ZR | Zoological Research

Article   **Open Access**

Genome-wide scan for selective footprints and genes related to cold tolerance in Chantecler chickens

Nai-Yi Xu[1], Wei Si[2], Ming Li[1], Mian Gong[1], Jean-Marc Larivière[3], Hojjat Asadollahpour Nanaei[1], Pei-Pei Bian[1], Yu Jiang[1,*], Xin Zhao[1,2,*]

[1] *Key Laboratory of Animal Genetics*, *Breeding and Reproduction of Shaanxi Province*, *College of Animal Science and Technology*, *Northwest A & F University*, *Yangling*, *Shaanxi* 712100, *China*

[2] *Department of Animal Science*, *McGill University*, *Montreal*, *Québec H9X3V9*, *Canada*

[3] *Institut de technologie agroalimentaire du Québec*, *La Pocatière*, *Québec G0R1Z0*, *Canada*

ABSTRACT

The Chantecler chicken, a unique Canadian indigenous breed, is well adapted to extremely cold environments. However, its genetic characteristics have not been well studied. Here, we analyzed the whole genomes of 10 Chantecler chickens and 121 worldwide chickens, which indicated that Chantecler chickens were derived from commercial chickens and exhibit a high level of inbreeding. Based on a genome-wide scan, we identified two vital candidate regions containing *ME3* and *ZNF536*, which are related to fat metabolism and nervous system in cold adaptation, respectively. We also found that the genetic mechanism of cold adaptation in Chantecler chickens differed from that of chickens from other cold regions, such as northern China. Our study indicated that specialized commercial chickens in the early 20th century contained sufficient genetic diversity to adapt to extreme cold environments over a very short time. These findings enrich our understanding of the adaptive potential of commercial species.

Keywords: Chantecler chickens; Whole-genome

resequencing; Cold adaptation; *ME3*; *ZNF536*

INTRODUCTION

Elucidating the genetic mechanisms underlying organismal adaptation to extreme environments, such as cold climate, is of great interest. Livestock and poultry, such as cattle, goats, sheep, chickens, and ducks, which are distributed worldwide, have adapted to various environmental conditions and production systems through both natural and artificial selection. Climate-mediated selection pressure has shaped phenotypic diversity and left a genetic "imprint" on the genomes of animals raised in different geographical regions. Understanding the genetic basis of adaptation has become a central focus of evolutionary biology. Rapid progress in high-throughput technologies, computational methods, and analytical techniques has made it possible to study candidate genes and causal variations that contribute to the genetic architecture of complex phenotypes in different organisms (Ai et al., 2015; Rubin et al., 2010; Yang et al., 2016).

To understand the genetic basis of animal adaptations to the natural environment, studies have focused on various species, including humans, living in different climates, e.g., chickens (Wang et al., 2015), yaks (Qiu et al., 2012), and canids (Li et al., 2014; Miao et al., 2017; Signore et al., 2019) at high elevation (Tibetan Plateau of China), camels (Wu et

al., 2014) in desert regions, and Inuits (Fumagalli et al., 2015), Siberians (Hallmark et al., 2019), and Yakutian horses (Librado et al., 2015) in cold environments. Low temperature is a major environmental factor that limits animal growth and can threaten animal survival. Animals have evolved various physiological and biochemical mechanisms to adapt to cold environments. To date, studies on humans (Fumagalli et al., 2015) and different animal species, such as sled dogs (Sinding et al., 2020), sheep (Lv et al., 2014), and goats (Cai et al., 2020), have identified several candidate genes with major effects on cold adaptation. For example, mitochondrial uncoupling protein genes *UCP1* and *UCP3* are related to non-shivering thermogenesis (NST) in brown adipose tissue (BAT) and skeletal muscle (Cannon & Nedergaard, 2004). Several potential candidate genes in the human genome, i.e., *FADS1*, *FADS2*, *FADS3*, *TBX15*, and *WARS2*, are associated with adaptation to a polyunsaturated fatty acid (PUFA)-rich diet and cold climates (Fumagalli et al., 2015). Some specific haplotype genes, including *TRPC4*, *TRPV2*, *CACNA1A*, *SLC25A40*, and *APOO*, are potentially related to arctic adaptation in sled dogs (Sinding et al., 2020). Comprehensive analysis of sheep suggests that strong selection of the *TBC1D12* gene could act as a key point for studying the biological processes and potential mechanisms of genetic adaptation to the environment (Lv et al., 2014). Modern goat populations from southern and northern China show two major genes, *FGF5* and *EDA2R*, associated with hair follicle development, which may be related to local adaptive evolution (Cai et al., 2020). Although chickens are widely distributed and adapted to the harsh environmental stresses prevalent in many rural areas, there are scant published studies on their adaptation to cold climates.

The Chantecler (CA) chicken is a unique dual-purpose breed from Canada and is well adapted to the cold climate of Quebec, where the average temperature in winter ranges from about −10 to −25 ℃. They exhibit several unique cold-adapted characteristics such as dense feathers and a thick down layer against the body. Furthermore, they have a distinctive small cushion comb and wattles, which may help them withstand harsh climate conditions (Ekarius, 2007). However, the excellent adaptability of CA chickens to extreme cold environments remains poorly explored. Thus, research on the mechanisms underlying their adaptation to such environments will provide new insight into cold adaptation in poultry.

In the present study, we resequenced the whole genome of 10 CA chickens and downloaded the resequencing data of 121 different chickens distributed worldwide (Li et al., 2017; Wang et al., 2015, 2020; Yi et al., 2014) to analyze the mechanisms underlying cold adaptation in CA chickens and to explore their genetic characteristics. These results should provide a theoretical basis for improvement in chicken breeds.

MATERIALS AND METHODS

Sample collection and genome sequencing
We sampled a total of 10 CA chickens from a local breeder in Quebec, Canada. DNA was extracted from the blood of each individual using the standard phenol-chloroform protocol. DNA

samples were sequenced with Illumina HiSeq technology at the McGill University and Genome Quebec Innovation Centre. Moreover, whole-genome resequencing data of 121 chickens from all over the world were collected, including Europe and America (Commercial, C (*n*=10)), Middle East (Iran, ME (*n*=30)), East Asia (China, EA (*n*=25)), Tibet, China (TB (*n*=10)), Yunnan, China, and Southeast Asia (Yunnan and Indonesia, YN.SEA (*n*=35)), South Asia (India, SA (*n*=10)), and Indonesia (green jungle fowl, GJF (*n*=1)) (Supplementary Table S1).

Sequence alignment and variant calling
Raw sequencing reads were trimmed using Trimmomatic v0.38 (Bolger et al., 2014) to filter low-quality bases and sequences using default parameters. All clean reads were then mapped to the chicken reference genome (GRCg6a) using the Burrows-Wheeler Aligner "BWA-MEM" algorithm with default parameters (Li & Durbin, 2009). On average, 99.17% of the reads sequenced in this study were mapped and the final average sequencing coverage was 20.37× (ranging from 15.36× to 25.26×) per individual. Using Picard Tools, reads were sorted and merged, and all duplicate reads were removed successively. Variant calling of sequence data was performed using the Genome Analysis Toolkit (GATK, v3.6-0-g89b7209) (McKenna et al., 2010). The applied criteria to all single nucleotide polymorphisms (SNPs) were: (1) <1/3× mean sequencing depth (over all included individuals)<3×; (2) mapping quality (MQ)>40.0; (3) Quality by Depth (QD)>2.0; (4) Fisher Strand (FS)>60.0; (5) MQRankSum>−12.5; (6) ReadPosRankSum>−8. Haplotype phasing was implemented using BEAGLE v4.1 (Browning & Browning, 2007). Identified SNPs were further classified by ANNOVAR (Wang et al., 2010) based on the gene annotation of GRCg6a.

