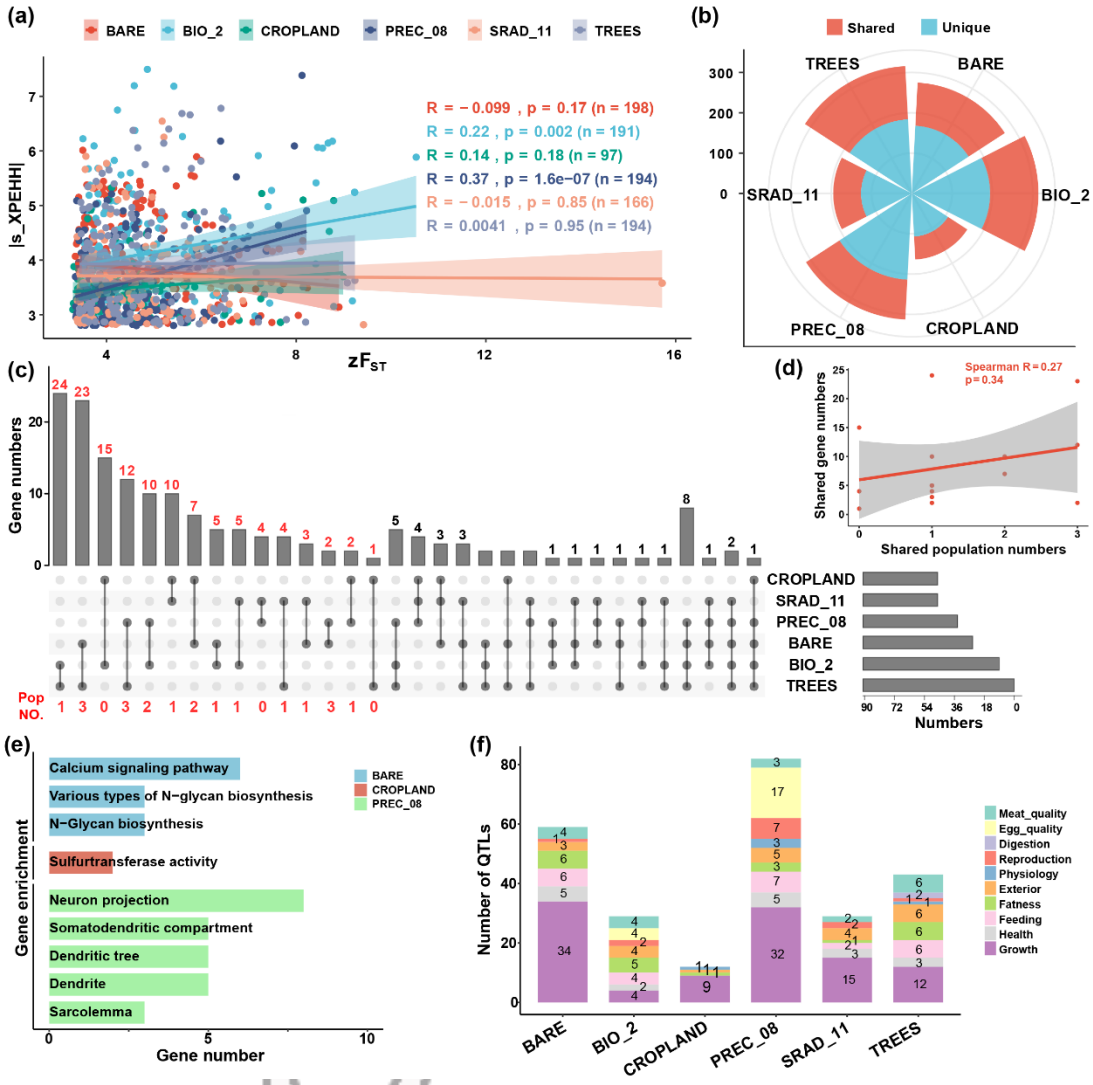

Figure 4. Selection signature analysis between high and low group for six key environmental variables. (a). Correlations between top 1% standardize values of F_{ST} (x-axis) and XPEHH (y-axis). The number of top 1% regions is listed in the parentheses. **(b).** Shared and unique top genes, defined as the top 0.1% selection signatures in either method and top 1% selected genes shared by F_{ST} and XPEHH for each environmental variable. **(c).** Upset plot for top genes across environmental variables. Red numbers mark the shared genes and chicken populations between two environmental variables, used in the correlation analysis **(d)** using Spearman method. **(e).** Significantly enriched GO biological processes and KEGG pathways for top genes associated with each

environmental variable. **(f).** QTL annotation for top genes based on the chicken QTLdb.

