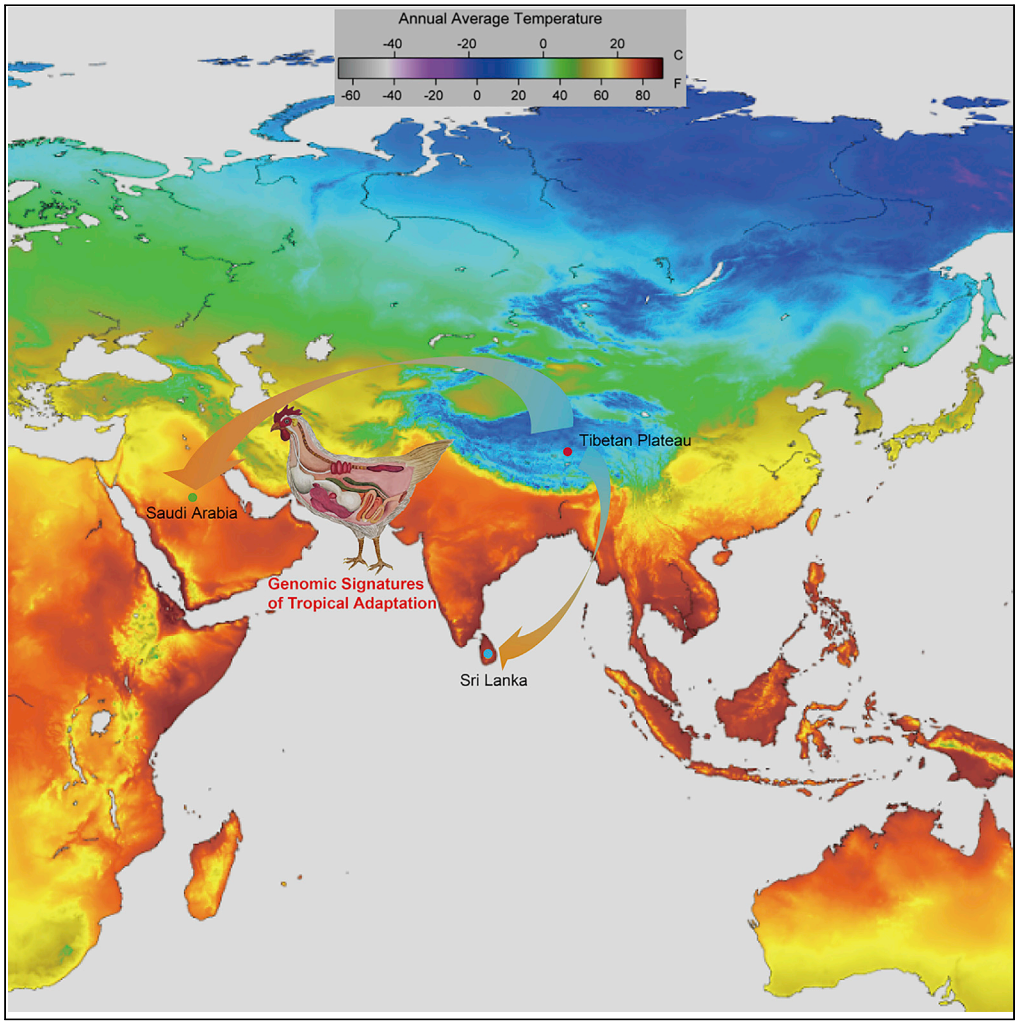
iScience

CellPress
OPEN ACCESS

Article

Genomic Analyses Reveal Genetic Adaptations to Tropical Climates in Chickens



Shilin Tian, Xuming Zhou, Tashi Phuntsok, ..., Chunyou Ning, Diyan Li, Huabin Zhao

diyanli@sicau.edu.cn (D.L.)
huabinzhao@whu.edu.cn (H.Z.)

HIGHLIGHTS

Gene flow was detected from Tibetan chickens to Sri Lankan and Saudi Arabian chickens

Twelve genes showed chicken adaptations to both tropical desert and monsoon climates

Advantageous alleles of *TLR7* and *ZC3HAV1* showed immune adaptation in chickens

Convergent adaptation to tropical desert climate was found between birds and mammals

Tian et al., iScience *23*, 101644
November 20, 2020
https://doi.org/10.1016/j.isci.2020.101644

iScience

CellPress
OPEN ACCESS

Article

Genomic Analyses Reveal Genetic Adaptations to Tropical Climates in Chickens

Shilin Tian,[1,6] Xuming Zhou,[3,6] Tashi Phuntsok,[4] Ning Zhao,[4] Dejing Zhang,[5] Chunyou Ning,[2] Diyan Li,[2,*] and Huabin Zhao[1,4,7,*]

SUMMARY

The genetic footprints of adaptations to naturally occurring tropical stress along with domestication are poorly reported in chickens. Here, by conducting population genomic analyses of 67 chickens inhabiting distinct climates, we found signals of gene flow from Tibetan chickens to Sri Lankan and Saudi Arabian breeds and identified 12 positively selected genes that are likely involved in genetic adaptations to both tropical desert and tropical monsoon island climates. Notably, in tropical desert climate, advantageous alleles of *TLR7* and *ZC3HAV1*, which could inhibit replication of viruses in cells, suggest immune adaptation to the defense against zoonotic diseases in chickens. Furthermore, comparative genomic analysis showed that four genes (*OC90*, *PLA2G12B*, *GPR17* and *TNFRSF11A*) involved in arachidonic acid metabolism have undergone convergent adaptation to tropical desert climate between birds and mammals. Our study offers insights into the genetic mechanisms of adaptations to tropical climates in birds and other animals and provides practical value for breeding design and medical research on avian viruses.

INTRODUCTION

Chickens were domesticated from red jungle fowl (RJF) at least 4,000–4,500 years ago and have multiple origins within the geographic range of their wild ancestors in South and Southeast Asia (Fumihito et al., 1996; Larson and Fuller, 2014; Liu et al., 2006). Domestication of chickens is closely related to human activities. For instance, the complex and ancient interactions resulting from the Silk Route have facilitated maritime and terrestrial intercontinental translocation of several domestic chicken breeds among the Middle East and South, Southeast, and East Asia (Boivin and Fuller, 2009; Boulnois, 2004; Fuller and Boivin, 2009). After a long period of artificial and natural selection by humans, chicken breeds have evolved genetic adaptations to their local environmental conditions, such as hot and arid climates and high-altitude environments (Lawal et al., 2018; Wang et al., 2015). It is well known that the detection of genomic differences can shed light on the genetic basis of adaptation to diverse environments and provide insights into functionally important genetic variants (Andersson and Georges, 2004). The completion of the chicken genome sequence and genetic variation maps has promoted the exploration and research of chicken genetic mechanisms using whole-genome-based strategies (Hillier et al., 2014; Johnsson et al., 2012; Li et al., 2017a; Qanbari et al., 2019; Rubin et al., 2010; Wong et al., 2004). Conversely, considering their short reproductive and growth periods and wide distribution, chickens can be used as ideal models to study genetic adaptations to environments.

Environmental pressure is an important driver shaping the animal genome, which is also affected by pathogens (Fumagalli et al., 2011), disease (Sabeti et al., 2007), diet (Jiao et al., 2019), altitude (Huerta-Sanchez et al., 2013), and climate (Ai et al., 2015; Frichot et al., 2013; Lv et al., 2014). High altitudes are among the most challenging extreme environments, and high-altitude adaption has been widely studied in humans (Simonson et al., 2010), Tibetan chickens (TCs) (Wang et al., 2015; Zhang et al., 2016), yaks (Qiu et al., 2012), pigs (Li et al., 2013), ground tit (Qu et al., 2013), and sheep (Yang et al., 2016). Recently, genomic studies have also reported the genomic responses to acclimation to hot and arid environments in camels (Jirimutu et al., 2012; Wu et al., 2014), sheep (Yang et al., 2016), and goats (Kim et al., 2016). Although the acute heat response has been studied in commercial chickens (Lara and Rostagno, 2013), to the best of our knowledge, genetic adaptations to naturally occurring hot and arid environments in domestic chickens

[1]Department of Ecology, Tibetan Centre for Ecology and Conservation at WHU-TU, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, 299 Bayi Road, Wuhan, Hubei 430072, China

[2]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu 611130, China

[3]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Beijing 100101, China

[4]Laboratory of Extreme Environmental Biological Resources and Adaptive Evolution, Research Center for Ecology, College of Science, Tibet University, Lhasa 850000, China

[5]Novogene Bioinformatics Institute, Beijing 100015, China

[6]These authors contributed equally

[7]Lead Contact

*Correspondence: diyanli@sicau.edu.cn (D.L.), huabinzhao@whu.edu.cn (H.Z.)

https://doi.org/10.1016/j.isci.2020.101644

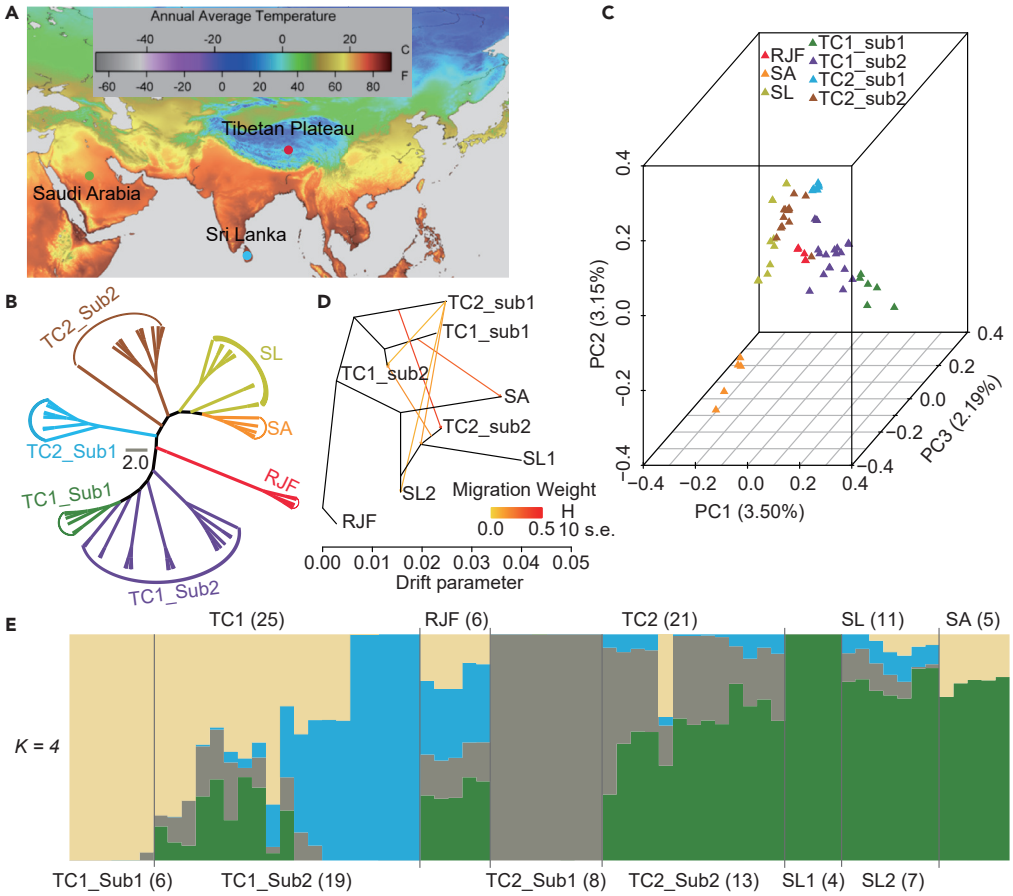


Figure 1. Phylogeny, Genetic Structure, and Inferred Migration Events of Chicken Breeds/Clusters

(A) Geographic variation of annual average temperature for three indigenous chickens.

(B) Phylogenetic tree constructed by the maximum likelihood (ML) algorithm. Different colors represent various breeds/clusters of chickens. Abbreviations of chicken populations: SL, Sri Lanka; SA, Saudi Arabia; RJF, red jungle fowl; TC1_Sub1, subpopulation 1 of Tibetan chicken population 1; TC1_Sub2, subpopulation 2 of Tibetan chicken population 1; TC2_Sub1, subpopulation 1 of Tibetan chicken population 2; TC2_Sub2, subpopulation 2 of Tibetan chicken population 2. See also Table S5 for more detail of each population.

(C) Principal-component analysis (PCA) plot for chickens. The fraction of the variance explained is 3.50% for eigenvector 1, 3.15% for eigenvector 2, and 2.19% for eigenvector 3.

(D) Phylogenetic network of the inferred migration events among the eight chicken breeds/clusters with six migration events. Migration arrows are colored according to their weight. The scale bar shows 10 times the average standard error of the entries in the sample covariance matrix.

(E) Genetic structure of chickens with K = 4 ancestral populations.

remain poorly known. In view of the current and future global effects of climate change, the genetic basis of adaptation to hot or arid climates in chickens is of practical value for breeding design and medical research (Diamond, 2002).

In this study, we performed a large-scale genomic analysis of chicken populations, including 36 high-altitude TCs from our work (Li et al., 2017a), 5 RJFs (Wang et al., 2015), 10 TCs (Wang et al., 2015), 11 Sri Lankan indigenous chickens (SL) (Lawal et al., 2018), and 5 Saudi Arabian indigenous chickens (SA) (Lawal et al., 2018), from habitats in harsh environments (high-altitude, tropical desert, and tropical monsoon island climates) (Figure 1A and Table S1). Genome-wide analysis for selective sweeps was performed to elucidate the genetic mechanisms of cutaneous melanin formation (thermotolerance), hormonal regulation, angiogenesis, vasodilation, and mitochondrial respiration, which may work together to enable chickens to adapt to tropical desert and tropic monsoon island climates. We also revealed the immune adaptation to the defense against zoonotic diseases in chickens. Additionally, we revealed convergent adaptation to tropical

Category	Core Set	RJF	TC1_Sub1	TC1_Sub2	TC2_Sub1	TC2_Sub2	SL	SA
Number of individuals	67	5	6	19	8	13	11	5
Number of SNPs	10,691,650	8,409,391	8,254,044	10,313,588	8,456,758	9,727,520	9,341,260	7,631,063
Unique SNPs	Not applicable	1,154,609 (13.73%	253,399 (3.07%)	854,996 (8.29%)	405,079 (4.79%)	516,531 (5.31%)	560,476 (6.00%)	312,110 (4.09%)
Nonsynonymous[a]	56,026	39,294	41,938	53,591	43,016	50,360	47,525	37,518
Nonsyn/Syn ratio[b]	0.453	0.416	0.438	0.450	0.439	0.447	0.444	0.433
Number of InDels	2,085,509	1,179,662	952,921	1,389,263	1,018,893	1,224,798	1,165,978	877,956
Unique InDels	Not applicable	21,706 (1.84%)	4,002 (0.42%)	15,282 (1.10%)	7,947 (0.78%)	9,798 (0.8%)	10,494 (0.90%)	4,390 (0.50%)
Frameshift InDels[c]	6,973	2,784	2,531	3,424	2,621	3,162	2,557	1,844

Table 1. Summary of Genomic Variations in 67 Chickens

[a]Nonsynonymous variants include missense, stop-gain, and stop-loss variants.