Population genetics analysis
An individual-based neighbor-joining (NJ) tree was constructed from the matrix of pairwise genetic distances of autosomal SNP data from 81 chickens (10 CA chickens, 10 samples from each group of C, ME, TB, and SA, 15 samples from each group of EA and YN.SEA, and one sample of GJF) using PLINK software (Purcell et al., 2007). The NJ tree was visualized by iTOL (Letunic & Bork, 2019). Using the smartpca program in EIGENSOFT v5.052 (Patterson et al., 2006), principal component analysis (PCA) was carried out based on pruned SNP data and eigenvector significance was detected by the Tracy-Widom test. Genetic structure clustering was deduced with the model-based assignment program ADMIXTURE v1.3.0 (Alexander et al., 2009) by estimating the ancestry of each individual using the genome-wide unlinked SNP dataset. The number of assumed genetic clusters (K) ranged from 2 to 7, with 200 bootstrap replicates. Using the qp3Pop program in the AdmixTools package (Patterson et al., 2012), we calculated outgroup *f3* statistics, using GJF as the outgroup.

Run of homozygosity (ROH), genetic diversity, and linkage disequilibrium (LD) detection
The ROH for each group was identified using the --homozyg option implemented in PLINK. The following criteria were applied for ROH identification: (1) required minimum density

(--homozyg-density) 50; (2) number of missing calls of a window (--homozyg-window-missing) 5; (3) number of heterozygotes of a window (--homozyg-window-het) 3. Length and average number of ROHs of each group were estimated, with length divided into four categories, i.e., 0.5–1 Mb, 1–2Mb, 2–4 Mb, and >4 Mb. Nucleotide diversity of each group was also calculated with a window size of 40 kb and step size of 20 kb using VCFtools (Danecek et al., 2011). The decay of LD was calculated via the squared correlation coefficient (r^2) using PopLDdecay (Zhang et al., 2019). Genomic inbreeding coefficients based on ROH (F_{ROH}) were estimated for each bird, according to: $F_{ROH}=L_{ROH}/L_{total}$ (McQuillan et al., 2008).

Whole-mitochondrial genome phylogeny

The BAM files of the mitochondrial genomes of 10 CA chickens were transformed to FASTQ files and then assembled using Mapping Iterative Assembler v1.0 (Briggs et al., 2009). Using the maximum-likelihood (ML) model within IQ-TREE (Nguyen et al., 2015), we analyzed the mitochondrial phylogenetic relationships between CA chickens and 58 previously published wild fowls and domestic chickens (Miao et al., 2013) (Supplementary Table S2). TIM2+F+R2 was chosen as the best-fit model for the complete mitochondrial genome sequences using ModelFinder (Kalyaanamoorthy et al., 2017). Bootstrap support values of ML analysis were generated with 1 000 replicates (-bb 1 000). The final tree topology was visualized using iTOL (Letunic & Bork, 2019).

Genome-wide selective sweep analysis

To identify genomic regions of selective sweeps associated with cold adaptation, we used several methods to investigate selection signatures, including population differentiation index (*F*st) (Weir & Cockerham, 1984), nucleotide diversity (π), and cross-population extended haplotype homozygosity (XPEHH) (Szpiech & Hernandez, 2014). These three methods are often used in positive natural selection analyses. Statistics were calculated using a sliding window approach with a window size of 40 kb and step size of 20 kb. We calculated the average *F*st, π-Ratio, and XPEHH values of SNPs in each window and used the outlier method to obtain windows with the top 1% of *F*st, π-Ratio, and XPEHH values. Protein-coding genes in these outlier windows were annotated using ANNOVAR software (Wang et al., 2010) and were treated as candidate positively selected genes (PSGs). Using the KOBAS v3.0 tool (Xie et al., 2011), we performed Gene Ontology (GO) pathway analysis of the potential candidate genes to better understand their biological functions. Pathways with a false discovery rate (FDR)-corrected *P*-value<0.05 were considered significantly enriched.

RESULTS

Genome resequencing and variation

A total of 10 CA chickens were collected for genome resequencing (average depth of 20.37×). To facilitate comparisons with other chicken breeds, we combined our data with available whole-genome resequencing data of 121 individuals from seven geographical regions according to their locations, including C, ME, EA, TB, YN.SEA, SA, and GJF

(phylogenetic outgroups) (Figure 1A; Supplementary Table S1). Functional annotation of the polymorphic sites identified a total of 4 681 038 SNPs in the 10 CA genomes, which included most intergenic regions (50.3%) and intronic regions (33.8%). Exons contained 2.0% of the total SNPs, with 23 911 nonsynonymous SNPs and 67 932 synonymous SNPs.
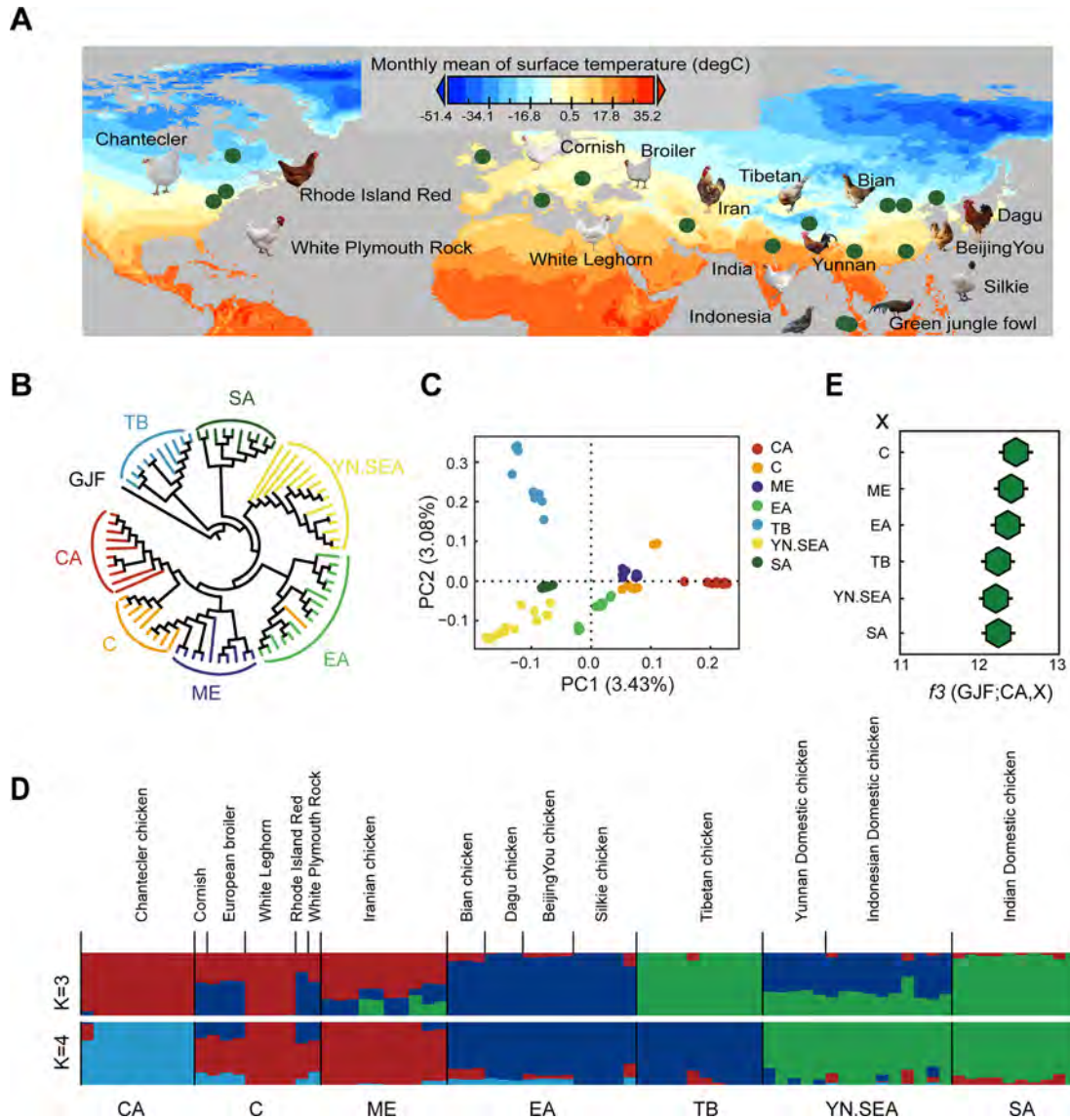
Population genetic structure of CA chickens