Science Press
ZR Zoological Research

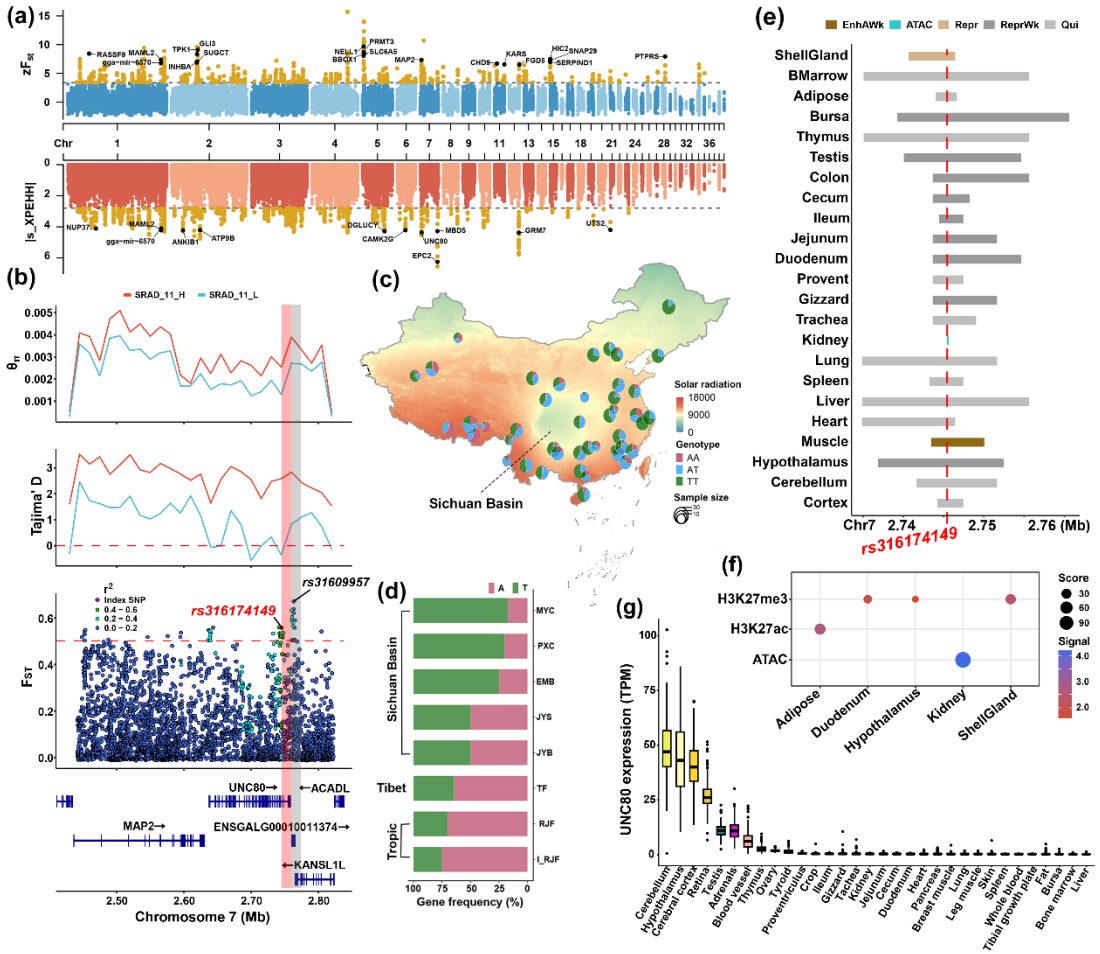

Figure 5. Functional SNPs associated with adaptation to the solar radiation. (a). Genome-wide selective signals of standardized F_{ST} and XP-EHH for chicken populations from high versus low solar radiation regions. Grey dashed lines highlight the top 1% threshold. Orange points represent candidate regions, and black points mark top annotated genes associated with solar radiation. **(b).** Selection signatures for the gene cluster on Chromosome 7 that identified by both gene-environment association analysis and selection signature analysis. θπ and Tajima's D are calculated in a 15 kb window with a 15 kb step. F_{ST} is calculated for each individual SNP. Grey shadow highlights genomic region harbored the index SNP, while the red shadow highlights candidate region harbored SNP *rs316174149* used for the next analysis. **(c).** Genotype frequencies of *rs316174149* across 44 chicken populations. **(d).** Allelic frequencies of

rs316174149 across eight chicken populations distributed in Sichuan basin, Tibet and tropical area. A and T represent the alleles of this SNP. MYC, PXC, EMB, JYS, JYB, TF, RJF and I_RJF denote Miyi chicken, Pengxian yellow chicken, Emei black chicken, Jinyang silky chicken, Jiuyuan black chicken, Tibetan fowl, Red jungle fowl and Red jungle fowl in Indonesian. **(e).** Regulatory activity across 23 different tissues annotated for *rs316174149*. **(f).** Epigenomic annotation for *rs316174149* based on the H3K4me3, H3K4me1, H3K27ac, H3K27me3, CTCF, ATAC-seq and DNase-seq. **(g).** Expression profile of gene *UNC80* across 32 chicken tissues.