[b]Nonsyn/Syn ratio indicates nonsynonymous to synonymous ratio.

[c]Frameshift InDels include frameshift InDels and stop-gain and stop-loss variants.

desert climate between birds and mammals. Our study adds to knowledge on genetic mechanisms of adaptations to tropical climates in vertebrates and is of fundamental importance for the potential application in functional genomics, breeding design, and the development of medical models for avian viruses.

RESULTS AND DISCUSSION

Characterization of Genomic Variants

After performing strict data quality filtering, a total of 1.56 Tb (terabases) of high-quality genome sequences from 67 chickens were mapped to the chicken reference genome (GRCg6a, GCF_000002315.6), resulting in an average of 94.74% non-gap genome coverage and ~20.25-fold average depth (Table S2). We identified 12.78 million biallelic high-quality population genomic variants (PGVs) that contained 10.69 million SNPs and 2.09 million InDels with lengths ranging from 1 to 306 bp (Tables S3 and S4). Overall, more than half of the PGVs (52.63% and 52.65% for total SNPs and InDels, respectively) were located in intronic regions, followed by intergenic regions, with 39.28% and 40.00% for SNPs and InDels, respectively. The densities of the PGVs also showed a higher ratio in the promoter region than in the coding region (Figures S2 and S3). We annotated 180,729 SNPs (1.67% of the total) and 10,901 InDels (0.52% of the total) located within 15,370 protein-coding genes (89.31% of the total 17,209 genes). We identified a total of 55,480 missense, 6,798 frameshift, 629 stop-gain, and 92 stop-loss PGVs that caused amino acid changes, transcript elongation, or premature stopping, leading to structural changes in 11,407 genes (66.29% of total genes). We also observed an enrichment of $\leq$30 bp in coding sequences (95.13% of the total) that are multiples of 3 bp, which is expected to preserve the open reading frame (Figure S3).

The ratio of nonsynonymous to synonymous changes was estimated to be lower in chickens (0.45) than in other agricultural economic animals (Table 1), such as pigs (0.68) (Li et al., 2013), sheep (0.66) (Yang et al., 2016), and cattle (1.21) (Daetwyler et al., 2014), possibly indicating a weaker artificial selection accumulation effect in chickens than in other animals due to the short domestication time (Larson and Fuller, 2014). We also found a greater proportion of unique PGVs (13.73% and 1.84% for SNPs and InDels, respectively) in RJF than in other chicken groups (Table 1), which is consistent with the loss of genetic diversity during domestication (Larson and Fuller, 2014).

Phylogeny and Introgression in Chickens

On the basis of the maximum likelihood tree and principal-components analysis (PCA), there were five distinct major clusters (i.e., TC1, TC2, SL, SA, and RJF) (Figures 1B and 1C). We further used an admixture model (FRAPPE) to explore the level of shared ancestry among individuals within each breed (Tang et al., 2005). As the ancestry component (*K* value) increased, we observed a simultaneous increase in the composition of ancestors in the RJF (Figure S4). This result is expected because the genetic materials of other

chicken breeds were from components of RJF ancestry (Figure 1C). The subcluster ancestry composition showed considerable genetic difference between the TC1 and TC2 clusters, which is consistent with a previous report showing that multiple origins of TCs can be traced to two independent sources (Wang et al., 2015). We also observed that some SL individuals always maintained a single ancestry component and others showed a simultaneous increase in the composition of ancestors with increasing *K* value. We speculated that this single ancestor might be the result of a continent-island genetic model in Sri Lanka. We further identified four admixture category clusters (TC1_Sub1, TC1_Sub2, TC2_Sub1 and TC2_Sub2) on the basis of the two distinct TC clusters that were previously observed (Wang et al., 2015), according to phylogenetic analysis (Figure 1B). We found that the TC2_Sub2 cluster had more shared ancestry components with the SA and SL breeds than the other clusters (Figure 1E). Based on TreeMix (Pickrell and Pritchard, 2012), we observed signals of genetic flow from TCs to SA and SL breeds (Figures 1D and S5). Subsequently, the *D*-statistic (Durand et al., 2011; Green et al., 2010; Patterson et al., 2012) also indicated that the SA and SL breeds have a significant level of introgression with TC2_Sub2 chickens (Figure S6, Tables S6 and S7). Previous studies showed that Southwest China was one of the origin regions of chickens (Liu et al., 2006; Wang et al., 2020) and that trade activities have facilitated maritime and terrestrial intercontinental translocation of several domestic chicken breeds among the Middle East and South, Southeast, and East Asia (Figure S7) (Boivin and Fuller, 2009; Boulnois, 2004; Fuller and Boivin, 2009). Our admixture analysis supported the hypothesis that Chinese indigenous chicken breeds may have been transported to Sri Lanka and Saudi Arabia via the land and maritime Silk Route voyages (Fuller and Boivin, 2009).

Genomic Sweep Footprints in Chickens

The SA indigenous breed is well adapted to hot and arid environments because those breeding centers are located in the Eastern Province of Saudi Arabia and belong to a tropical desert climate zone with high solar radiation, very large diurnal ambient temperature differences, high daytime temperatures (ranging from 21.2°C to 50.8°C), frequent floating dust weather, and extremely low rainfall (74 mm per year) (Figures 1A and Table S1). The SL indigenous breed is from the Puttalam district of Sri Lanka, which is a tropical monsoon island country with an average annual rainfall of ~1,000 mm and a temperature of 27°C, whereas TCs live in a temperate humid plateau climate in China. To identify genomic sweep footprints in chickens, we applied a modified population-branch statistic (PBS) method, which was very powerful in detecting incomplete selective sweeps over short divergence times (Simonson et al., 2010; Zhan et al., 2014). This approach was designed to take advantage of three outgroups (TC2_Sub2, TC2_Sub1 and RJFs), aiming to identify genomic regions under selection in SA/SL breeds. We identified 24.18 Mb (2.30% of the genome) and 15.58 Mb (1.48% of the genome), encompassing 723 and 464 positively selected genes (PSGs), in SA and SL breeds, respectively. Next, we employed three methods: log2 ($\theta\pi$ ratio) (Nei and Li, 1979), cross-population extended haplotype homozygosity (XP-EHH) (Sabeti et al., 2007), and cross-population composite likelihood ratio (XP-CLR) (Chen et al., 2010), although these methods could not remove noise signals from domestication if compared with RJFs (the ancestral population of domestic chickens). Ideally, for our purpose of detecting tropical adaptation, the best control group would be those breeds with a similar level of domestication and a distinct level of environmental temperature compared with study groups. We selected TC2_Sub2 as the control group, because our analyses of phylogeny and introgression showed that there is significant genetic mixing between SA/SL breeds and TC2_Sub2 cluster (Figures 1B and 1D). Thus, using the three methods, we performed a genome-wide selection sweep scanning with two comparisons, i.e., SA and SL breeds were compared with TC2_Sub2 cluster, respectively. Specifically, 376 and 396 PSGs were identified by XP-CLR, 233 and 305 from $\log_2$ ($\theta\pi$ ratio), and 383 and 267 from XP-EHH in SA and SL breeds, respectively (Figure 2). Combined with the PBS method, only 87 and 45 PSGs were detected by all four methods in SA and SL breeds, respectively (Figures S8 and S9). The low number of overlapped PSGs could have resulted from different signatures of population variations in the four methods (Akey, 2009; Sabeti et al., 2006; Wang et al., 2016). Finally, we obtained 942 PSGs in SA breed and 923 PSGs in SL breed after integrating all candidate selection regions from the four methods, and 150 of them were found in both SA and SL breeds (Figure S10).

Genetic Adaptations to Tropical Climates in Chickens

Among the PSGs in the SA and SL breeds, we found significant enrichment of functional pathways ($p < 0.05$) that directly or indirectly influence several traits that are critical for survival under heat stress in tropical climates, e.g., 18 PSGs involved in the vascular smooth muscle contraction pathway (VSMC, 12 PSGs in SA breed and 9 PSGs in SL breed), 19 PSGs located in the melanogenesis pathway (13 in SA and 9 in SL); 27 PSGs located in the adrenergic signaling in cardiomyocytes pathway (12 in SA and 18 in SL), and 26

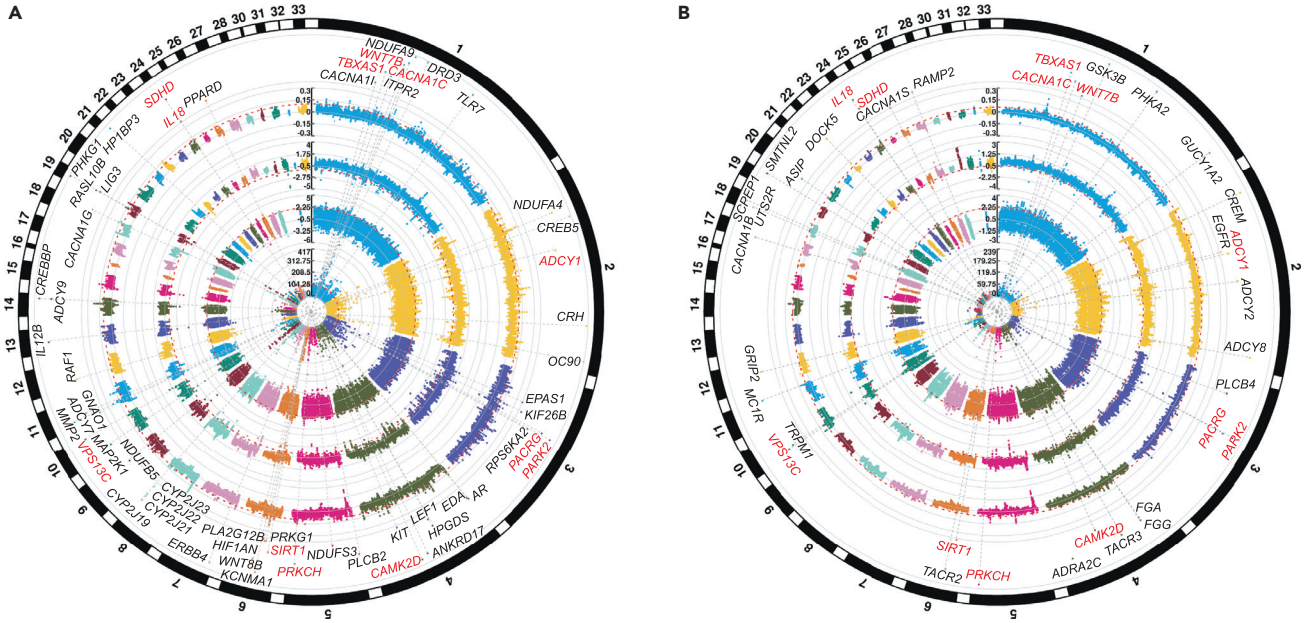


Figure 2. Selective Sweep Signals in Two Indigenous Chickens

(A) Circular Manhattan plot of genome-wide selection sweep signals for SA breed compared with the TC2_Sub2 cluster.

(B) Circular Manhattan plot of genome-wide selection sweep signals for SL breed compared with the TC2_Sub2 cluster. From inside to outside, the circular Manhattan plots show the PBS value, $\log_2\theta\pi$ ratio, XP-CLR score, and XP-EHH value. The thresholds for identifying candidate selection sweep regions were $p < 0.01$ (Z-test) for the PBS values and 1% outliers for $\log_2\theta\pi$ ratio values, XP-CLR score, and XP-EHH values Twelve PSGs simultaneously present in the SA and SL breeds are shown in red.

PSGs involved in the calcium signaling pathway (13 in SA and 17 in SL) (Tables S8 and S10). We found five PSGs (*ADCY1, ADCY2, ADCY7, ADCY8,* and *ADCY9*) associated with adenylyl cyclase (AC), affecting the formation of cAMP; these genes also play roles in the above four pathways (Figure 3). Adrenergic signaling in cardiomyocytes plays a pivotal role in regulating cardiac function in response to ever-changing environments. Systolic and diastolic function and heart rate are primarily controlled by the adrenergic and muscarinic systems in response to varying physiological demands (Schaub et al., 2006). The calcium signaling pathway acts as an intracellular messenger for signal transduction that is primarily modulated by adrenergic control of phosphorylation involved in intracellular Ca^{2+} handling (Schaub et al., 2006). We identified five PSGs (*CACNA1B, CACNA1C, CACNA1G, CACNA1I,* and *CACNA1S*) that are critical for calcium channel function, especially *CACNA1C,* which encodes the L-type voltage-gated calcium channel Cav1.2 and is under positive selection in both SA and SL breeds. This protein activates the VSMC pathway (Figure 3). The *CAMK2D* gene is under selection in chickens living in different tropical climates. It encodes calmodulin-dependent protein kinase-II (CaMKII), which is activated in response to adrenergic stimulation (Grimm and Brown, 2010), affects a variety of Ca^{2+}-handling proteins (Schaub et al., 2006) and regulates melanin synthesis (Figure 3). In addition, we found that 68 PSGs are involved in central nervous system developmental processes, such as regulation of nervous system development, neurogenesis, neuron remodeling, cholinergic synapse, synaptic transmission regulation, cerebellar development, and axonal transport (Tables S9 and S11). The central nervous system of avian organisms exhibits high thermosensitivity in the control of thermoregulatory responses (Boulant and Dean, 1986). A deficit of thermal stimulation causes alterations, retardation, and deleterious effects in the central nervous system (Ahmed, 2005). Therefore, PSGs involved in these relevant functional ontologies suggested that nerve development, cutaneous melanin formation, hormonal regulation, angiogenesis, and vasodilation may have worked together to enable chickens to adapt to heat stress in tropical climates, reflecting common and local adaptations.