To determine the population structure and genetic relationships of the different chicken groups, we conducted a series of analyses, including phylogenetic reconstructions, PCA, and Bayesian clustering, using whole-genome resequencing data. In total, 81 chickens were used to perform genetic relationships (see Supplementary Table S1 for detailed sample information). The NJ tree, rooted with GJF, showed that the CA group was closest to the C group, followed by the ME group (Figure 1B). Similarly, the genetic relationships between CA and the other chickens were also confirmed by PCA (Figure 1C). The PCA results showed that the CA group deviated from the other groups, suggesting that the CA group has unique characteristics. The admixture results provided additional corroborating evidence for the genetic relationships of CA and other chickens. In clustering analysis, when K=3, the CA group shared similar ancestries as chickens from the C and ME groups; when K=4, the CA group was separated independently (Figure 1D). We also used outgroup *f3* (Figure 1E) to investigate the genetic relationships between CA and other chicken groups. Results from the outgroup *f3* statistics unequivocally indicated that CA chickens have a higher genetic affinity to chickens from the C group, followed by the ME chickens. These results indicate that CA chickens are genetically closest to the C group, followed by the ME group.

To examine variations in mitochondrial DNA (mtDNA), we constructed a ML phylogenetic tree based on the complete mitochondrial genome sequences of 10 CA chickens and 58 downloaded chickens (Supplementary Table S2) that cover almost all known mitochondrial haplotypes. Results showed that CA chickens belonged to the mtDNA haplotype E1 (Supplementary Figure S1), which is the most widely distributed matrilineal lineage (Miao et al., 2013).

Patterns of genomic variations and LD

To explore the patterns of genomic variation, LD, and F_{ROH} in CA and other chickens, we analyzed ROH size, average number of ROHs, π, LD, and F_{ROH}. The SNPs (6 Mb) of each population were used for ROH analysis. The genetic variation parameters showed a basically consistent pattern in the different populations. Specifically, CA, C, and TB chickens exhibited larger ROHs and higher ROH numbers, whereas the opposite genomic variations were observed in the YN.SEA and SA chickens (Figure 2A, B). In addition, CA chickens not only exhibited lower π values but also a slow decay rate and visibly high level of LD (Figure 2C, D). The genomic diversity and LD patterns strongly suggest that CA chickens are likely affected by inbreeding or population founder effects. To further explore the reasons for the above observations, we calculated F_{ROH}. Of note, the higher F_{ROH} found for the CA chickens indicated a higher level of inbreeding in this group (Figure 2E).

Figure 1 Population genetic analyses of CA chickens

A: Geographic variation in monthly mean surface temperature for location of all chickens used in the study. B: NJ tree of CA chickens sequenced in this study and publicly available whole-genome sequences of representative chickens worldwide. C: PCA with first (PC1) and second (PC2) principal components. D: Population structure analysis. E: Outgroup *f3* statistics (GJF; CA, X), with higher *f3* values suggesting more ancestral alleles shared by CA and X, and thus their closer relationship. CA, Chantecler chickens; C, commercial chickens; ME, domestic chickens from Middle East; EA, domestic chickens from East Asia; TB, Tibetan chickens; YN.SEA, domestic chickens from Yunnan, China, and Southeast Asia; SA, domestic chickens from South Asia; GJF, green jungle fowl from Indonesia.

Genome-wide selective sweep analyses of CA chickens

To identify candidate population stratification genes potentially related to cold adaptation in the CA group, we performed genome-wide selective sweep analyses based on the genomic variation data in this study. Typically, those genomic regions related to traits of interest that have undergone selection pressure in one group should show strong differences in comparison to other groups, including high *F*st, significantly reduced π, and long-range extended haplotype homozygosity (EHH) (Sabeti et al., 2006). Therefore, we used the abovementioned statistical methods to identify potential candidate genes that may play a significant role in the regulation of stress response to cold adaptation in CA chickens. We used 10 and 30 chickens from the CA and ME groups, respectively, for genome-wide selective sweep analyses. In the analyses, the top 1% was used as a threshold to identify potential candidate regions, and genes annotated in these regions were considered as potential candidate genes. In total, 310 genes with significantly higher $F\mathrm{st}_{\text{CA-to-ME}}$ values (top 1%) were identified as candidate PSGs in the CA chickens (Figure 3A; Supplementary Table S3). As high population differentiation is affected by both positive selection and demographic history (Sabeti et al., 2006), the π-$\text{Ratio}_{\text{ME/CA}}$ and $\text{XPEHH}_{\text{CA-to-ME}}$ statistical methods were also
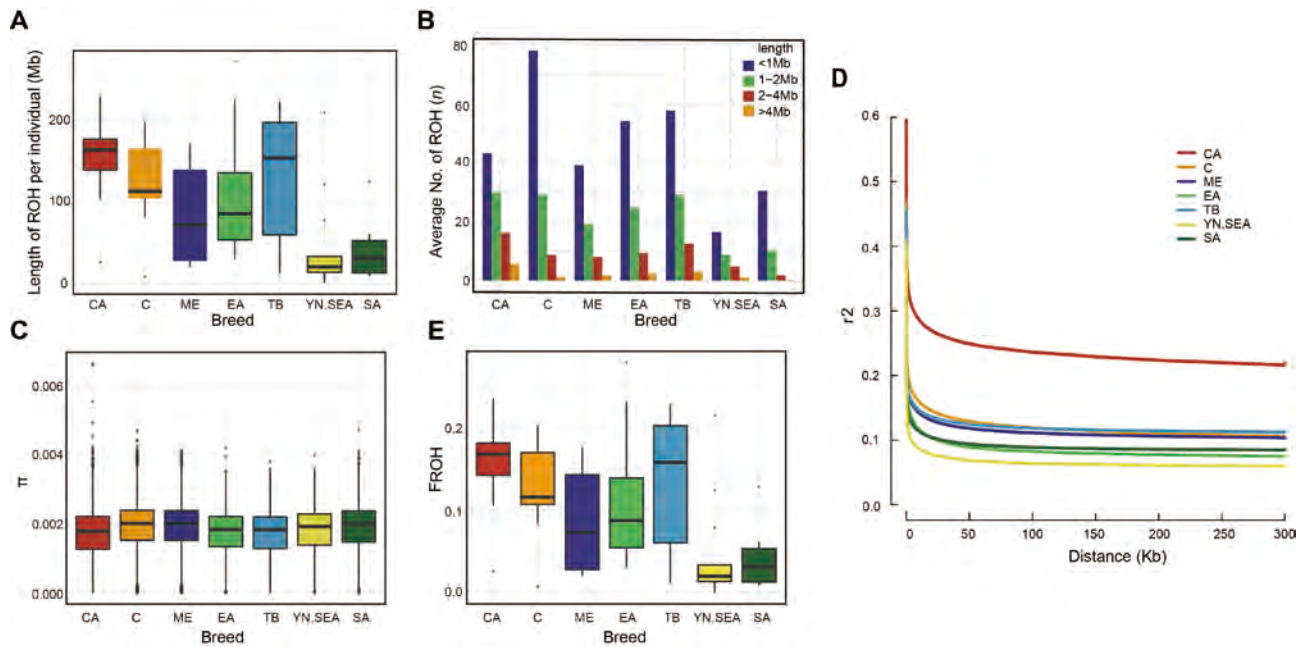
Figure 2 Analysis of genomic variation, LD, and F$_{ROH}$ for CA chickens