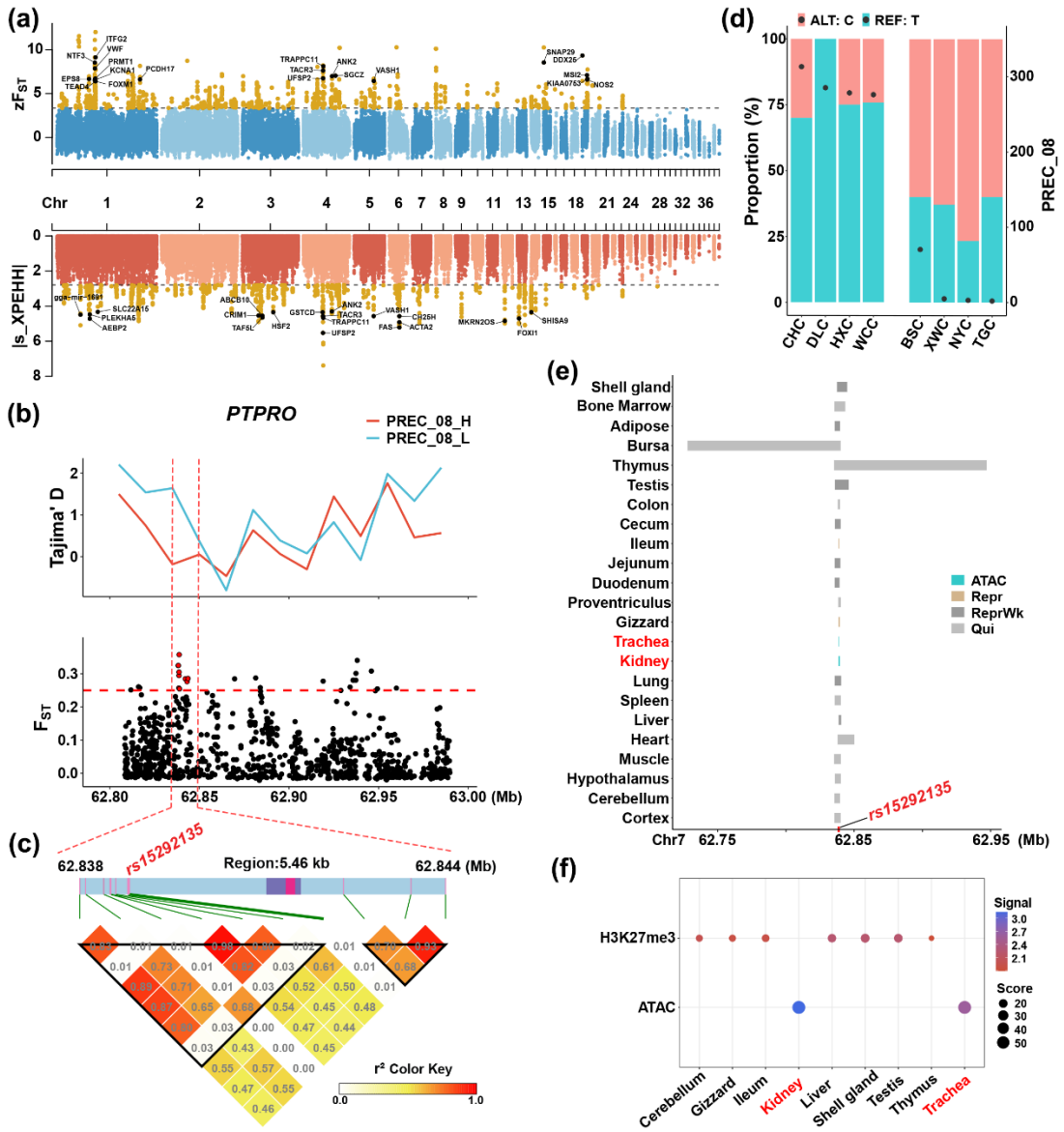

Figure 6. Functional SNPs associated with adaptation to the rainfall. (a). Genome-wide selective signals of standardized F_{ST} and XP-EHH for chicken populations from high- and low-precipitation areas. Grey dashed lines highlight the top 1% quantiles. Orange points represent candidate regions, and black points mark top genes associated with precipitation. **(b).** Selection signatures for gene *PTPRO*, identified by both gene-environment association analysis and selection signature analysis. θπ and Tajima's D are calculated in a 15 kb window with a 15 kb step. F_{ST} is calculated for each individual SNP. **(c).** Linkage disequilibrium (LD) block for SNPs with $F_{ST} \geqslant 0.25$ in the selected

region, shown as mean *r*2 for paired SNPs. **(d).** Allelic frequencies of the associated SNP *rs15292135* across chicken populations from areas with extremely high and low precipitation. **(e).** The regulatory activity across 23 different tissues annotated for *rs15292135*. **(f).** Epigenomic annotation for *rs15292135* based on the H3K4me3, H3K4me1, H3K27ac, H3K27me3, CTCF, ATAC-seq and DNase-seq.