PSGs Associated with Thermotolerance

The cutaneous melanin pigment plays a critical role in protecting animals against the harmful effects of solar radiation, and Wnt signaling plays a crucial role in orchestrating epidermal stratification. Of the 19 PSGs located in the melanogenesis pathway, three PSGs (*WNT7B, WNT8B,* and *GNAO1*) in SA breed and two (*WNT7B and GSK3B*) in SL breed are involved in Wnt signaling, which induces epidermal stratification and regeneration

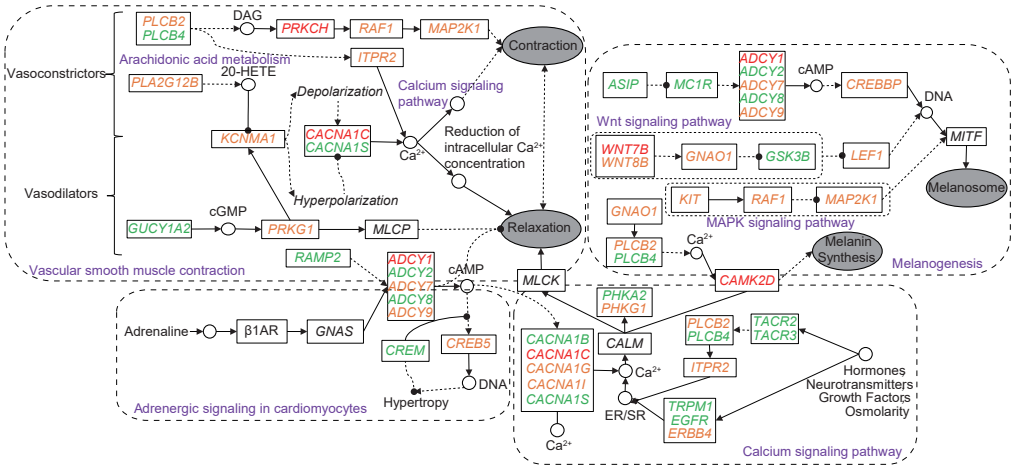


Figure 3. Genetic Adaptations of Chickens to Tropical Climates

Dotted arrows indicate an indirect effect. Dotted lines indicate indirect inhibition. Solid arrows indicate activation. Solid lines indicate association. The PSGs simultaneously present in the SA and SL breeds are shown in red. The PSGs in SA breed are shown in brown. The PSGs in SL breed are shown in green. The names of the KEGG pathways are shown in purple.

(Zhu et al., 2014); three PSGs (*KIT*, *RAF1*, and *MAP2K1*) located only in SA breed are involved in the MAPK signaling pathway, which is an upstream regulator of melanogenesis and melanoma angiogenesis (Figure 3). In particular, the interaction of *KIT* with *MITF* has been reported to confer resistance to melanosis and the ability to cope with solar radiation (Satzger et al., 2008). One PSG (*LIG3*) in SA breed was involved in "negative regulation of mitochondrial DNA replication" and "base-excision repair, DNA ligation" (Table S9); two PSGs (*SIRT1* and *MC1R*) of SL were significantly enriched in "UV-damage excision repair," of which *SIRT1* was selected in SA breed simultaneously based on the XP-EHH value (Table S11). We found significantly higher ($p < 10^{-16}$, Mann–Whitney U test) selective pressure in the SA breed than in the SL breed for *SIRT1*, indicating that the genome characteristics were sensitive to solar radiation (Figure S11).

PSGs Associated with the Circulatory and Respiratory Systems

The living climates of the SA and SL breeds are characterized by very large diurnal ambient temperature differences and high daytime temperatures caused by intense solar radiation (Table S1). Indigenous chickens use a variety of solutions to maintain core body temperature and internal homeostasis, such as convective and evaporative heat loss through diastolic blood vessels (Lara and Rostagno, 2013; Mutaf et al., 2009). Therefore, the balance between vascular contraction and relaxation plays a pivotal role in dissipating excessive heat through regulation and maintenance of blood pressure to maintain core body temperature and internal homeostasis in chickens. Eighteen PSGs involved in the VSMC pathway encode seven vasoconstrictors, eight vasodilators, two voltage-dependent calcium channel L-type alpha proteins, and one calcium-activated potassium channel subunit alpha protein. Notably, three genes (*PRKCH*, *ADCY1*, and *CACNA1C*) were simultaneously selected by SA and SL breeds and separately activated vasoconstrictors, vasodilators, and voltage-dependent calcium channels (Figure 4A). Furthermore, we performed quantitative real-time PCR (qRT-PCR) to investigate changes in gene expression level between heat stress and normal temperature using the C2C12 cell line. We found significantly higher expression levels of *CACNA1C* under heat treatments (Figure 4B). Subsequently, we analyzed the population genotype frequency in a large dataset of 845 *Gallus gallus* accessions (Wang et al., 2020). We first divided 845 accessions into four groups, of which 109 domesticated chickens live in tropical climates (TRCs) (Table S12), 112 are TCs, 475 are other domesticated chickens (non-tropical chickens, or NTRCs), and 149 are RJFs. Based on the recently released population SNP data of chickens (http://bigd.big.ac.cn/chickensd/), we screened out 50,582 SNPs within *CACNA1C* gene and up-/downstream 500-kb regions based on 845 accessions. We discovered directional increased or reduced genotype frequency through a priority order with TRCs → NTRCs → TCs, as demonstrated by a linear regression model (Transparent Methods); especially 147 mutations (29.05%) within top 1% outlier of absolute *K* value belonged to the region of *CACNA1C* gene (Figure 4C). Indeed, 11 PSGs in SA breed and 14 PSGs in SL breed were significantly enriched ($p < 0.05$) in GO terms "regulation of blood circulation" (Table S9). More specifically, the significantly enriched GO terms

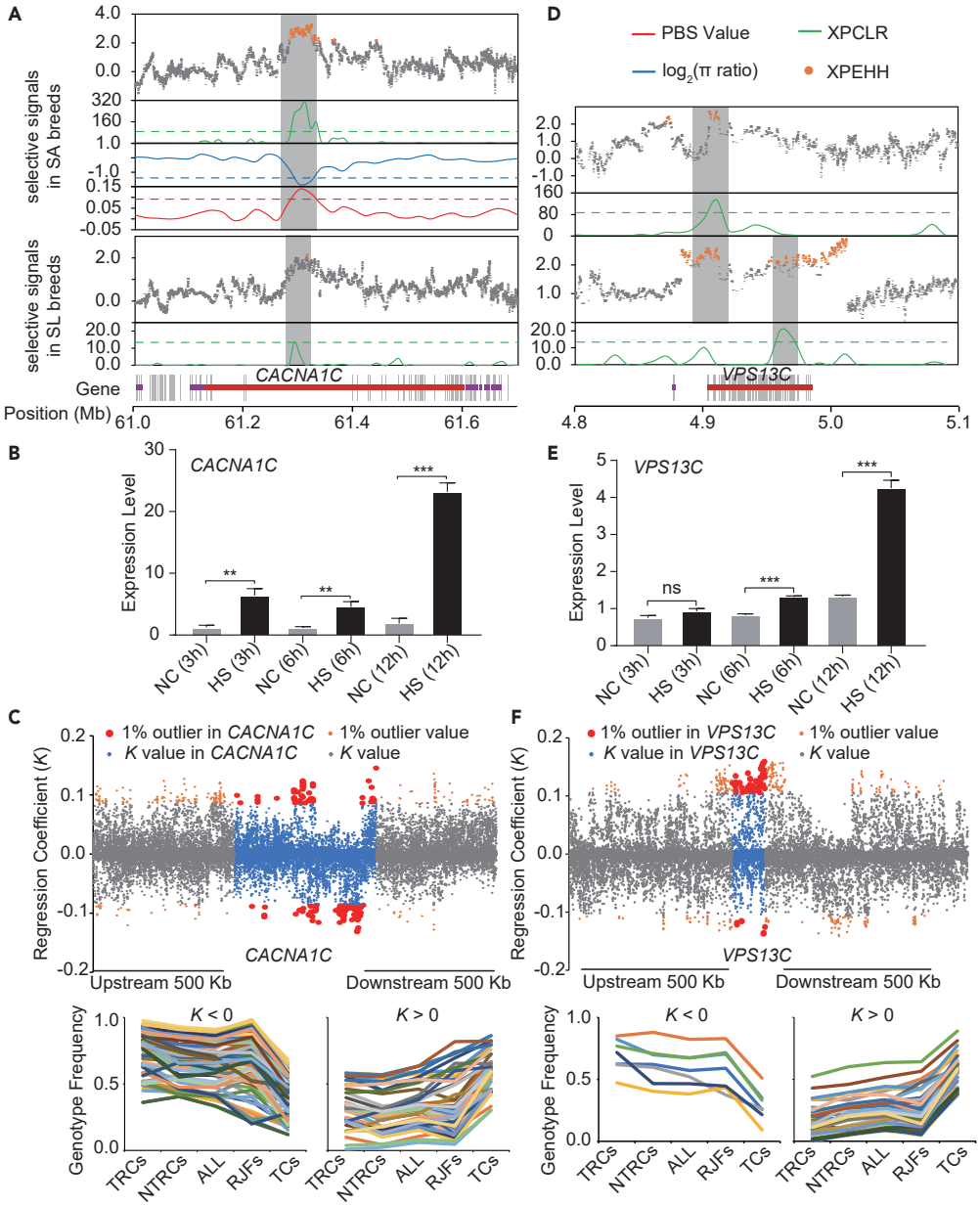


Figure 4. Genomic Regions with Strong Selective Sweep Signals in SA/SL Breeds

(A) Example of genes with strong selective signals in SA and SL breeds. Genomic regions located above the upper horizontal red, green, and yellow dashed lines and under the horizontal blue dashed line were termed regions with strong selective sweep signals (gray regions). Genome annotations are shown in the middle (gray bar, coding sequence; purple bar, gene body). The boundary of *CACNA1C* is marked in red.

(B) Quantitative real-time PCR (qRT-PCR) for the *CACNA1C* gene in C2C12 cells under 42°C for heat stress (HS) and 37°C for the normal control (NC) groups (n = 3 replicates per group; two-sided Student's t test (***: $p < 0.001$, **: $0.001 < p < 0.01$, ns: non-significant); data are represented as mean ± SEM).

(C) Genotype frequency analysis detects selection signals upstream and downstream 500 Kb around *CACNA1C* gene based on 845 chickens. Below left showed directional reduced frequency of 52 mutations and below right showed directional reduced frequency of 95 mutations across five chicken populations in *CACNA1C* gene.

(D) Selection signals in the SA breed for the *VPS13C* gene. The upper half is similar to (A).

(E) qRT-PCR for the *VPS13C* gene in C2C12 cells under HS and NC groups (n = 3 replicates per group; two-sided Student's t test (***: $p < 0.001$, **: $0.001 < p < 0.01$, *: $0.01 < p < 0.05$, ns: non-significant); data are represented as mean ± SEM).

Figure 4. *Continued*
(F) P genotype frequency analysis detects selection signals upstream and downstream 500 Kb around *VPS13C* gene based on 845 chickens. Below left showed directional reduced frequency of seven mutations and below right showed directional reduced frequency of 52 mutations across five chicken populations in *VPS13C* gene.

were "regulation of systemic arterial blood pressure by hormone" (*RASL10B, DRD3* and *RPS6KA2*) and "blood vessel maturation" (*MMP2* and *ANKRD17*) in SA breed, whereas GO terms included "regulation of blood vessel diameter" (*ADRA2C, FGG, FGA, SMTNL2, DOCK5, GRIP2, UTS2R,* and *SCPEP1*) in SL breed.

Heat stress causes excessive oxygen demand in animals (Portner, 2001), leading to increased blood flow in circulation and optimized oxygen delivery. Consistent with this phenomenon in SA breed, we discovered that three PSGs (*EPAS1, CREBBP,* and *HIF1AN*) are significantly involved in the "cellular response to hypoxia," whereas five PSGs (*SIRT1, PPARD, HP1BP3, SDHD,* and *HIF1AN*) were involved in the "cellular response to decreased oxygen levels" (Table S9). Among those genes, *EPAS1* was detected by PBS (PBS value = 0.096; p = 0.007), and was known as hypoxia-inducible factor 2α (*HIF-2α*), and *HIF1AN* was involved in transcriptional repression through interaction with *HIF1α*, *VHL,* and histone deacetylases. Also, a unique set of regulatory targets was acted by each of *HIF-1α* and *EPAS1* (Hu et al., 2003). It has been suggested that there is a close relationship between *EPAS1* and the regulation of red blood cell production (Percy et al., 2008; Yi et al., 2010). Indeed, we found one PSG (*KIT*) that is involved in "cellular response to erythropoietin" (Table S9). One SNP of *EPAS1* with the greatest genotype frequency difference between SA and TCs is located in the sixth intron, with a derived allele at 0% frequency in TCs and 70% in the SA breed, as well as $\leq$ 10% in SL breed and RJFs (Table S13). This is consistent with a previous finding that the intron 5 of *EPAS1* harbored a highly differentiated SNP associated with erythrocyte abundance to positively influence adaptation to hypoxia (Yi et al., 2010). Similarly, a previous study also found that a non-synonymous SNP in the exon 7 of *EPAS1* may contribute to the hypoxia adaptation of TCs (Li et al., 2017b). To test whether positive selection has acted on different regions of *EPAS1* for coping with heat stress and high altitude, we analyzed genotype frequency of *EPAS1* from 845 *G. gallus* accessions (Wang et al., 2020). In addition to one SNP in the sixth intron of EPAS1 identified in our dataset, two more SNPs located in the sixth intron with a significant frequency preference were found in tropical chickens using the large dataset of 845 accessions (Table S14). By contrast, the corresponding genotypes in TCs and RJFs were fixed, indicating that this trait was retained before the ancient divergence of chickens. However, after comparing these genotypes in 109 TRCs and 475 NTRCs with the background of RJFs and all 845 chickens, we observed significant deviations of p values using chi-square tests, indicating that the two SNPs in tropical chickens are not random and lineage-specific positive selection should have favored the tropical chickens (Table S15). Indeed, our PBS method was considered to be robust for identifying recent natural selection. Therefore, although the regulation mechanism of *EPAS1* by the intronic mutations remains to be discovered, these findings supported that the naturally occurring tropical stress may have led to the alleles targeted by natural selection, which may confer a functionally relevant adaptation to excessive oxygen demand. This result also indicated that different mutations of same genes may have occurred to meet similar physiological needs to cope with different environment stresses.

Additionally, chickens can expel internal heat by evaporation of moisture by increasing the frequency of panting (Collier and Gebremedhin, 2015). Correspondingly, we found that one PSG (*EDA*) was involved in "trachea submucosa development" in SA breed (Table S9). These functions play an important role in the regeneration and repair of airway epithelial cells (Tata et al., 2013). Therefore, the synergistic action of these biological functions may have enabled chickens to cope with heat stress in the heart, a trachea burden, and vascular pressure.