Length (A) and average number (B) of ROHs in CA chickens and other representative chickens of worldwide geographical distribution. C: Distributions of π in different groups. D: Decay of LD in different groups, with one line per breed. E: F$_{ROH}$ in different chicken groups.

adopted in this study. These two methods are less sensitive to population demographic events than other approaches (Innan & Kim, 2008). In total, 247 and 212 PSGs in the top 1% were identified using the π-Ratio and XPEHH methods, respectively (Figure 3B, C; Supplementary Tables S4, S5). Overall, 622 unique PSGs were found using the three methods, with only 36 shared by all three (Figure 3D; Supplementary Table S6). Unsurprisingly, the overlap among detected PSGs using different statistical methods was underwhelming. Different statistical methods based on different principles can yield contradictory results from the same dataset, which may lead to this rare overlap phenomenon (Sabeti et al., 2006). Importantly, all 36 PSGs were found by at least one of the above methods when CA chickens were regarded as the target population compared to domestic chickens from Indonesia (ID, Supplementary Table S6). These results indicate that all 36 PSGs are reliable PSGs in the CA chickens.

To further elucidate the genetic mechanisms related to several candidate genes, functional enrichment analyses were applied to the three different PSG lists. GO analysis of the PSGs showed several significantly enriched categories related to various metabolic processes (corrected P<0.041), biological processes (corrected P<0.022), cellular processes (corrected P<0.0009), and biosynthetic processes (corrected P<0.047) (Supplementary Figures S2, S3 and Tables S7, S8), which may contribute to thermostatic maintenance in cold environments. Based on gene annotation, the identified PSGs were also related to different physiological functions and metabolic processes, such as neuronal development or migration, lipid metabolism, hair bundle morphogenesis, vasoconstriction and vasodilation, developmental processes,

kidney development, skin pigmentation, and circadian rhythm. These processes are reported to play important roles in cold-induced thermoregulation (Adolph & Molnar, 1946; Barnett, 1959; Boulant & Dean, 1986; Lynch et al., 2015; Masoro, 1966; Smith, 1962; Yang et al., 2018).

ME3 and *ZNF536* genes in CA chickens

To determine which of the 36 PSGs shared by the three methods in the CA group were most likely to be related to cold adaptation, we conducted haplotype and nonsynonymous mutation analyses and carefully searched the literature. Interestingly, we found a strong selective sweep spanning a 250 kb region (190 090 000–190 340 000 bp) on chromosome 1, which exhibited high Fst$_{CA-to-ME}$, π-Ratio$_{ME/CA}$, and XPEHH$_{CA-to-ME}$ values (Figure 4A). This region contained five PSGs, including *PRSS23*, *ME3*, *FAM181B*, *PRCP*, and *DDIAS*, which may be related to cold adaptability in CA chickens. Further analyses revealed that the CA chickens had a core haplotype in this region, which was distinct from that of other chickens (Figure 4B). In particular, the genotype frequency of the nonsynonymous mutation in the *ME3* coding region (c. A287G), which results in an amino acid substitution (p.H96R), was 90% in the CA chickens but was less than 36% in the other chickens (Figure 4C). In addition, the genotype frequency of a nonsynonymous mutation (p.A9S) in the *DDIAS* gene was 90% in the CA chickens, but less than 32% in the other chickens (Figure 4D). *ME3* is an isoform of the malic enzyme, which is an oxidative decarboxylase that catalyzes the conversion of malate to pyruvate, simultaneously generating NADPH (Jiang et al., 2013; Pongratz et al., 2007). Malic enzymes participate in pyruvate cycling and insulin secretion (Pongratz et al., 2007). Compared to other tissues, malic enzymes are more abundant
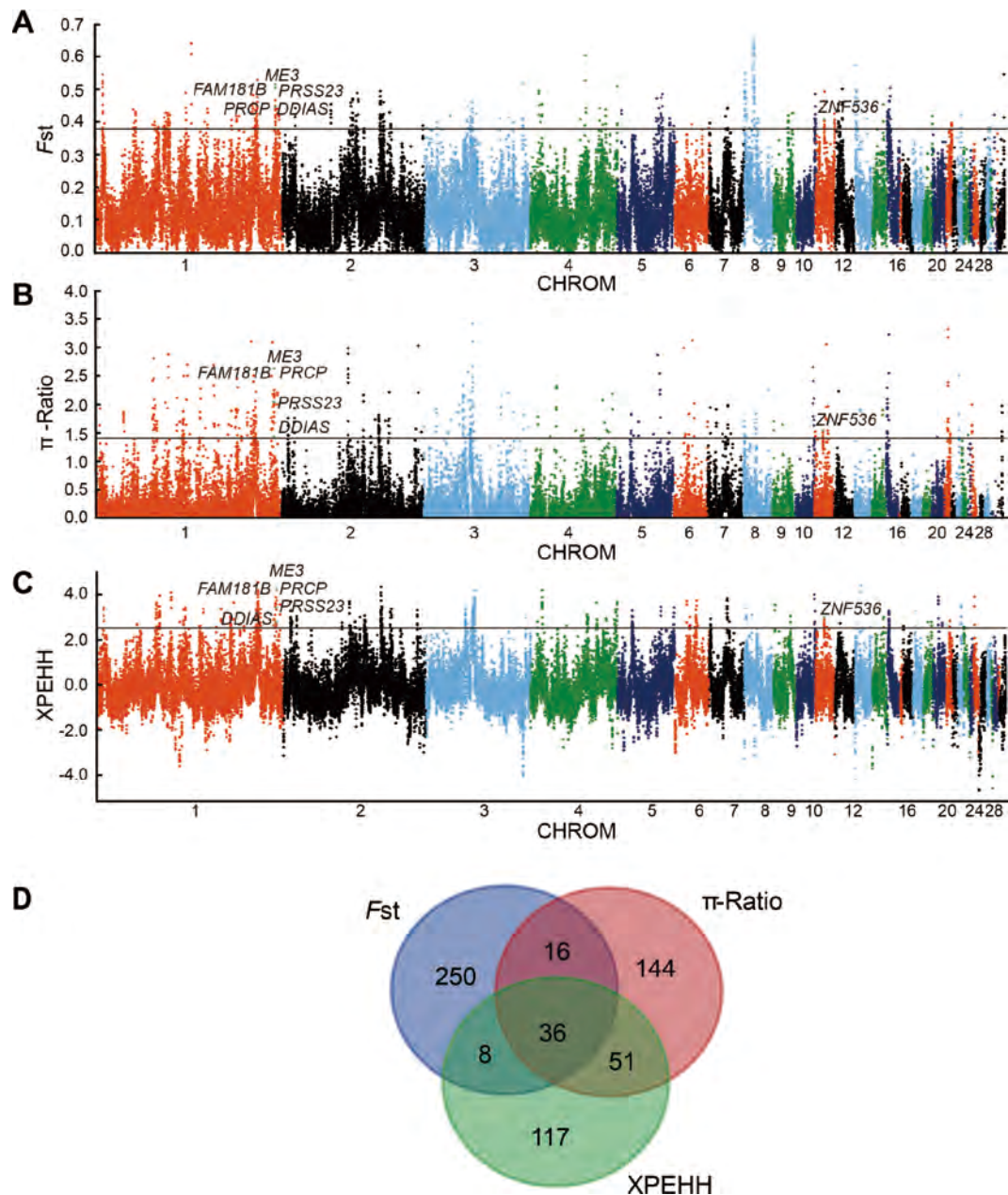
Figure 3 Genome-wide selection scan for PSGs in CA chickens using sliding window analysis (40 kb window size, 20 kb step size, kb increment, 99th percentile cutoff)