Science Press
ZR Zoological Research

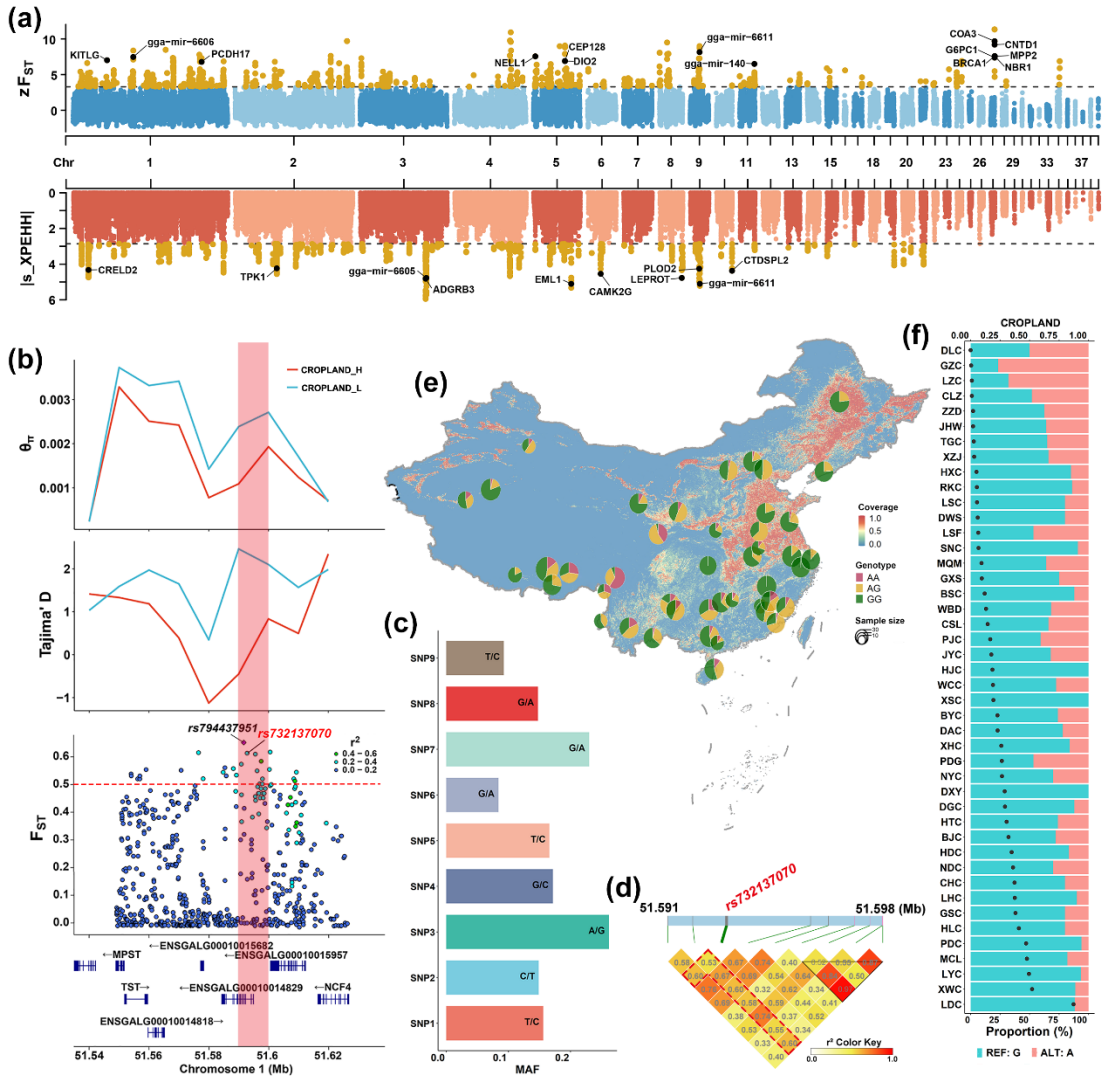

Figure 7. Functional SNPs associated with adaptation to the cropland coverage. **(a).** Genome-wide selective signals of standardized F_{ST} and XP-EHH for chicken populations from areas with extremely high and low cropland use. Grey dashed lines highlight the top 1% threshold. Orange points represent candidate regions, and black points mark top genes associated with cropland coverage. **(b).** Selection signatures for gene cluster that identified by both gene-environment association analysis and selection signature analysis. θπ and Tajima's D are calculated in a 15 kb window with a 15 kb step. F_{ST} is calculated for each individual SNP. **(c).** The minor allele frequency of SNPs surrounding the index SNP *rs794437951*. (d). Linkage disequilibrium (LD) block for

SNPs with F_{ST} ⩾0.5 in the selected region, shown as mean *r*2 for paired SNPs. The selected SNP *rs732137070* is in high LD with surrounding SNPs. (e). SNP genotype frequencies of *rs732137070* across 44 chicken populations in China. **(f).** Allelic frequencies of *rs732137070* across 44 chicken populations.