PSGs Associated with Mitochondrial Respiration

We also found that 15 PSGs and 5 PSGs are associated with the cellular component term "mitochondrial membrane" in SA and SL breeds, respectively (Tables S9 and S11). Under heat stress, oxygen levels in chicken body fluids may decrease, reflecting excessive oxygen demand (Portner, 2001). The adjustment of mitochondrial densities in addition to molecular or membrane adjustments appears crucial for maintaining aerobic scope and for shifting thermal tolerance. Of the eight PSGs *(NDUFS3, NDUFA4, NDUFB5, NDUFA9, SDHD, VPS13C, PARK2,* and *PACRG*) involved in the response to the mitochondrial respirasome based on the annotations in the PANTHER and UniProt databases, four (*SDHD, VPS13C, PARK2,* and *PACRG*) were under selection in both SA and SL breeds. *VPS13C* and *PARK2* play a role in mitochondrial maintenance, such as in the regulation of

mitochondrial respiration rates, suggesting compensatory adaptation aimed at preserving mitochondrial transmembrane potential levels (Lesage et al., 2016; Mortiboys et al., 2008). In our study, *SDHD* was involved in complex II of the mitochondrial electron transport chain, and this gene is responsible for transferring electrons from succinate to ubiquinone (Tables S9 and S11). We hypothesize that *VPS13C* regulates parkin-mediated stimulation of mitophagy in response to mitochondrial depolarization, thereby inhibiting the generation of reactive oxygen species and reducing irreversible mitochondrial damage (Figure 4D) (Youle and van der Bliek, 2012). The higher expression level of *VPS13C* in the C2C12 cell line under heat treatment may also support this hypothesis (Figure 4E). In addition, genotype frequency analysis of 845 accessions showed that 59 mutations (16.95%) within top 1% outlier of absolute *K* value are located within the region of *VPS13C* gene (Figure 4F; Transparent Methods) (Wang et al., 2020).

PSGs Associated with Defense against Zoonotic Diseases

Avian influenza and salmonellosis are caused by pathogens capable of infecting humans and animals and cause significant morbidity and mortality worldwide. Chickens are an important medical model for studying low-pathogenicity avian influenza viruses due to partial immunity owing to previous exposure (Seo et al., 2002). The Saudi Arabian climate, as a high-temperature climate, is conducive to the propagation of pathogenic bacteria in chicken coops, resulting in poultry that are vulnerable to infection. Cytokines are key players in the regulation of the immune response, particularly during infection, joint inflammation, and endocrinological autoimmune diseases (Evans, 1993). Among 942 PSGs of SA breed, we found that 60 were significantly (< 0.05) involved in several functional categories related to infection and defense against zoonotic diseases, including the *influenza A* pathway (14 PSGs), the *Salmonella infection* pathway (11 PSGs), cytokine-cytokine receptor interaction (13 PSGs), immune system development (26 PSGs), and immune response (25 PSGs) (Tables S8 and S9). These functional terms indicate that the thermal climate can indirectly affect autoimmune regulation in animals through its effects on the environments that chickens inhabit.

Of these PSGs, *TLR7* is a key component of innate and adaptive immunity and is involved in RNA virus recognition, especially in the recognition of highly pathogenic avian influenza viruses. Small antiviral compounds activate immune cells via the *TLR7-MyD88*-dependent signaling pathway. *TLR7* may activate *IRF7* through activation of *MyD88*, *BTK*, and *TRAF6*, thus inducing antiviral responses via the production of *IFN-α*. *IFNGR2* interacts with *IFNGR1* to form a receptor for the cytokine IFN-γ (Soh et al., 1994). A previous study reported that loxoribine induces antiviral gene expression, such as the expression of type I IFNs (*IFN-α* and *IFN-β*) and *IFN-γ* in primary chicken splenocytes, and can inhibit influenza A replication *in vitro* and *in vivo* in a dose-dependent manner (Stewart et al., 2012). The PSG *IL12B* is a cytokine that can act as a growth factor for T and NK cell activation, which can enhance the lytic activity of natural killer (NK)/lymphokine-activated killer cells, and stimulate the production of *IFN-γ* by resting peripheral blood mononuclear cells (Oppmann et al., 2000). Another PSG *IL18* is a proinflammatory cytokine primarily involved in polarized T-helper 1 cell and NK cell immune responses (Tominaga et al., 2000). Another study found that members of the Toll-like receptor (TLR) family are critical for the recognition and clearance of *Salmonella* (Talbot et al., 2009; Vazquez-Torres et al., 2004). One consequence of *Salmonella*-induced TLR activation is the production of inflammatory cytokines and antimicrobial compounds, including *pro–IL-1β*, *pro–IL-18*, *IFN-γ*, *TNF-α*, and reactive oxygen species, which are critical mediators for the control of bacterial growth in host tissues (Eckmann and Kagnoff, 2001).

Remarkably, within the *TLR7* gene, we found that two nearly fixed missense mutations in SA breed, i.e., Val121-Ile (G361-A) and Thr135-Ser (A403-T), have significantly higher frequencies than those observed in the other breeds (Figure 5A), possibly due to strong selective pressure. In addition, a similar pattern was identified in the gene *ZC3HAV1*, which has high breed genotype spectrum frequencies and harbors three missense mutations (Figure 5B). *ZC3HAV1* encodes an antiviral protein that inhibits the replication of viruses by recruiting cellular RNA degradation machinery to degrade viral mRNAs (Hayakawa et al., 2011). In addition, our genotype frequency analysis of 845 accessions showed that 12 mutations within top 1% outlier of absolute *K* value are located within the region of these two genes (*TLR7* and *ZC3HAV1*) (Wang et al., 2020) (Figures 5C–5F; Transparent Methods).

Thus, the immune adaptation of the chickens was shaped by the hot environment for resistance to invasion by pathogens. Our findings suggest the importance of specific advantageous alleles for the defensive response to pathogens in chickens.

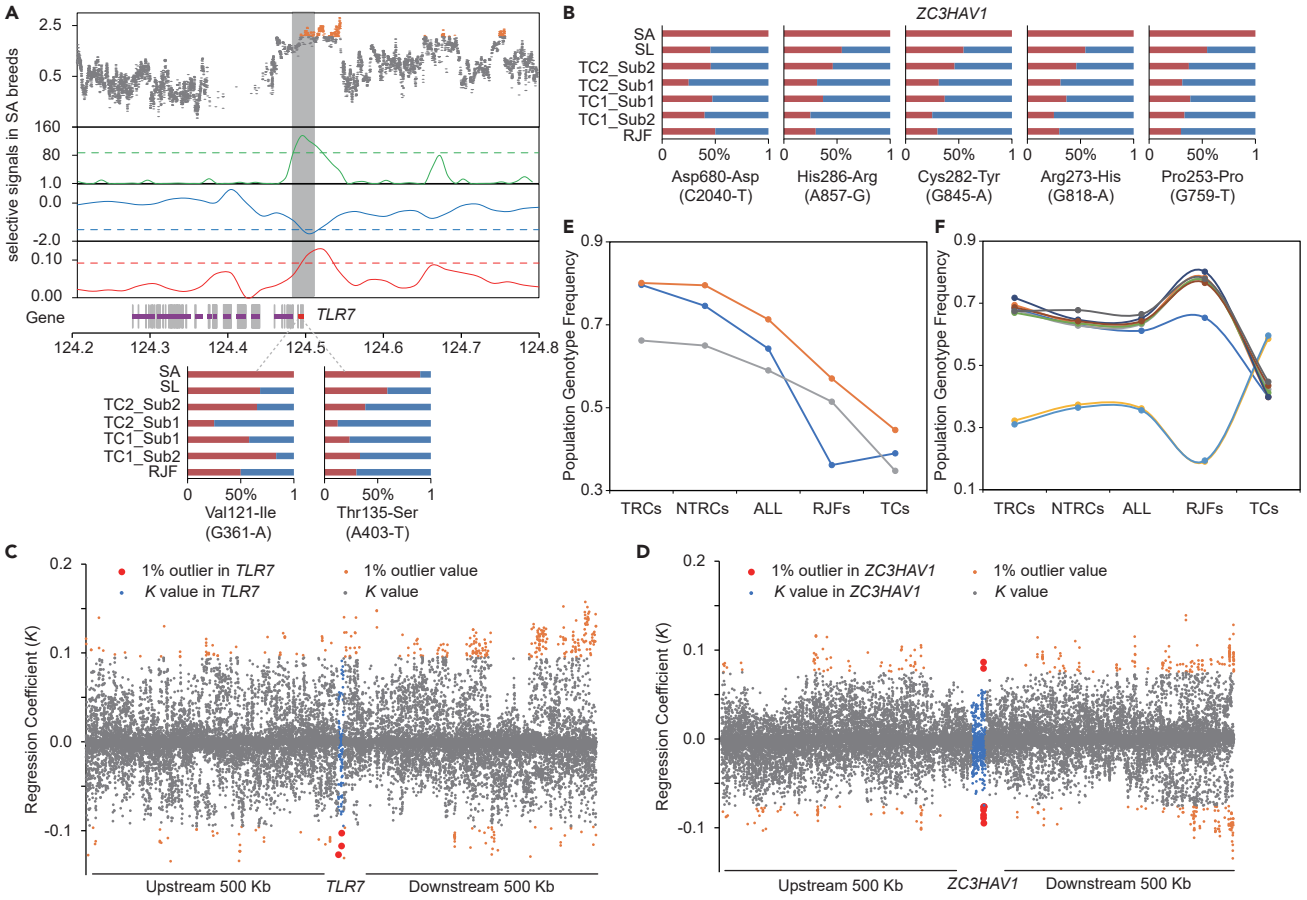


Figure 5. Selective Sweep Signals with two PSGs Associated with Defense Against Zoonotic Diseases

(A) Selective signals in SA breed for the *TLR7* gene. The upper half is similar to Figure 4A. Allele frequencies of variants within the *TLR7* gene across seven chicken subclusters are shown at the bottom.

(B) Allele frequencies of five SNPs within the *ZC3HAV1* gene across seven chicken subclusters.

(C) Genotype frequency analysis detects selection signals upstream and downstream 500 Kb around *TLR7* gene based on 845 chickens.

(D) Genotype frequency analysis detects selection signals upstream and downstream 500 Kb around *ZC3HAV1* gene based on 845 chickens.

(E) Regional plots for three directional mutations across five chicken populations in *TLR7* gene.

(F) Regional plots for nine directional mutations across five chicken populations in *ZC3HAV1* gene.

Convergent Evolution to Tropical Desert Climate between Birds and Mammals

Intriguingly, we identified seven PSGs (*CYP2J19, CYP2J21, CYP2J22, CYP2J23, PLA2G12B, TBXAS1*, and *HPGDS*) that are significantly enriched in the pathway of "Arachidonic acid metabolism" and three PSGs (*OC90, DRD3*, and *PLA2G12B*) related to the GO term "arachidonic acid secretion (GO: 0050482, $p < 0.05$)" specific in SA breed (Tables S8 and S9). The arachidonic acid metabolism pathway plays a key role in converting arachidonic acid to hydroxyeicosatetraenoic acids (HETE) and exhibits strong expression in the kidney (Rouault et al., 2003). Members of the CYP2J family could help convert arachidonic acid into 19(S)-HETE, which has been demonstrated to be a potent vasodilator of renal preglomerular vessels for stimulating water reabsorption (Carroll et al., 1996; Croft et al., 2000; Escalante et al., 1991). In addition, *CYP2J2* is regulated by high-salt diet and its suppression can lead to high blood pressure (Carroll et al., 1996). We also observed specific genotype blocks of the two genes (*OC90* and *PLA2G12B*) in SA breed, which have significantly higher or lower frequencies than those observed in the other breeds (Figures 6A and 6B). Indeed, *PLA2G12B* catalyzes the conversion of arachidonic acid to 20-HETE eicosanoid, which is a potent vasoconstrictor produced in vascular smooth muscle cells that depolarizes and stimulates angiogenic responses *in vivo* (Figure 6C) (Miyata and Roman, 2005). Consistently, we also identified in the SA breed two PSGs (*KIF26B* and *AR*) involved in urogenital system development and one PSG (*CRH*) harboring positive regulation of defecation (Table S9), which may be involved in the response to renal

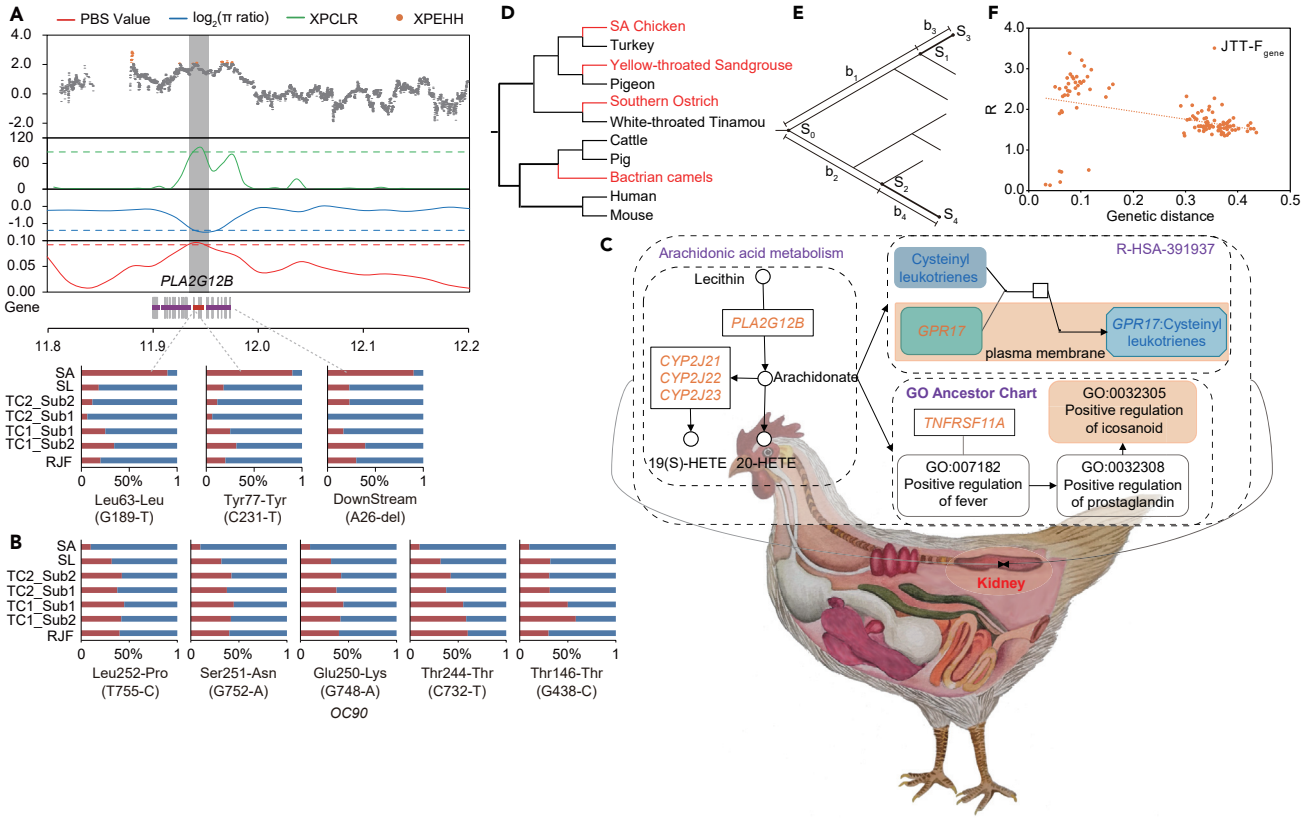


Figure 6. Adaptively Convergent Genes Linked to Arachidonic Acid Metabolism

(A) Selective signals around the *PLA2G12B* locus. Genomic regions located above the upper horizontal red, green, and yellow dashed lines and under the horizontal blue dashed line were termed regions with strong selective signals for SA breed (gray shading). The structure of this gene is shown in the middle of the figure panel (gray bar, coding sequence; purple bar, gene body), and the region of this gene showing selection signals is marked in red. Allele frequencies of variants within the *PLA2G12B* gene across seven chicken subclusters are shown at the bottom.