A: Selection signatures in CA chickens for $Fst_{CA-to-ME}$. B: Selection signatures in CA chickens for $\pi\text{-Ratio}_{ME/CA}$. C: Selection signatures in CA chickens for $XPEHH_{CA-to-ME}$. Threshold (top 1%) of $Fst_{CA-to-ME}$, $\pi\text{-Ratio}_{ME/CA}$, and $XPEHH_{CA-to-ME}$ is marked with a horizontal black line. D: Number of PSGs identified in CA chickens by the three methods listed in each Venn diagram component.

in lipogenic, white and brown adipose, and liver tissues, and may play a role in lipogenesis by providing NADPH (Chung et al., 1999; Wise & Ball, 1964). Our results indicated that the *ME3* gene, which is related to fat production, likely plays an important role in cold adaptation in CA chickens. Although *DDIAS* functions as an anti-apoptotic protein in response to DNA repair and *PRSS23* expression is regulated by estrogens, the other functions of these two genes remain to be elucidated (Huan et al., 2014; Im et al., 2016). *PRCP* and *FAM181B* may participate in cold adaptation by regulating

angiogenesis and nervous system development (Hagedorn, 2011; Marks et al., 2016). In addition, *ZNF536*, a gene on chromosome 11, showed strong positive selection in the CA chickens, as exhibited by the high $Fst_{CA-to-ME}$, $\pi\text{-Ratio}_{ME/CA}$, and $XPEHH_{CA-to-ME}$ values (Figure 5A), while shared haplotypes were rarely observed in the CA and other chickens (Figure 5B). A nonsynonymous mutation (c.G3685A) was also found in the *ZNF536* coding region, generating an amino acid substitution (p.A1229T) in a conserved domain (Supplementary Figure S4). The derived genotype frequency
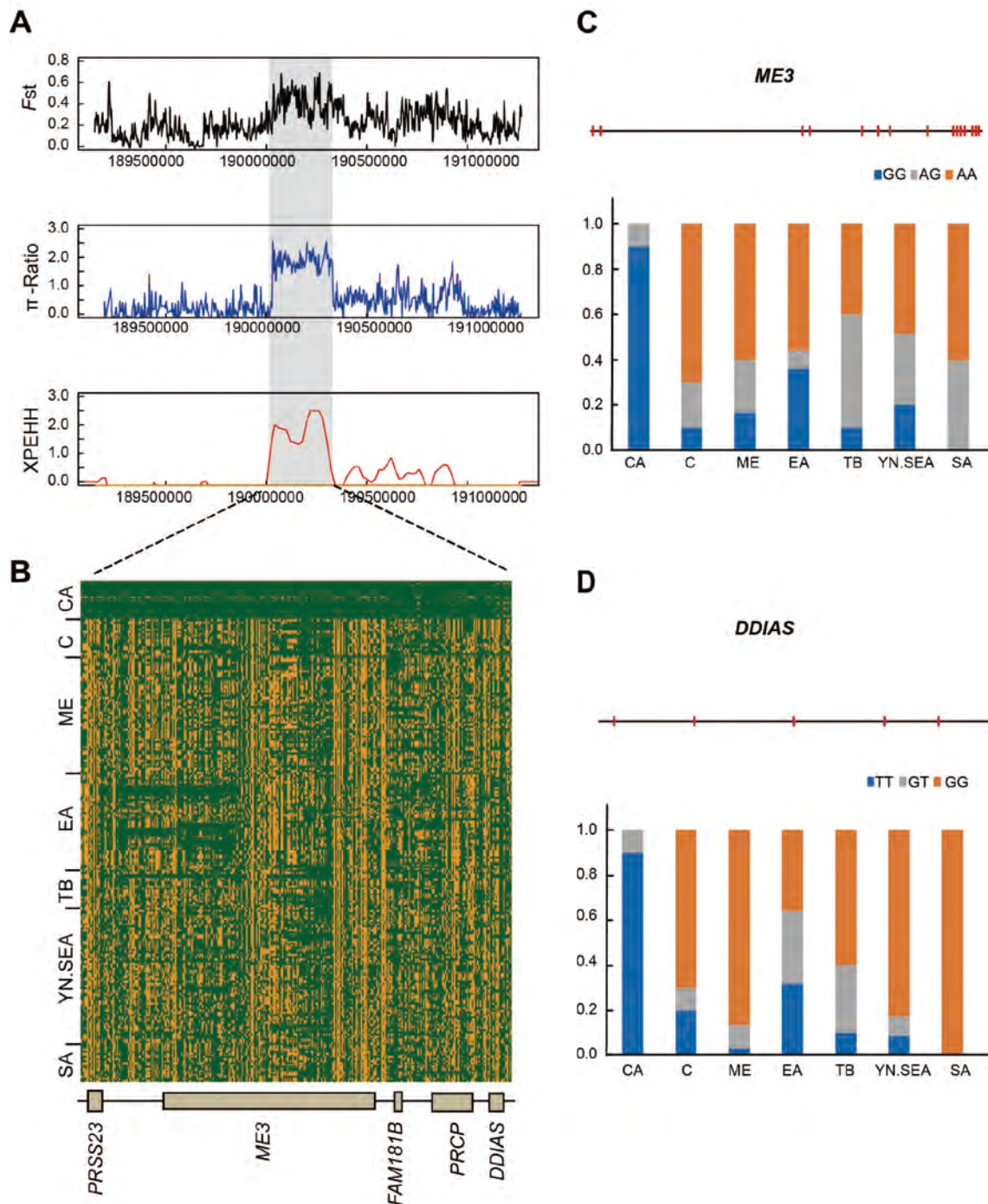
**Figure 4 Selection signals in 250 kb region of chromosome 1 in CA chickens**

A: Zoom of selective analyses in 250 kb region of chromosome 1. $Fst_{CA\text{-}to\text{-}ME}$, π-Ratio$_{ME/CA}$, and XPEHH$_{CA\text{-}to\text{-}ME}$ were plotted as lines based on sliding window analysis (4 kb window and 2 kb increment). B: Degree of haplotype sharing between CA and other chickens in 250 kb region of chromosome 1, which contained five genes. Structure diagram of *ME3* (C) and *DDIAS* (D) genes, with red bars representing exons, and allele frequency of mutant *ME*3 and *DDIAS* loci. Blue, gray, and orange represent homozygous mutant, heterozygous mutant, and homozygous wild-type, respectively.

in the CA chickens was 80% but was less than 8% in the other chickens (Figure 5C). *ZNF536* is a highly conserved zinc finger protein that negatively regulates neuronal differentiation (Qin et al., 2009). It plays an important role in the development of forebrain neurons involved in social behavior and stress

(Thyme et al., 2019). Thus, *ZNF536* likely participates in cold adaptation in CA chickens.