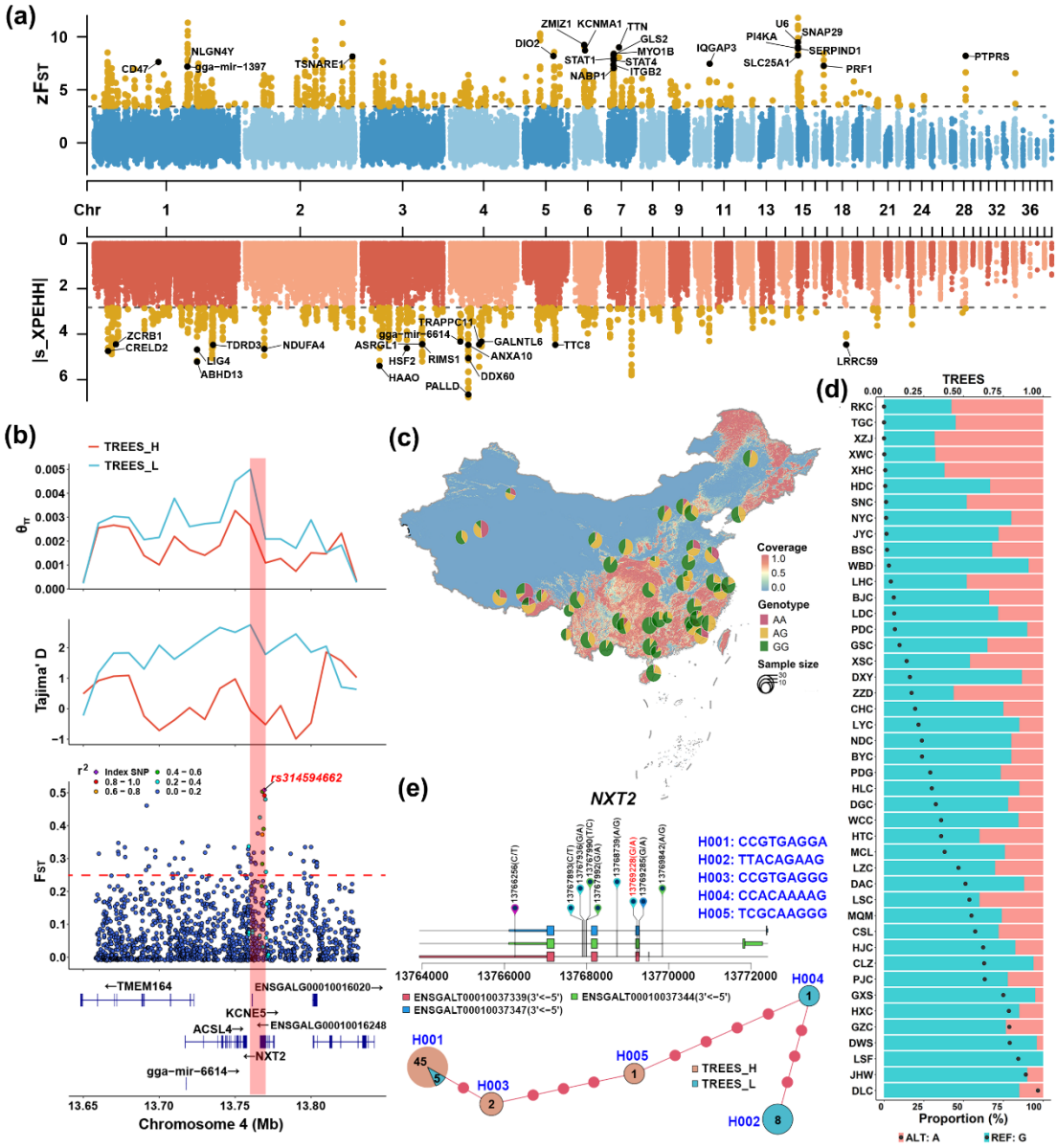

Figure 8. Functional SNPs associated with adaptation to tree coverage environment. (a). Genome-wide selective signals of standardized F_{ST} and XP-EHH for chicken populations from areas with high and low tree coverage. Grey dashed lines highlight the top 1% quantiles. Orange points represent candidate regions, and black points mark top genes associated with tree coverage. **(b).** The selection signatures for gene cluster that identified by both gene-environment association analysis and selection signature analysis. θπ and Tajima's D are calculated in a 15 kb window with a 15 kb

step. F_{ST} is calculated for each individual SNP. The red shadow highlight genomic region under extensive selection with an index SNP *rs314594662* located in gene *NXT2*. **(c).** Genotype frequencies of *rs314594662* across 44 chicken populations. **(d).** Allelic frequencies of *rs314594662* across 44 chicken populations in China. **(e).** Genomic position of candidate SNPs and a haplotype network in gene *NXT2*. The symbols on the line between haplotypes represent number of variants.