(B) Allele frequencies of five SNPs within the *OC90* gene across seven chicken subclusters.

(C) Schematic diagram showing the metabolic pathways of arachidonic acid and its derivatives. The left panel represents a part of the arachidonic acid metabolic pathway. The upper right represents the Reactome R-HSA-391937 pathway for derivatives, and the bottom right indicates the GO Ancestor chart of derivatives.

(D) Phylogenetic tree for the 11 species of birds and mammals used for identification of molecular convergence.

(E) A tree model illustrating the counting of the numbers of observed and expected molecular convergences between two thick branches. For a given position, the amino acids at nodes 0–4 are indicated by S0–S4. The relevant branch lengths are indicated by the b values.

(F) Negative correlation between the observed number of molecular convergences relative to the expected number (R) and the genetic distance between the two branches concerned. Each dot represents one branch pair, and different colors show the results under different substitution models. The R values under the JTT-f_{gene} are based on 4,844 proteins. Genetic distance is the number of amino acid substitutions per site between the two younger ends of the two branches considered. Lines show linear regressions.

stress. In conclusion, those genes are important for the adaptation of chickens to tropical desert climates. Additionally, the arachidonic acid metabolism pathway was reported to be important for desert adaptation in the Bactrian camel (Jirimutu et al., 2012) and sheep (Yang et al., 2016), indicating possible convergent evolution between chickens and mammals for inhabiting similar tropical desert climates.

Molecular convergence is viewed as evidence for common adaptations in divergent lineages. We used 11 species, including six birds and five mammals (Figure 6D), to further identify convergent amino acid substitutions for tropical climate adaptation by applying an established method for detecting convergent and parallel evolution (Zhang and Kumar, 1997). Of these species, three also experience desert stress, such as the southern ostrich (*Struthio camelus australis*) (Zhang et al., 2015), which is abundant in dry areas with low rainfall in Africa; the yellow-throated sandgrouse (*Pterocles gutturalis*), which was collected from the Sharjah Breeding Center in the United Arab Emirates (Zhang et al., 2014); and Bactrian camels (Jirimutu et al., 2012), which live in desert regions.

We first replaced the bases of gene sequences from the reference genome with the corresponding fixed genotypes in SA chickens, where the fixed genotype was defined if the allele frequency was more than 0.8. We identified 4,844 single-copy orthologs using OrthoMCL (Li et al., 2003) and then compared the observed number of convergent sites with the random expected number for each gene under the JTT-f_{gene} amino acid substitution models (Figures 6D and 6E) (Zou and Zhang, 2015). As expected, the ratio of molecular convergence between the numbers of observed and expected values declined with the increase in genetic distance between the two branches compared (Figure 6F). Only genes with a significantly higher number of observed convergent sites were considered to be truly under convergent adaptation ($p < 0.05$, Poisson test) (Thomas and Hahn, 2015). Subsequently, we screened convergent genes harboring nonsynonymous amino acids on the basis of PGVs in SA breed. We observed 11, 32, and 37 adaptively convergent genes (ACGs) between SA chickens and Bactrian camels, the southern ostrich, and the yellow-throated sandgrouse, respectively. Gene function enrichment analyses revealed several significant terms involved in the regulation of the response to stimulus and stress (Table S16), especially terms related to "immune system development," "cardiac muscle adaptation," "glomerular development," and "blood vessel development."

Among the 11 ACGs detected between SA chickens and Bactrian camels, *GPR17* has been identified as a dual uracil nucleotide and cysteinyl-leukotriene (CysLT) receptor and is involved in the Reactome pathway named "UDP/CysLT receptor can bind cysteinyl leukotrienes" (R-HSA-391937) (Ciana et al., 2006). CysLTs are derived from the ubiquitous membrane constituent arachidonic acid and are members of a large family of molecules known as eicosanoids. Another convergently evolved gene, *TNFRSF11A*, functions upstream of the biological process "positive regulation of eicosanoid secretion" (GO: 0032305) (Figure 6C). These findings confirmed that arachidonic acid metabolism was an important molecular adaptive convergence mechanism for the tropical desert climate (Table S16). The kidney is a key tissue for regulating water retention and reabsorption. Interestingly, we found several genes associated with kidney size and function, such as *ASXL1*, which are involved in "regulation of kidney size" (GO: 0035564) in the southern ostrich, and *KLF15*, which is involved in "glomerular epithelium development" (GO: 0072010) in the yellow-throated sandgrouse. Additionally, *GPR17* and *TNFRSF11A* were also involved in the biological process "regulation of inflammatory response" (GO:0050727), and similar convergence mechanisms were also found in the southern ostrich and the yellow-throated sandgrouse, which indicated that heat stress could increase inflammatory signaling (Table S16) (Ganesan et al., 2016; Pearce et al., 2013).

Conclusions

In sum, selective sweep analyses of the chicken genomes revealed a variety of important genes, pathways, and GO categories associated with genetic adaptations of chicken to tropical desert/monsoon island climates. Specifically, 12 PSGs (*ADCY1, CACNA1C, CAMK2D, PACRG, PARK2, PRKCH, SDHD, SIRT1, WNT7B, TBXAS1, IL18,* and *VPS13C*) were detected to play roles in chicken adaptations to both tropical desert and tropical monsoon island climates. Our study revealed that nerve development, cutaneous melanin formation, hormonal regulation, angiogenesis, and vasodilation may have worked together to enable chickens to adapt to heat stress. We also discovered different mutations of the same *EPAS1* gene to cope with different environment stresses. Furthermore, a number of PSGs are functionally related to the immune adaptation to the defense against zoonotic diseases and water reabsorption in the tropical desert climate; some of these findings were confirmed by analysis of population genotype frequency in a large dataset of 845 chicken accessions just released. In addition, we provided insights into the genetic mechanism of water reabsorption that have resulted in adaptive convergence between birds and mammals inhabiting tropical desert climate. In view of the current and future global effects of climate change, our study adds to knowledge on genetic mechanisms of adaptations to tropical climates in vertebrates and also has inestimable value for breeding heat-, drought-, and stress-tolerant chicken lines and/or breeds, as well as for medical research related to zoonotic diseases.

Limitations of the Study

To make up for the adverse effect caused by the small sample size, we combined four classic methods to independently identifying genome-wide selection signals for two chicken breeds inhabiting both types of tropical climates; this could reduce the false-positive rate caused by the deviation of a single method. We subsequently revealed genetic adaptations to tropical climates in chickens of two populations inhabiting two similar tropical climates and confirmed some of these results by analysis of the population genotype frequency in a large dataset of 845 chicken accessions just released. Furthermore, we employed comparative genomic analysis between six birds and five mammals, combining with population genomics analysis,

to explain shared genetic mechanisms associated with adaptations to the tropical desert climate from divergent lineages. Due to the protection of local chicken breeds abroad, it is difficult to obtain samples with similar genetic backgrounds. In the future, more samples from diverse breeds are needed to further understand the genetic mechanisms of adaptations to tropical climates in chickens or other vertebrates.

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Huabin Zhao (huabinzhao@whu.edu.cn).

Data Availability

The sequence data of 67 chicken genomes in this study are available under the NCBI accession numbers of SRP067615, PRJNA241474, and PRJNA453469.

Materials Availability

This study did not generate new reagents or other materials.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101644.

ACKNOWLEDGMENTS

The work was supported by the National Natural Science Foundation of China (31722051 and 31672272), Natural Science Foundation of the Hubei Province (2019CFA075), Sichuan Provincial Department of Science & Technology Program (2019JDTD0009, 20GJHZ0069 and 20SYSX0249), and Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (161026).

AUTHOR CONTRIBUTIONS

H.Z., D.L., and S.T. designed the study. S.T. and D.Z. performed the bioinformatics analysis. C.N. performed the qRT-PCR experiment. T.P. and N.Z. made an effort to sample Tibetan chickens. S.T. and H.Z. wrote the manuscript. X.Z. and D.L. revised the paper.

DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

Received: May 26, 2020
Revised: August 19, 2020
Accepted: September 30, 2020
Published: November 20, 2020

REFERENCES

Ahmed, R.G. (2005). Heat stress induced histopathology and pathophysiology of the central nervous system. Int. J. Dev. Neurosci. *23*, 549–557.

Ai, H., Fang, X., Yang, B., Huang, Z., Chen, H., Mao, L., Zhang, F., Zhang, L., Cui, L., He, W., et al. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. Nat. Genet. *47*, 217–225.

Akey, J.M. (2009). Constructing genomic maps of positive selection in humans: where do we go from here? Genome Res. *19*, 711–722.

Andersson, L., and Georges, M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. Nat. Rev. Genet. *5*, 202–212.

Boivin, N., and Fuller, D.Q. (2009). Shell middens, ships and seeds: exploring coastal subsistence, maritime trade and the dispersal of domesticates in and around the ancient Arabian Peninsula. J. World Prehistory *22*, 113–180.

Boulant, J.A., and Dean, J.B. (1986). Temperature receptors in the central nervous system. Annu. Rev. Physiol. *48*, 639–654.

Boulnois, L. (2004). Silk Road: Monks, Warriors & Merchants on the Silk Road (WW Norton & Co Inc)).

Carroll, M.A., Balazy, M., Margiotta, P., Huang, D.D., Falck, J.R., and McGiff, J.C. (1996). Cytochrome P-450-dependent HETEs: profile of biological activity and stimulation by vasoactive peptides. Am. J. Phys. *271*, R863–R869.

Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. Genome Res. *20*, 393–402.

Ciana, P., Fumagalli, M., Trincavelli, M.L., Verderio, C., Rosa, P., Lecca, D., Ferrario, S., Parravicini, C., Capra, V., and Gelosa, P. (2006). The orphan receptor GPR17 identified as a new dual uracil nucleotides/cysteinyl-leukotrienes receptor. EMBO J. *25*, 4615–4627.

Collier, R.J., and Gebremedhin, K.G. (2015). Thermal biology of domestic animals. Annu. Rev. Anim. Biosci. *3*, 513–532.

Croft, K.D., McGiff, J.C., Sanchez-Mendoza, A., and Carroll, M.A. (2000). Angiotensin II releases 20-HETE from rat renal microvessels. Am. J. Physiol. Ren. Physiol. *279*, F544–F551.

Daetwyler, H.D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brondum, R.F., Liao, X., Djari, A., Rodriguez, S.C., Grohs, C., et al. (2014). Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. *46*, 858–865.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. Nature *418*, 700–707.

Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. Mol. Biol. Evol. *28*, 2239–2252.

Eckmann, L., and Kagnoff, M.F. (2001). Cytokines in host defense against Salmonella. Microbes Infect. *3*, 1191–1200.

Escalante, B., Erlij, D., Falck, J.R., and McGiff, J.C. (1991). Effect of cytochrome P450 arachidonate metabolites on ion transport in rabbit kidney loop of Henle. Science *251*, 799–802.

Evans, C.H. (1993). Cytokines: molecular keys to homeostasis, development, and pathophysiology. J. Cell. Biochem. *53*, 277–279.

Frichot, E., Schoville, S.D., Bouchard, G., and François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. Mol. Biol. Evol. *30*, 1687–1699.

Fuller, D.Q., and Boivin, N. (2009). Crops, Cattle and Commensals across the Indian Ocean. Current and Potential Archaeobiological Evidence (Etudes Ocean Indien), pp. 13–46.

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L., and Nielsen, R. (2011). Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. PLoS Genet. *7*, e1002355.

Fumihito, A., Miyake, T., Takada, M., Shingu, R., Endo, T., Gojobori, T., Kondo, N., and Ohno, S. (1996). Monophyletic origin and unique dispersal patterns of domestic fowls. Proc. Natl. Acad. Sci. U S A *93*, 6792–6795.

Ganesan, S., Reynolds, C., Hollinger, K., Pearce, S.C., Gabler, N.K., Baumgard, L.H., Rhoads, R.P., and Selsby, J.T. (2016). Twelve hours of heat stress induces inflammatory signaling in porcine skeletal muscle. Am. J. Physiol. Regul. Integr. Comp. Physiol. *310*, R1288–R1296.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H., et al. (2010). A draft sequence of the Neandertal genome. Science *328*, 710–722.

Grimm, M., and Brown, J.H. (2010). β-Adrenergic receptor signaling in the heart: role of CaMKII. J. Mol. Cell. Cardiol. *48*, 322–330.

Hayakawa, S., Shiratori, S., Yamato, H., Kameyama, T., Kitatsuji, C., Kashigi, F., Goto, S., Kameoka, S., Fujikura, D., and Yamada, T. (2011). ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. Nat. Immunol. *12*, 37.

Hillier, L.W., Miller, W., Birney, E., Warren, W., Hardison, R.C., Ponting, C.P., Bork, P., Burt, D.W., Groenen, M.A., and Delany, M.E. (2014). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature *423*, 695–777.

Hu, C.J., Wang, L.Y., Chodosh, L.A., Keith, B., and Simon, M.C. (2003). Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation. Mol. Cell Biol. *23*, 9361–9374.

Huerta-Sanchez, E., Degiorgio, M., Pagani, L., Tarekegn, A., Ekong, R., Antao, T., Cardona, A., Montgomery, H.E., Cavalleri, G.L., Robbins, P.A., et al. (2013). Genetic signatures reveal high-altitude adaptation in a set of ethiopian populations. Mol. Biol. Evol. *30*, 1877–1888.

Jiao, H., Zhang, L., Xie, H., Simmons, N., Liu, H., and Zhao, H. (2019). Trehalase gene as a molecular signature of dietary diversification in mammals. Mol. Biol. Evol. *36*, 2171–2183.

Jirimutu, Wang, Z., Ding, G., Chen, G., Sun, Y., Sun, Z., Zhang, H., Wang, L., Hasi, S., Zhang, Y., et al. (2012). Genome sequences of wild and domestic bactrian camels. Nat. Commun. *3*, 1202.

Johnsson, M., Gustafson, I., Rubin, C.-J., Sahlqvist, A.-S., Jonsson, K.B., Kerje, S., Ekwall, O., Kämpe, O., Andersson, L., and Jensen, P. (2012). A sexual ornament in chickens is affected by pleiotropic alleles at HAO1 and BMP2, selected during domestication. PLoS Genet. *8*, e1002914.

Kim, E.S., Elbeltagy, A.R., Aboul-Naga, A.M., Rischkowsky, B., Sayre, B., Mwacharo, J.M., and Rothschild, M.F. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. Heredity (Edinb) *116*, 255–264.

Lara, L.J., and Rostagno, M.H. (2013). Impact of heat stress on poultry production. Animals (Basel) *3*, 356–369.