Parallelism in cold-tolerant chickens

To explore whether the genetic mechanisms underlying cold
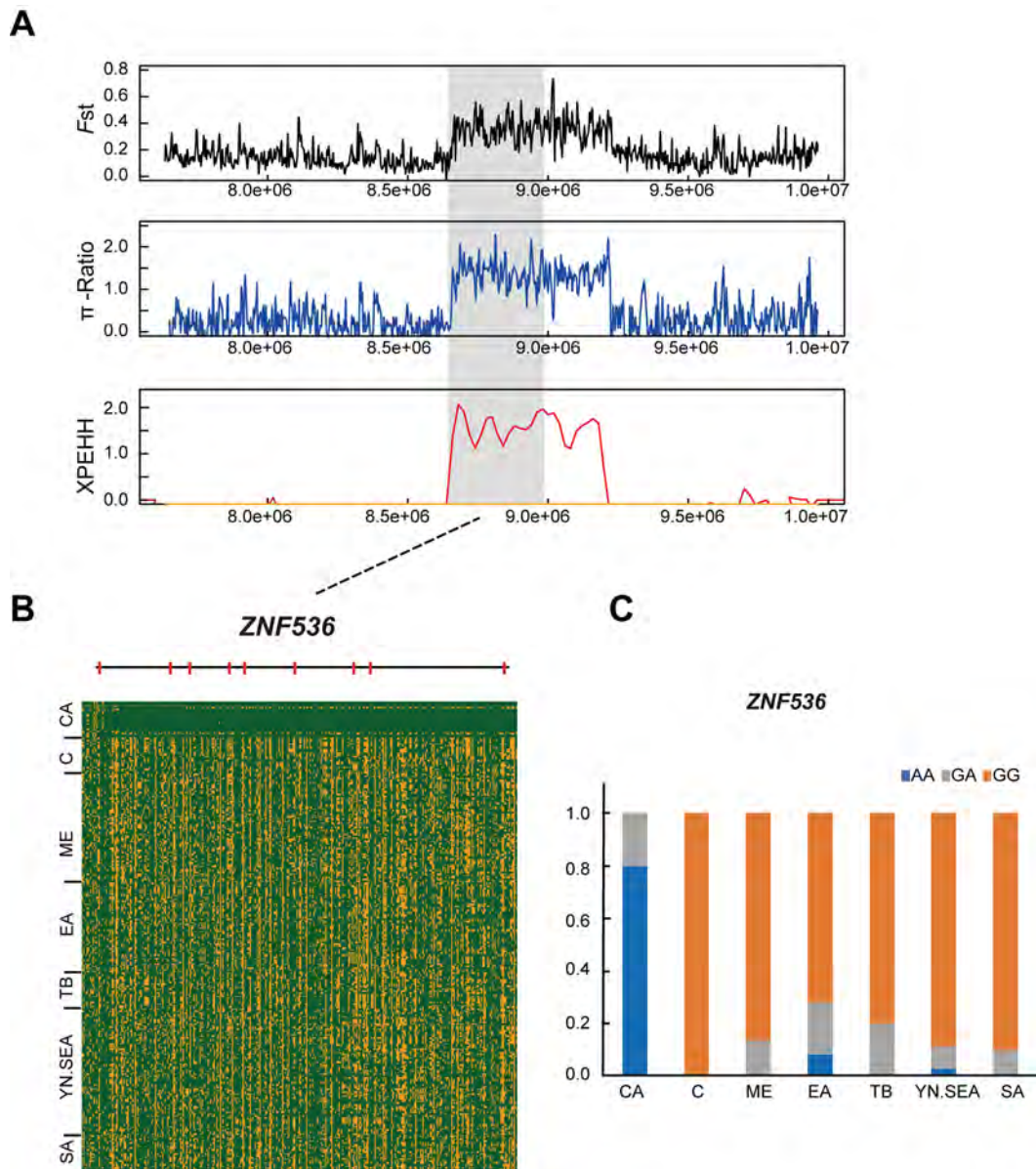
Figure 5 Selective signature on *ZNF536* in CA chickens

A: Zoom of selective analyses in *ZNF536* region. $F_{st_{CA-to-ME}}$, π-Ratio$_{ME/CA}$, and XPEHH$_{CA-to-ME}$ were plotted as lines based on sliding window analysis (4 kb window and 2 kb increment). B: Structure diagram of *ZNF536* gene, with red bars representing exons, and degree of haplotype sharing between CA and other chickens in *ZNF536* gene. C: Allele frequency of mutant *ZNF536* loci. Blue, gray, and orange represent homozygous mutant, heterozygous mutant, and homozygous wild-type, respectively.

tolerance in chickens in different regions are the same, we conducted genome-wide selective sweep analyses on representative cold-resistant chickens (i.e., Bian and Dagu chickens) in northern China (NCN). We used eight Bian chickens, eight Dagu chickens, and 30 ME chickens for analysis of selection signals. Based on selection signals analysis of the chickens, we detected 208, 1 023, and 228 PSGs in the top 1% using $F_{st_{NCN-to-ME}}$, π-Ratio$_{ME/NCN}$, and XPEHH$_{NCN-to-ME}$, respectively (Supplementary Figure S5 and Tables S9–S11). GO analyses of the three PSG lists showed that the significantly enriched categories not only included cell-related processes (corrected *P*<0.043), biosynthetic

processes (corrected *P*<0.049), and metabolism-related processes (corrected *P*<0.049), but also regulation of neuronal differentiation, genesis, and development (corrected *P*<0.036) (Supplementary Figures S6, S7 and Tables S12, S13). Although the cold-tolerant NCN and CA chickens shared 68 PSGs, none of the 36 PSGs detected using the three methods in the CA chickens were found in the PSG list of NCN chickens (Supplementary Figure S8 and Table S6). These results suggest that parallel genetic mechanisms are likely responsible for cold adaptation in these two chicken breeds.

DISCUSSION

The primary breeding objective of animal husbandry has been to increase growth and production traits (e.g., meat, eggs, and milk) to meet the rapidly increasing demands of human consumption. Although climate has a profound impact on the survival of organisms, limited previous research has focused on animal adaptations to extreme climate conditions, such as cold adaptation. Recently, with increasing societal needs and technology development, as well as changes in the environment, animal adaptability has attracted increasing attention from both a breeding and biological perspective. Several studies have been conducted on cold adaptability in humans and animals (Fumagalli et al., 2015; Hallmark et al., 2019; Librado et al., 2015; Liu et al., 2014; Lynch et al., 2015). However, very few studies have focused on cold adaptability in poultry. In this study, 10 cold-tolerant CA chickens were used for whole-genome resequencing to understand the genetic mechanisms underlying cold adaptability in chickens. Our results represent a valuable resource and provide an excellent opportunity for exploring cold adaptability in chickens.