Science Press
ZR Zoological Research

中文摘要

地方鸡品种环境适应性的遗传基础

袁经纬[1]，贾晓旭[2]，黄晓隆[1]，李云雷[1]，倪爱心[1]，陈国宏[3]，郑嫩珠[4]，和胜[5]，孙研研[1]，易国强[6]，韩威[2]，陈继兰[1]*
*通讯作者：陈继兰(chen.jilan@163.com)
[1]动物生物育种国家重点实验室，农业和农村部动物（家禽）遗传育种与繁殖重点实验室，中国农业科学院北京畜牧兽医研究所，北京 100193
[2]江苏省家禽科学研究所，扬州 225125，中国
[3]教育部农业和农产品安全国际联合研究实验室，扬州大学动物科学与技术学院，扬州 225009
[4]福建省农业科学院畜牧兽医研究所，福州 350013
[5]怒江畜牧科技推广站，怒江 673199，中国
[6]中国农业科学院深圳农业基因组研究所，深圳 518124，中国

地方鸡遗传资源丰富，是生物多样性的重要组成部分，也是人类重要的优质蛋白来源。然而，不断变化的饲养环境正对它们的生存和生产性能构成严重影响。因此，评估不同环境下的种群适应特定环境的能力，对于培育环境友好型畜禽品种和保护珍稀濒危遗传资源至关重要。本研究中，我们整合了来自 44 个不同地理位置的中国地方鸡品种群体，共计 1022 只鸡的生态学数据和全基因组重测序数据，来揭示局部适应的基因组特征。应生态位建模方法，我们首先从 87 个农业气候变量中，筛选出 8 个重要的环境因子影响地方鸡的生态位，这其中包括太阳辐射、降水、温差，以及 5 个土地覆盖指标（农田面积、水域面积、森林覆盖率、裸地和灌木覆盖率）。通过群体景观基因组和比较基因组学分析，我们挖掘到 *UNC80*、*PTPRO*、*NCOR2*、*CSF2RB*、*NXT2* 和 *PALLD* 等候选基因，分别与太阳辐射、降水、昼间温度范围、农田面积、树木覆盖和裸地环境适应性关联。进一步对候选 SNP 的调控功能注释，发现这些基因中与适应性选择关联的非编码 SNP 的基因频率在群体中表现出地理空间的规律变化。他们通过染色质可及性（ATAC）和 DNA 甲基化来充当调节元件，以组织特异性的方式影响地方鸡对特定环境的适应性。研究结果揭示中国地方鸡品种的丰富遗传多样性，阐明了地方鸡适应特定环境的基因组学分子机制，为畜禽遗传资源保护和育种提供了理论参考和研究范式。

关键词：环境适应性；生态位建模；景观基因组学；群体遗传学；地方鸡