Larson, G., and Fuller, D.Q. (2014). The evolution of animal domestication. Annu. Rev. Ecol. Evol. Syst. *45*, 115–136.

Lawal, R.A., Al-Atiyat, R.M., Aljumaah, R.S., Silva, P., Mwacharo, J.M., and Hanotte, O. (2018). Whole-genome resequencing of red junglefowl and indigenous village chicken reveal new insights on the genome dynamics of the species. Front. Genet. *9*, 264.

Lesage, S., Drouet, V., Majounie, E., Deramecourt, V., Jacoupy, M., Nicolas, A., Cormier-Dequaire, F., Hassoun, S.M., Pujol, C., Ciura, S., et al. (2016). Loss of VPS13C function in autosomal-recessive parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent mitophagy. Am. J. Hum. Genet. *98*, 500–513.

Li, D., Che, T., Chen, B., Tian, S., Zhou, X., Zhang, G., Li, M., Gaur, U., Li, Y., Luo, M., et al. (2017a). Genomic data for 78 chickens from 14 populations. Gigascience *6*, 1–5.

Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. *13*, 2178–2189.

Li, M., Tian, S., Jin, L., Zhou, G., Li, Y., Zhang, Y., Wang, T., Yeung, C.K., Chen, L., Ma, J., et al. (2013). Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nat. Genet. *45*, 1431–1438.

Li, S., Li, D., Zhao, X., Wang, Y., Yin, H., Zhou, L., Zhong, C., and Zhu, Q. (2017b). A non-synonymous SNP with the allele frequency correlated with the altitude may contribute to the hypoxia adaptation of Tibetan chicken. PLoS One *12*, e0172211.

Liu, Y.P., Wu, G.S., Yao, Y.G., Miao, Y.W., Luikart, G., Baig, M., Beja-Pereira, A., Ding, Z.L., Palanichamy, M.G., and Zhang, Y.P. (2006). Multiple maternal origins of chickens: out of the Asian jungles. Mol. Phylogenet. Evol. *38*, 12–19.

Lv, F.H., Agha, S., Kantanen, J., Colli, L., Stucki, S., Kijas, J.W., Joost, S., Li, M.H., and Ajmone Marsan, P. (2014). Adaptations to climate-mediated selective pressures in sheep. Mol. Biol. Evol. *31*, 3324–3343.

Miyata, N., and Roman, R.J. (2005). Role of 20-hydroxyeicosatetraenoic acid (20-HETE) in vascular system. J. Smooth Muscle Res. *41*, 175–193.

Mortiboys, H., Thomas, K.J., Koopman, W.J., Klaffke, S., Abou-Sleiman, P., Olpin, S., Wood, N.W., Willems, P.H., Smeitink, J.A., Cookson, M.R., et al. (2008). Mitochondrial function and morphology are impaired in parkin-mutant fibroblasts. Ann. Neurol. *64*, 555–565.

Mutaf, S., Kahraman, N.S., and Firat, M.Z. (2009). Intermittent partial surface wetting and its effect on body-surface temperatures and egg production of white and brown domestic laying hens in Antalya (Turkey). Br. Poult. Sci. *50*, 33–38.

Nei, M., and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. U S A *76*, 5269–5273.

Oppmann, B., Lesley, R., Blom, B., Timans, J.C., Xu, Y., Hunte, B., Vega, F., Yu, N., Wang, J., and Singh, K. (2000). Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. Immunity *13*, 715–725.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

Pearce, S.C., Mani, V., Boddicker, R.L., Johnson, J.S., Weber, T.E., Ross, J.W., Rhoads, R.P., Baumgard, L.H., and Gabler, N.K. (2013). Heat stress reduces intestinal barrier integrity and

favors intestinal glucose transport in growing pigs. PLoS One *8*, e70215.

Percy, M.J., Furlow, P.W., Lucas, G.S., Li, X., Lappin, T.R., McMullin, M.F., and Lee, F.S. (2008). A gain-of-function mutation in the HIF2A gene in familial erythrocytosis. N. Engl. J. Med. *358*, 162–168.

Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. *8*, e1002967.

Portner, H.O. (2001). Climate change and temperature-dependent biogeography: oxygen limitation of thermal tolerance in animals. Naturwissenschaften *88*, 137–146.

Qanbari, S., Rubin, C.J., Maqbool, K., Weigend, S., Weigend, A., Geibel, J., Kerje, S., Wurmser, C., Peterson, A.T., Brisbin, I.L., Jr., et al. (2019). Genetics of adaptation in modern chicken. PLoS Genet. *15*, e1007989.

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., and Larkin, D.M. (2012). The yak genome and adaptation to life at high altitude. Nat. Genet. *44*, 946–949.

Qu, Y.H., Zhao, H.W., Han, N.J., Zhou, G.Y., Song, G., Gao, B., Tian, S.L., Zhang, J.B., Zhang, R.Y., Meng, X.H., et al. (2013). Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. Nat. Commun. *4*, 2071.

Rouault, M., Bollinger, J.G., Lazdunski, M., Gelb, M.H., and Lambeau, G. (2003). Novel mammalian group XII secreted phospholipase A2 lacking enzymatic activity. Biochemistry *42*, 11494–11503.

Rubin, C.J., Zody, M.C., Eriksson, J., Meadows, J.R., Sherwood, E., Webster, M.T., Jiang, L., Ingman, M., Sharpe, T., Ka, S., et al. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. Nature *464*, 587–591.

Sabeti, P.C., Schaffner, S.F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T.S., Altshuler, D., and Lander, E.S. (2006). Positive natural selection in the human lineage. Science *312*, 1614–1620.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. Nature *449*, 913–918.

Satzger, I., Schaefer, T., Kuettler, U., Broecker, V., Voelker, B., Ostertag, H., Kapp, A., and Gutzmer, R. (2008). Analysis of c-KIT expression and KIT gene mutation in human mucosal melanomas. Br. J. Cancer *99*, 2065–2069.

Schaub, M.C., Hefti, M.A., and Zaugg, M. (2006). Integration of calcium with the signaling network in cardiac myocytes. J. Mol. Cell Cardiol. *41*, 183–214.

Seo, S.H., Peiris, M., and Webster, R.G. (2002). Protective cross-reactive cellular immunity to lethal A/Goose/Guangdong/1/96-like H5N1 influenza virus is correlated with the proportion of pulmonary CD8+ T cells expressing gamma interferon. J. Virol. *76*, 4886–4890.

Simonson, T.S., Yang, Y., Huff, C.D., Yun, H., Qin, G., Witherspoon, D.J., Bai, Z., Lorenzo, F.R., Xing, J., Jorde, L.B., et al. (2010). Genetic evidence for high-altitude adaptation in Tibet. Science *329*, 72–75.

Soh, J., Donnelly, R.J., Kotenko, S., Mariano, T.M., Cook, J.R., Wang, N., Emanuel, S., Schwartz, B., Miki, T., and Pestka, S. (1994). Identification and sequence of an accessory factor required for activation of the human interferon gamma receptor. Cell *76*, 793–802.

Stewart, C.R., Bagnaud-Baule, A., Karpala, A.J., Lowther, S., Mohr, P.G., Wise, T.G., Lowenthal, J.W., and Bean, A.G. (2012). Toll-like receptor 7 ligands inhibit influenza A infection in chickens. J. Interferon Cytokine Res. *32*, 46–51.

Talbot, S., Tötemeyer, S., Yamamoto, M., Akira, S., Hughes, K., Gray, D., Barr, T., Mastroeni, P., Maskell, D.J., and Bryant, C.E. (2009). Toll-like receptor 4 signalling through MyD88 is essential to control Salmonella enterica serovar Typhimurium infection, but not for the initiation of bacterial clearance. Immunology *128*, 472–483.

Tang, H., Peng, J., Wang, P., and Risch, N.J. (2005). Estimation of individual admixture: analytical and study design considerations. Genet. Epidemiol. *28*, 289–301.

Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S., and Sahay, A. (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature *503*, 218–223.

Thomas, G.W., and Hahn, M.W. (2015). Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. Mol. Biol. Evol. *32*, 1232–1236.

Tominaga, K., Yoshimoto, T., Torigoe, K., Kurimoto, M., Matsui, K., Hada, T., Okamura, H., and Nakanishi, K. (2000). IL-12 synergizes with IL-18 or IL-1beta for IFN-gamma production from human T cells. Int. Immunol. *12*, 151–160.

Vazquez-Torres, A., Vallance, B.A., Bergman, M.A., Finlay, B.B., Cookson, B.T., Jones-Carson, J., and Fang, F.C. (2004). Toll-like receptor 4 dependence of innate and adaptive immunity to Salmonella: importance of the Kupffer cell network. J. Immunol. *172*, 6202–6208.

Wang, M.S., Li, Y., Peng, M.S., Zhong, L., Wang, Z.J., Li, Q.Y., Tu, X.L., Dong, Y., Zhu, C.L., Wang, L., et al. (2015). Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. Mol. Biol. Evol. *32*, 1880–1889.

Wang, M.S., Thakur, M., Peng, M.S., Jiang, Y., Frantz, L.A.F., Li, M., Zhang, J.J., Wang, S., Peters, J., Otecko, N.O., et al. (2020). 863 genomes reveal the origin and domestication of chicken. Cell Res. 693–701.

Wang, M.S., Zhang, R.W., Su, L.Y., Li, Y., Peng, M.S., Liu, H.Q., Zeng, L., Irwin, D.M., Du, J.L., Yao, Y.G., et al. (2016). Positive selection rather than relaxation of functional constraint drives the evolution of vision during chicken domestication. Cell Res. *26*, 556–573.

Wong, G.K., Liu, B., Wang, J., Zhang, Y., Yang, X., Zhang, Z., Meng, Q., Zhou, J., Li, D., Zhang, J., et al. (2004). A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature *432*, 717–722.

Wu, H., Guang, X., Al-Fageeh, M.B., Cao, J., Pan, S., Zhou, H., Zhang, L., Abutarboush, M.H., Xing, Y., Xie, Z., et al. (2014). Camelid genomes reveal evolution and adaptation to desert environments. Nat. Commun. *5*, 5188.

Yang, J., Li, W.R., Lv, F.H., He, S.G., Tian, S.L., Peng, W.F., Sun, Y.W., Zhao, Y.X., Tu, X.L., Zhang, M., et al. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. Mol. Biol. Evol. *33*, 2576–2592.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., et al. (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. Science *329*, 75–78.

Youle, R.J., and van der Bliek, A.M. (2012). Mitochondrial fission, fusion, and stress. Science *337*, 1062–1065.

Zhan, S., Zhang, W., Niitepold, K., Hsu, J., Haeger, J.F., Zalucki, M.P., Altizer, S., de Roode, J.C., Reppert, S.M., and Kronforst, M.R. (2014). The genetics of monarch butterfly migration and warning colouration. Nature *514*, 317–321.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., Storz, J.F., Antunes, A., Greenwold, M.J., and Meredith, R.W. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. Science *346*, 1311–1320.

Zhang, J., Li, C., Zhou, Q., and Zhang, G. (2015). Improving the ostrich genome assembly using optical mapping data. GigaScience *4*, 24.

Zhang, J.Z., and Kumar, S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. Mol. Biol. Evol. *14*, 527–536.

Zhang, Q., Gou, W., Wang, X., Zhang, Y., Ma, J., Zhang, H., Zhang, Y., and Zhang, H. (2016). Genome resequencing identifies unique adaptations of Tibetan chickens to hypoxia and high-dose ultraviolet radiation in high-altitude environments. Genome Biol. Evol. *8*, 765–776.

Zhu, X.J., Liu, Y., Dai, Z.M., Zhang, X., Yang, X., Li, Y., Qiu, M., Fu, J., Hsu, W., Chen, Y., et al. (2014). BMP-FGF signaling axis mediates Wnt-induced epidermal stratification in developing mammalian skin. PLoS Genet. *10*, e1004687.

Zou, Z.T., and Zhang, J.Z. (2015). Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? Mol. Biol. Evol. *32*, 2085–2096.

iScience, Volume *23*

Supplemental Information

Genomic Analyses Reveal Genetic Adaptations to Tropical Climates in Chickens

Shilin Tian, Xuming Zhou, Tashi Phuntsok, Ning Zhao, Dejing Zhang, Chunyou Ning, Diyan Li, and Huabin Zhao

Supplemental Information for

Genomic analyses reveal genetic adaptations to tropical climates in chickens

Shilin Tian[1,6], Xuming Zhou[3,6], Tashi Phuntsok[4], Ning Zhao[4], Dejing Zhang[5], Chunyou Ning[2], Diyan Li[*, 2] and Huabin Zhao[*, 1,4,7]

[1]Department of Ecology, Tibetan Centre for Ecology and Conservation at WHU-TU, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan 430072, China

[2]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu 611130, China

[3]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Beijing 100101, China

[4]Laboratory of Extreme Environmental Biological Resources and Adaptive Evolution, Research Center for Ecology, College of Science, Tibet University, Lhasa 850000, China

[5]Novogene Bioinformatics Institute, Beijing 100015, China.

[6]These authors contributed equally

[7]Lead Contact

*Correspondence to

Huabin Zhao, E-mail: huabinzhao@whu.edu.cn;

Diyan Li, E-mail: diyanli@sicau.edu.cn

This supplementary file contains:

Figures S1-S11

Tables S1-S16 (Tables S9, S11 and S16 are provided as separate Excel files)

Transparent Methods

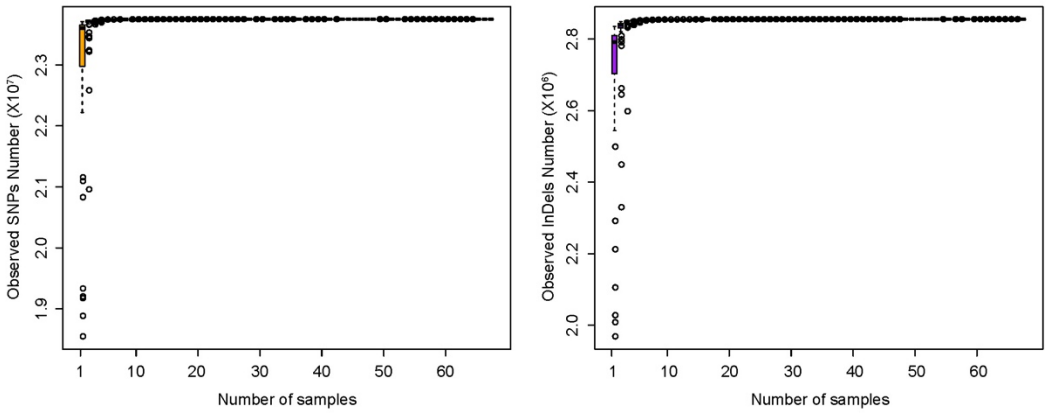


Figure. S1. Nonredundant SNPs (left) and InDels (right) increase as individuals are added. Related to Table 1. The number of genomic variants was plotted on the y axis, whereas the number of individuals sequentially added was shown on the x axis.