Although a pre-Columbian Polynesian introduction of chickens to the Americas has been suggested (Storey et al., 2007), the viewpoint that chickens were introduced to the Americas by European (Portuguese or Spanish) colonizers after 1500 AD has also been recognized (Herrera et al., 2020). The CA chicken was developed in the early 20th century for cold weather with high egg and meat production. Here, population genetic analysis indicated that CA chickens shared more ancestral components with C chickens (Figure 1) than the other chickens studied. Overall, 41 of the 622 PSGs identified in the CA chickens appear to be related to cold adaptation in humans and other animals. For example, *TBX15*, which plays a role in adipocytes (Gburcik et al., 2012), is reportedly related to cold adaptation in Inuits (Fumagalli et al., 2015), and was identified as a PSG in the CA chickens. *TRPC4* is involved in the temperature sensitivity pathway (Hofmann et al., 2002) and has been identified in sled dogs (Sinding et al., 2020) and CA chickens. *OCA2* and *TYR*, two key genes for skin lightening in humans (Shriver et al., 2003), were identified in CA chickens. We found limited PSG overlap between our study and previous reports, which may be due to parallel genetic mechanisms underlying cold adaptation in CA chickens and other animals. Our study and other results (Ai et al., 2015; Librado et al., 2015; Lv et al., 2014) show that cold adaptation is a complex process involving the nervous system, fat metabolism, hair formation, kidney development, vasoconstriction and vasodilation, pigmentation, and circadian rhythm. In selective sweep analysis, we identified 24 PSGs in the CA chickens involved in fat metabolism, including *ME3*, *TBX15*, *LRP2*, *PPARG*, *PNPLA2*, and *CERS6*, consistent with previous studies stating that fat metabolism plays an important role in cold adaptation in Greenlandic Inuits, Siberians, Yakutian horses, polar bears, and woolly mammoths (Fumagalli et al., 2015; Hallmark et al., 2019; Librado et al., 2015; Liu et al., 2014; Lynch et al., 2015). Of note, *ME3*, which was identified by several methods in our study (Figure 4), likely plays an important role in the cold adaptation of CA

chickens. In addition, 58 of the identified PSGs in CA chickens, including *ZNF536*, *CTNNA2*, *TIAM2*, *DISP3*, *KCNH7*, and *FBXO2*, participate in neural processes, such as regulation of neuronal development, migration, differentiation, and neurogenesis, consistent with previous findings that the neural system plays an important role in environmental adaptation in Chinese domestic pigs (Ai et al., 2015). In the current study, *ZNF536* was preliminarily shown to be related to cold adaptation in CA chickens (Figure 5). *DDIAS*, an anti-apoptotic protein involved in DNA repair (Im et al., 2016), may participate in cold adaptation in CA chickens, although other functions related to this gene need to be explored. The distinctive physical attributes of CA chickens, e.g., thicker and fluffier plumage, make them perfectly adapted to cold climates (Ekarius, 2007). Here, *FZD3*, which is involved in hair follicle development (Hung et al., 2001), was identified by selective sweep analysis in CA chickens. However, further research on this vital PSG is needed. We believe that *ME3* and *ZNF536* are more reliable PSGs and are more likely to participate in cold adaptation, although further functional experiments are needed to understand their role in CA chicken cold adaptation.

In this study, we performed selective sweep analysis between two cold-tolerant chicken breeds from Canada (CA chickens) and China (Bian and Dagu chickens) with chickens from other parts of the world. Several loci on different chromosomes were identified as potentially involved in cold adaptation traits. Our results suggest that parallel genetic mechanisms may underlie cold adaptation in these chicken breeds, although additional physiological and functional experiments are needed for verification. Cold adaptation in chickens appears to have evolved in parallel through different pathways involving different complements of genes, similar to high-altitude adaptation in Tibetan chickens (Wang et al., 2015). Overall, our study revealed the population genetic structure and some PSGs in CA chickens that may be important in cold adaptation, contributing to our understanding of cold adaptation in poultry and providing a molecular basis for poultry breeding.

DATA AVAILABILITY

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

X.Z. and Y.J. conceived and supervised the study. N.Y.X. performed most analyses with contributions from M.L., M.G.,

and P.P.B. W.S. and J.L. prepared the samples. N.Y.X. wrote the manuscript with input from the other authors. X.Z., Y.J., and H.A.N. revised the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Adolph EF, Molnar GW. 1946. Exchanges of heat and tolerances to cold in men exposed to outdoor weather. *American Journal of Physiology*, **146**(4): 507−537.

Ai HS, Fang XD, Yang B, Huang ZY, Chen H, Mao LK, et al. 2015. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nature Genetics*, **47**(3): 217−225.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**(9): 1655−1664.

Barnett SA. 1959. The skin and hair of mice living at a low environmental temperature. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*, **44**(1): 35−42.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114−2120.

Boulant JA, Dean JB. 1986. Temperature receptors in the central nervous system. *Annual Review of Physiology*, **48**: 639−654.

Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*, **325**(5938): 318−321.

Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*, **81**(5): 1084−1097.

Cai YD, Fu WW, Cai DW, Heller R, Zheng ZQ, Wen J, et al. 2020. Ancient genomes reveal the evolutionary history and origin of cashmere-producing goats in China. *Molecular Biology and Evolution*, **37**(7): 2099−2109.

Cannon B, Nedergaard J. 2004. Brown adipose tissue: function and physiological significance. *Physiological Reviews*, **84**(1): 277−359.

Chung SS, MacPhee KG, Goodridge AG. 1999. Effect of the CCAAT/enhancer binding protein on expression of the gene for chicken malic enzyme. *Archives of Biochemistry and Biophysics*, **364**(1): 30−41.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156−2158.

Ekarius C. 2007. Storey's Illustrated Guide to Poultry Breeds: Chickens, Ducks, Geese, Turkeys, Emus, Guinea Fowl, Ostriches, Partridges, Peafowl, Pheasants, Quails, Swans. Massachusetts: Storey Publishing.

Fumagalli M, Moltke I, Grarup N, Racimo F, Bjerregaard P, Jørgensen ME, et al. 2015. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*, **349**(6254): 1343−1347.

Hagedorn M. 2011. PRCP: a key to blood vessel homeostasis. *Blood*, **117**(14): 3705−3706.

Gburcik V, Cawthorn WP, Nedergaard J, Timmons JA, Cannon B. 2012. An essential role for Tbx15 in the differentiation of brown and "brite" but not white adipocytes. *American Journal of Physiology-Endocrinology and Metabolism*, **303**(8): E1053−E1060.

Hallmark B, Karafet TM, Hsieh PH, Osipova LP, Watkins JC, Hammer MF. 2019. Genomic evidence of local adaptation to climate and diet in indigenous Siberians. *Molecular Biology and Evolution*, **36**(2): 315−327.

Herrera MB, Kraitsek S, Alcalde JA, Quiroz D, Revelo H, Alvarez LA, et al. 2020. European and Asian contribution to the genetic diversity of mainland South American chickens. *Royal Society Open Science*, **7**(2): 191558.

Hofmann T, Schaefer M, Schultz G, Gudermann T. 2002. Subunit composition of mammalian transient receptor potential channels in living cells. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(11): 7461−7466.