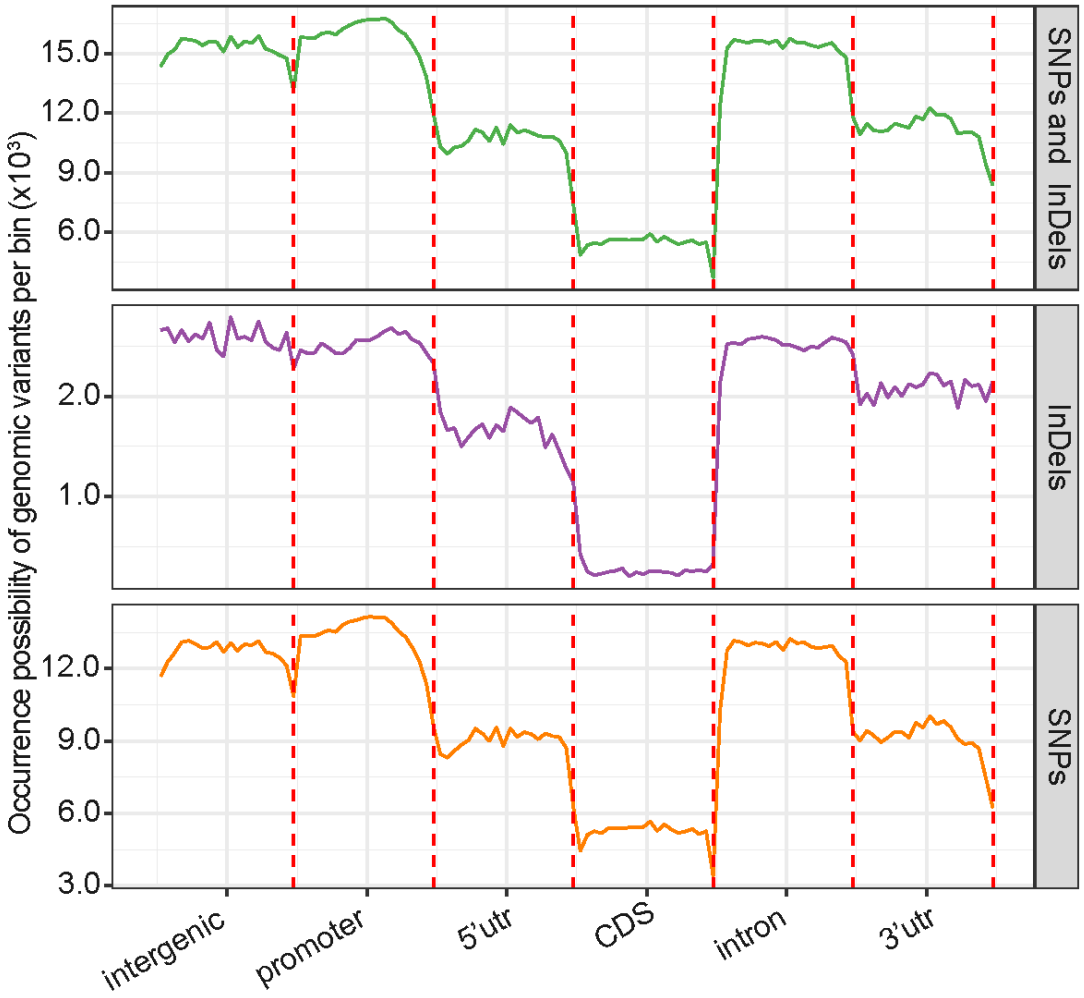


Figure. S2. Occurrence possibility of SNPs and InDels in different genomic elements. Related to Table 1.

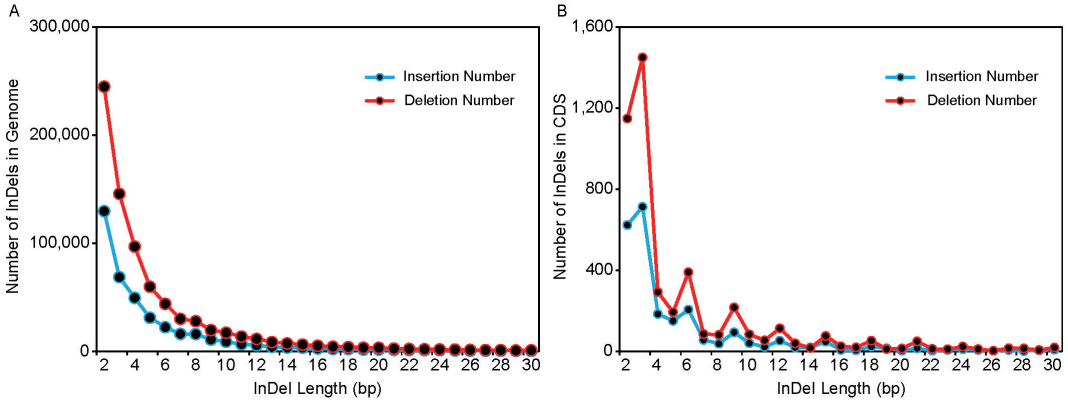


Figure. S3. Length distribution of small indels (1-30bp) in the whole genome (a) and coding sequence (CDS) regions (b). Related to Table 1.

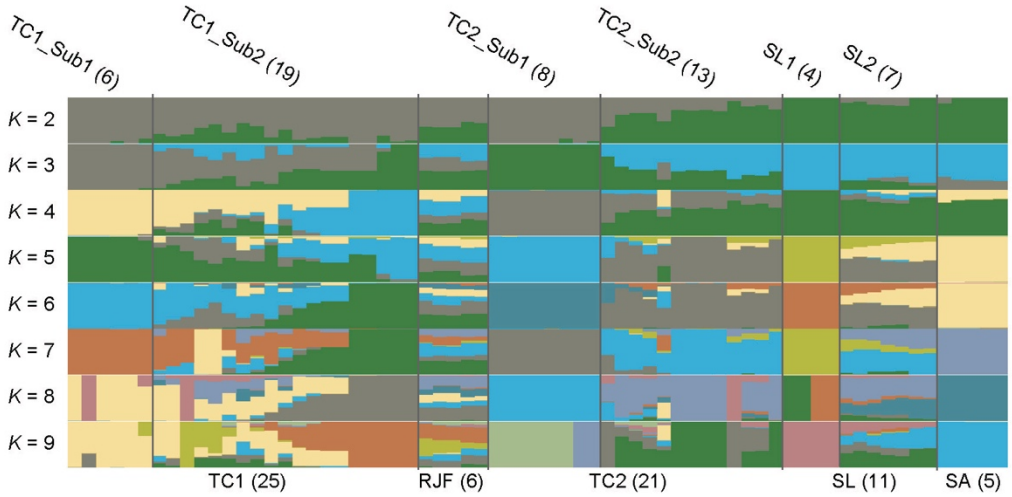


Figure. S4. Genetic structure of chickens. Related to Figure 1. Each colored segment represents the ancestral component in an individual from K = 2 to K = 9. The names of each breed/cluster are shown at the bottom of the figure, and the subcluster names are provided on the top.

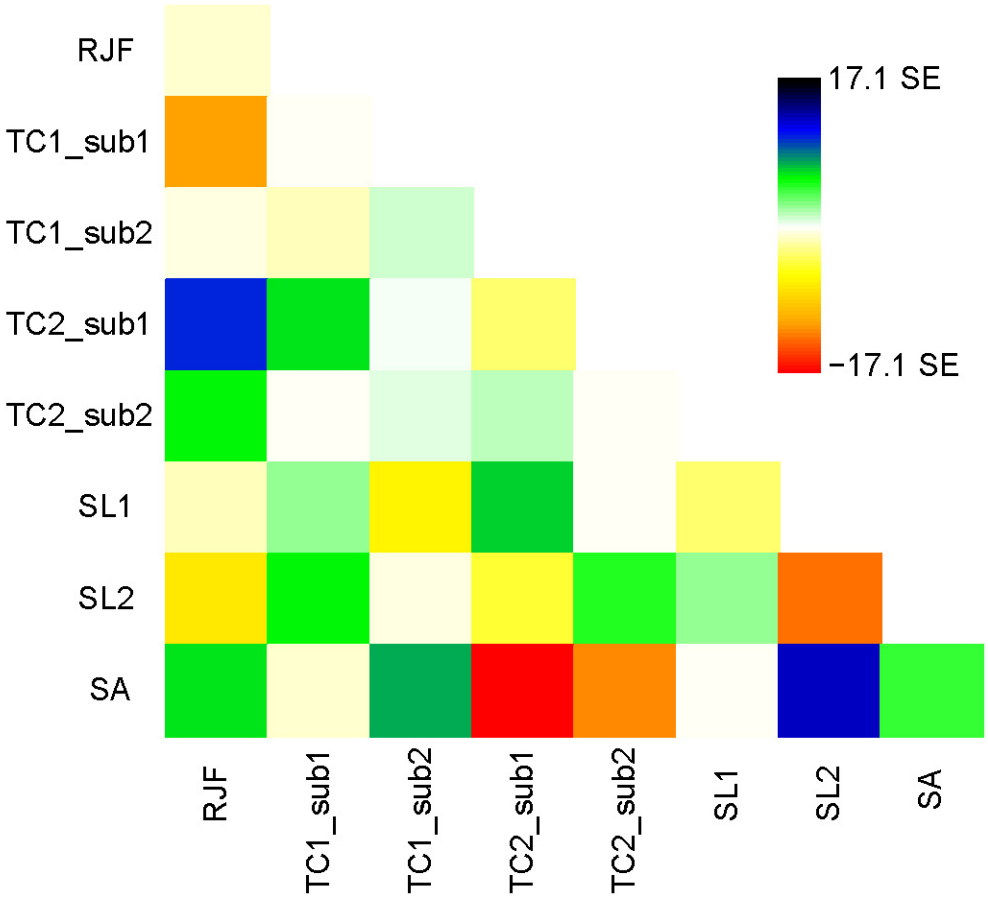


Figure. S5. **Residual fit. Related to Figure 1**. Plotted is the residual fit from the maximum likelihood tree in Figure. 1D. Residual values greater than zero represent populations that are more closely related to each other in the data than in the best-fit tree, which are thus considered candidates for admixture events.

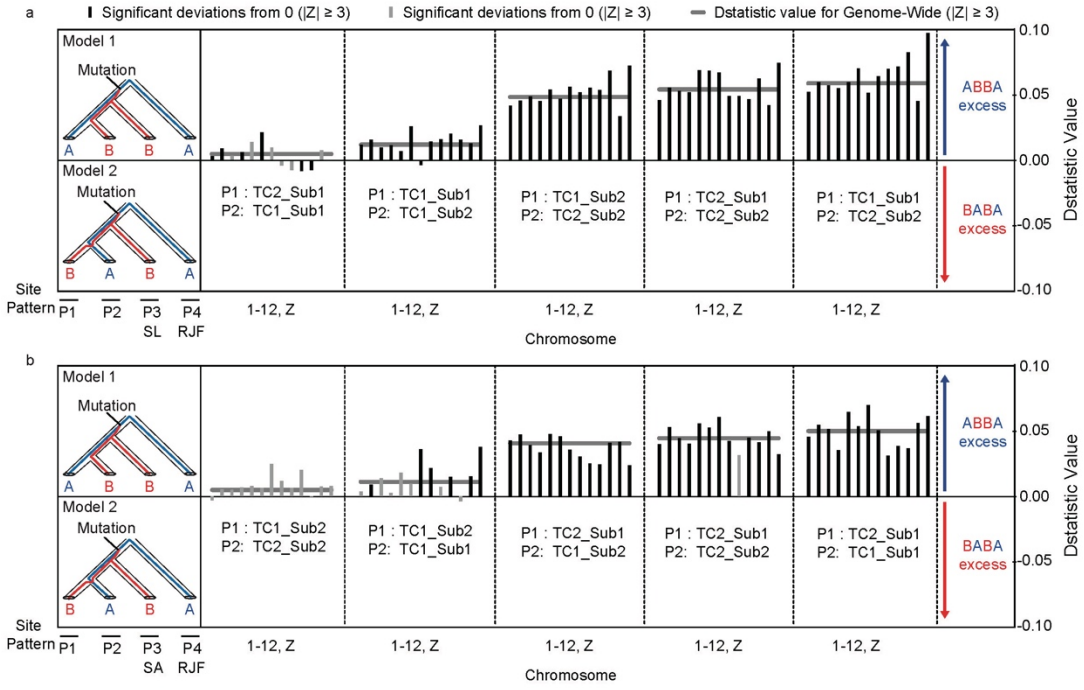


Figure. S6. Four-taxon ABBA/BABA test of introgression. Related to Figure 1.

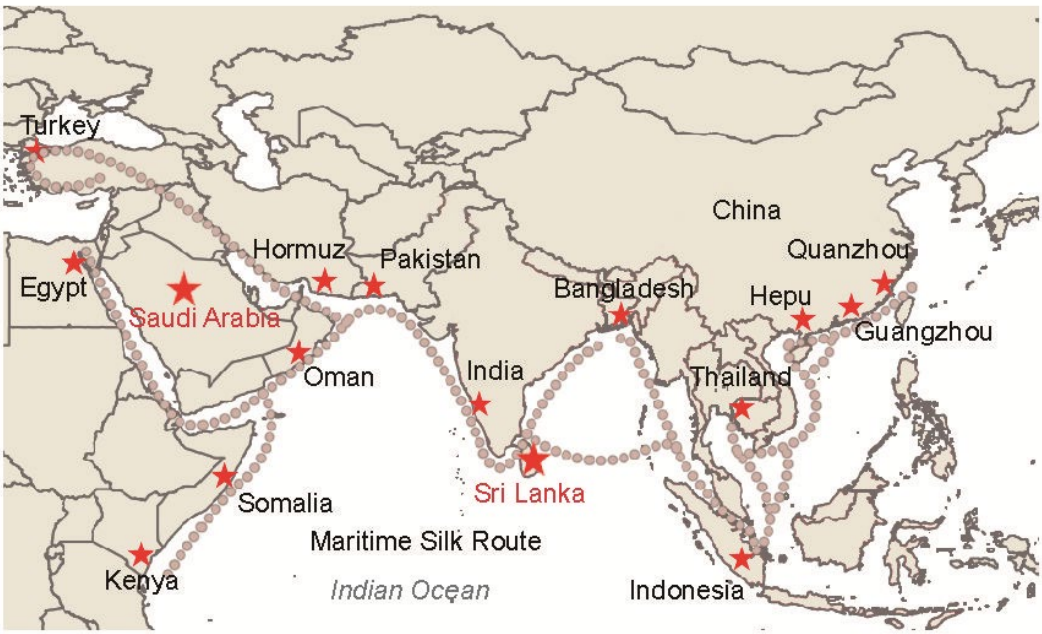


Figure. S7. The Maritime Silk Route. Related to Figure 1.