Huan JL, Wang LS, Xing L, Qin XJ, Feng LB, Pan XF, et al. 2014. Insights into significant pathways and gene interaction networks underlying breast cancer cell line MCF-7 treated with 17β-Estradiol (E2). *Gene*, **533**(1): 346−355.

Hung BS, Wang XQ, Rothnagel JA, Cam GR. 2001. Characterization of mouse *Frizzled-3* expression in hair follicle development and identification of the human homolog in keratinocytes. *Journal of Investigative Dermatology*, **116**(6): 940−946.

Im JY, Lee KW, Won KJ, Kim BK, Ban HS, Yoon SH, et al. 2016. DNA damage-induced apoptosis suppressor (DDIAS), a novel target of NFATc1, is associated with cisplatin resistance in lung cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, **1863**(1): 40−49.

Innan H, Kim Y. 2008. Detecting local adaptation using the joint sampling of polymorphism data in the parental and derived populations. *Genetics*, **179**(3): 1713−1720.

Jiang P, Du WJ, Mancuso A, Wellen KE, Yang XL. 2013. Reciprocal regulation of p53 and malic enzymes modulates metabolism and senescence. *Nature*, **493**(7434): 689−693.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**(6): 587−589.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, **47**(W1): W256−W259.

Li DY, Che TD, Chen BL, Tian SL, Zhou XM, Zhang GL, et al. 2017. Genomic data for 78 chickens from 14 populations. *GigaScience*, **6**(6): gix026.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**(14): 1754−1760.

Li Y, Wu DD, Boyko AR, Wang GD, Wu SF, Irwin DM, et al. 2014. Population variation revealed high-altitude adaptation of Tibetan mastiffs. *Molecular Biology and Evolution*, **31**(5): 1200−1205.

Librado P, Der Sarkissian C, Ermini L, Schubert M, Jónsson H, Albrechtsen A, et al. 2015. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(50): E6889−E6897.

Liu SP, Lorenzen ED, Fumagalli M, Li B, Harris K, Xiong ZJ, et al. 2014. Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell*, **157**(4): 785−794.

Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, et al. 2014. Adaptations to climate-mediated selective pressures in sheep. *Molecular Biology and Evolution*, **31**(12): 3324−3343.

Lynch VJ, Bedoya-Reina OC, Ratan A, Sulak M, Drautz-Moses DI, Perry GH, et al. 2015. Elephantid genomes reveal the molecular bases of woolly mammoth adaptations to the arctic. *Cell Reports*, **12**(2): 217−228.

Marks M, Pennimpede T, Lange L, Grote P, Herrmann BG, Wittler L. 2016.

Analysis of the *Fam181* gene family during mouse development reveals distinct strain-specific expression patterns, suggesting a role in nervous system development and function. *Gene*, **575**(2): 438−451.

Masoro EJ. 1966. Effect of cold on metabolic use of lipids. *Physiological Reviews*, **46**(1): 67−101.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297−1303.

McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. 2008. Runs of homozygosity in European populations. *American Journal of Human Genetics*, **83**(3): 359−372.

Miao BP, Wang Z, Li YX. 2017. Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. *Molecular Biology and Evolution*, **34**(3): 734−743.

Miao YW, Peng MS, Wu GS, Ouyang YN, Yang ZY, Yu N, et al. 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity*, **110**(3): 277−282.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**(1): 268−274.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan YP, et al. 2012. Ancient admixture in human history. *Genetics*, **192**(3): 1065−1093.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genetics*, **2**(12): e190.

Pongratz RL, Kibbey RG, Shulman GI, Cline GW. 2007. Cytosolic and mitochondrial malic enzyme isoforms differentially control insulin secretion. *Journal of Biological Chemistry*, **282**(1): 200−207.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**(3): 559−575.

Qin Z, Ren FL, Xu XL, Ren YM, Li HG, Wang YY, et al. 2009. ZNF536, a novel zinc finger protein specifically expressed in the brain, negatively regulates neuron differentiation by repressing retinoic acid-induced gene transcription. *Molecular and Cellular Biology*, **29**(13): 3633−3643.

Qiu Q, Zhang GJ, Ma T, Qian WB, Wang JY, Ye ZQ, et al. 2012. The yak genome and adaptation to life at high altitude. *Nature Genetics*, **44**(8): 946−949.

Rubin CJ, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, et al. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, **464**(7288): 587−591.

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, et al. 2006. Positive natural selection in the human lineage. *Science*, **312**(5780): 1614−1620.

Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, et al. 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Human Genetics*, **112**(4): 387−399.

Signore AV, Yang YZ, Yang QY, Qin G, Moriyama H, Ge RL, et al. 2019. Adaptive changes in hemoglobin function in high-altitude Tibetan canids were derived via gene conversion and introgression. *Molecular Biology and Evolution*, **36**(10): 2227−2237.

Sinding MHS, Gopalakrishnan S, Ramos-Madrigal J, De Manuel M, Pitulko VV, Kuderna L, et al. 2020. Arctic-adapted dogs emerged at the Pleistocene–Holocene transition. *Science*, **368**(6498): 1495−1499.

Smith RE. 1962. Cold acclimation—an altered steady state. *Journal of the American Medical Association*, **179**(12): 948−954.

Storey AA, Ramírez JM, Quiroz D, Burley DV, Addison DJ, Walter R, et al. 2007. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(25): 10335−10339.

Szpiech ZA, Hernandez RD. 2014. Selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution*, **31**(10): 2824−2827.

Thyme SB, Pieper LM, Li EH, Pandey S, Wang YQ, Morris NS, et al. 2019. Phenotypic landscape of schizophrenia-associated genes defines candidates and their shared functions. *Cell*, **177**(2): 478−491.

Wang K, Li MY, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, **38**(16): e164.

Wang MS, Li Y, Peng MS, Zhong L, Wang ZJ, Li QY, et al. 2015. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Molecular Biology and Evolution*, **32**(7): 1880−1889.

Wang MS, Thakur M, Peng MS, Jiang Y, Frantz LAF, Li M, et al. 2020. 863 genomes reveal the origin and domestication of chicken. *Cell Research*, **30**(8): 693−701.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**(6): 1358−1370.

Wise Jr EM, Ball EG. 1964. Malic enzyme and lipogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, **52**(5): 1255−1263.

Wu HG, Guang XM, Al-Fageeh MB, Cao JW, Pan SK, Zhou HM, et al. 2014. Camelid genomes reveal evolution and adaptation to desert environments. *Nature Communications*, **5**(1): 5188.

Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, et al. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research*, **39**(S2): W316−W322.

Yang J, Li WR, Lv FH, He SG, Tian SL, Peng WF, et al. 2016. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution*, **33**(10): 2576−2592.

Yang ZH, Shi H, Ma PC, Zhao SL, Kong QH, Bian TH, et al. 2018. Darwinian positive selection on the pleiotropic effects of *KITLG* explain skin pigmentation and winter temperature adaptation in Eurasians. *Molecular Biology and Evolution*, **35**(9): 2272−2283.

Yi GQ, Qu LJ, Liu JF, Yan YY, Xu GY, Yang N. 2014. Genome-wide patterns of copy number variation in the diversified chicken genomes using next-generation sequencing. *BMC Genomics*, **15**(1): 962.

Zhang C, Dong SS, Xu JY, He WM, Yang TL. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, **35**(10): 1786−1788.