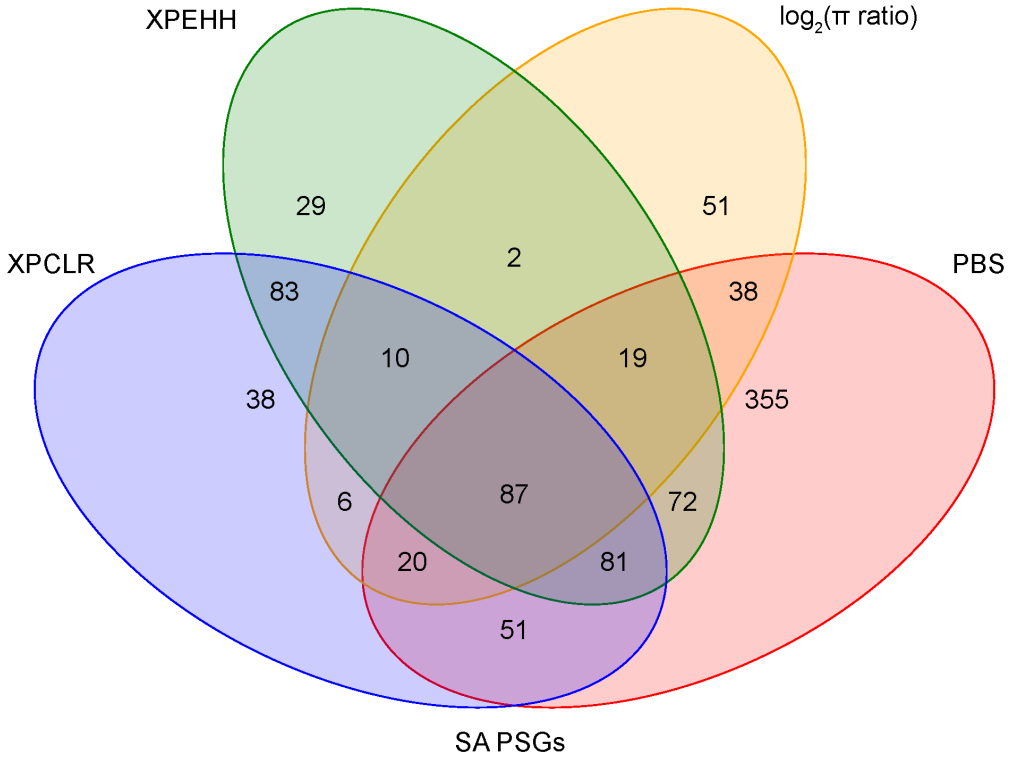


Figure. S8. Venn map of PSGs for hot adaptation in the SA breed detected by the four methods. Related to Figure 2.

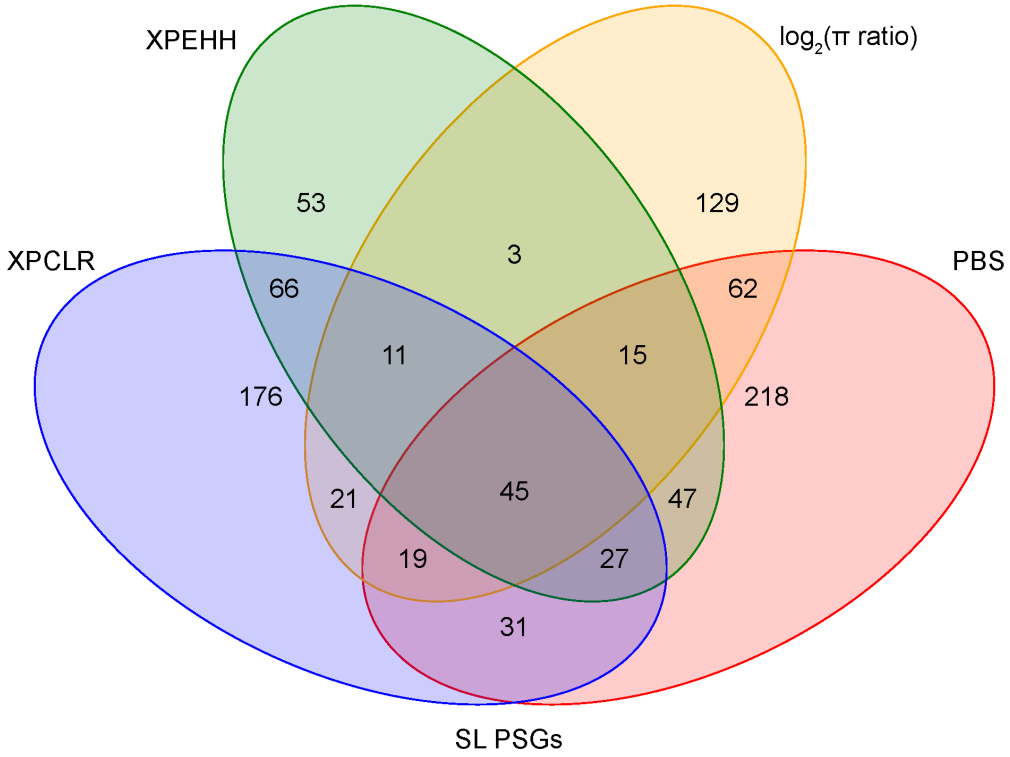


Figure. S9. Venn map of PSGs for hot adaptation in the SL breed detected by the four methods. Related to Figure 2.

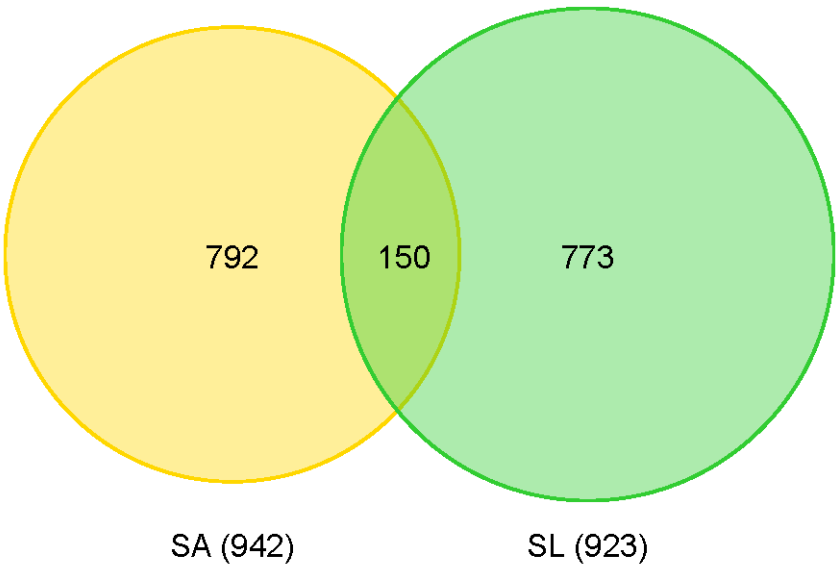


Figure. S10. Venn map of PSGs for hot adaptation in SA and SL breeds compared with TC2_Sub2. Related to Figure 2.

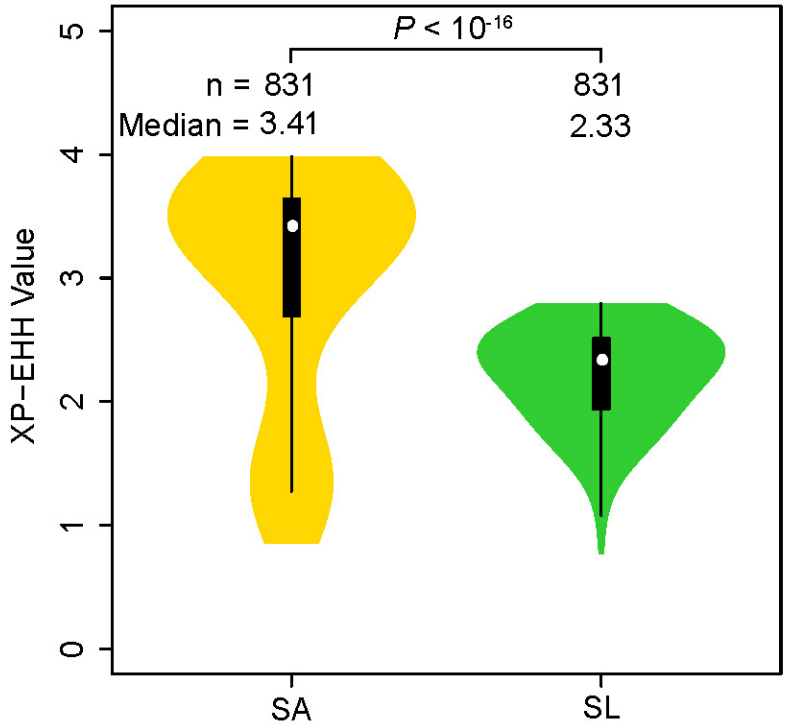


Figure. S11. XP-EHH values for the *SIRT1* gene between SA and SL breeds. Related to Figure 2.

Table S1. Sample information for chickens in this study. Related to Table 1 and Figure 1.

Sample ID	NCBI BioSample/RUN ID	Sampling location	Average annual rainfall	Average annual temperature	Altitude
SL1	SAMN08979083				
SL2	SAMN08979089				
SL3	SAMN08979090				
SL4	SAMN08979091				
SL5	SAMN08979092				
SL6	SAMN08979093	Sri Lankan	~1,000mm	27℃	60m
SL7	SAMN08979084				
SL8	SAMN08979085				
SL9	SAMN08979086				
SL10	SAMN08979087				
SL11	SAMN08979088				
SA1	SAMN08979081				
SA2	SAMN08979082		74mm	26℃ (ranging from 21.2 to 50.8℃)	
SA3	SAMN08979078	Saudi Arabian			110-130m
SA4	SAMN08979079				
SA5	SAMN08979080				
TC3	SAMN02712048				
TC4	SAMN02712049		50-5000m decreases from east to west	8℃ (ranging from -16 to 16℃)	
TC5	SAMN02712050	Tibet			3,000-6,000m
TC6	SAMN02712051				
TC7	SAMN02712052	Shigatse, Tibet	200-430mm	6.3 - 14℃	4000m+
TC1	SAMN02712045				
TC2	SAMN02712047				
TC8	SAMN02712053				
TC9	SAMN02712054				
TC10	SAMN02712046	Nyingchi, Tibet	650mm	8.7℃	3100m
TCLZ1	SRR3041450				
TCLZ2	SRR3041451				
TCLZ3	SRR3041452				
TCLZ4	SRR3041453				
TCLZ5	SRR3041454				
TCAB1	SRR3041433				
TCAB3	SRR3041434				
TCAB5	SRR3041435	Aba, Sichuan	705mm	9.3℃	3300m
TCAB6	SRR3041436				
TCAB7	SRR3041437				

TCDQ1	SRR3041438	Diqing, Yunnan	2261.2mm	4.7 - 16.5℃	3280m
TCDQ2	SRR3041439				
TCDQ3	SRR3041440				
TCDQ4	SRR3041441				
TCDQ5	SRR3041442				
TCDQ6	SRR3041443				
TCGZ1	SRR3041444	Ganzi, Sichuan	325-920mm	8℃	3390m
TCGZ10	SRR3041445				
TCGZ3	SRR3041446				
TCGZ4	SRR3041447				
TCGZ5	SRR3041448				
TCGZ6	SRR3041449				
TCQH1	SRR3041455	Haiyan, Qinghai	400mm	1.5℃	3260m
TCQH10	SRR3041456				
TCQH11	SRR3041457				
TCQH5	SRR3041458				
TCQH8	SRR3041504				
TCQH9	SRR3041573				
TCSN1	SRR3041620	Shannan, Tibet	450mm	6℃	3700m
TCSN3	SRR3041692				
TCSN4	SRR3041713				
TCSN5	SRR3041781				
TCSN6	SRR3041923				
TCSN7	SRR3041924				
TCSN8	SRR3041925				
TCSN9	SRR3041926				
RJF1	SAMN02712039	Yunnan	600-2,300mm	8 - 27℃	~3,000m
RJF2	SAMN02712040				
RJF3	SAMN02712041				
RJF4	SAMN02712042	Hainan	1639mm	22 - 27℃	~1,840m
RJF5	SAMN02712043	Yunnan, Xishuangbanna	1,136-1,513mm	18.9 - 22.6℃	~3,000m

Table S2. Summary of sequence quality control and alignment. Related to Table 1 and Figure 1.

Sample ID	Raw bases (Gb)	High-quality ratio (%)	Q20 (%)	Q30 (%)	Mapping rate (%)	Average sequencing depth (X)	Coverage
SL1	49.81	98.46	93.50	87.87	99.22	44.04	97.59
SL2	50.38	98.29	92.81	86.71	99.14	44.22	97.62
SL3	48.52	98.43	92.75	86.52	99.25	43.25	97.68
SL4	50.39	98.26	92.70	86.51	99.09	44.28	97.67
SL5	34.94	98.28	94.62	89.82	99.30	31.09	98.13
SL6	49.92	98.27	94.33	89.25	99.32	44.09	98.28
SL7	51.62	98.11	94.25	89.14	99.25	45.55	98.37
SL8	49.26	98.38	94.04	88.75	99.33	43.70	97.71
SL9	44.97	98.20	93.48	87.85	99.22	39.88	98.26
SL10	46.69	98.37	93.42	87.73	99.28	41.69	97.69
SL11	46.99	98.54	93.37	87.71	99.28	41.65	97.64
SA1	16.24	95.46	95.31	88.26	99.48	13.26	97.34
SA2	16.29	93.41	94.59	85.37	99.24	9.90	96.93
SA3	16.21	95.02	94.50	87.09	99.57	13.09	97.23
SA4	16.23	94.98	95.35	88.43	99.49	13.17	97.21
SA5	16.20	92.80	94.20	84.69	99.19	9.64	97.06
TC1	22.28	94.32	97.50	92.23	97.92	18.73	94.82
TC2	6.62	87.99	94.21	86.22	99.11	5.34	79.81
TC3	18.77	82.69	93.77	83.35	99.62	14.48	97.61
TC4	14.95	81.37	93.42	82.64	99.59	11.34	97.37
TC5	15.15	79.82	93.22	82.34	99.58	11.26	97.38
TC6	18.66	81.23	93.52	82.85	99.56	14.08	98.14
TC7	38.13	87.98	94.99	87.07	98.14	29.61	98.32
TC8	6.73	89.43	94.59	87.09	99.26	5.52	94.71
TC9	6.28	86.81	93.97	85.77	99.07	4.99	82.13
TC10	9.23	94.87	96.26	90.75	99.50	7.90	97.58
TCAB1	10.01	91.67	95.48	89.95	99.04	7.95	92.09
TCAB3	27.74	94.88	96.63	92.26	97.41	20.28	97.69
TCAB5	10.09	92.39	95.50	89.96	99.14	8.06	86.94
TCAB6	62.09	96.57	96.39	92.92	99.55	53.00	98.00
TCAB7	17.74	94.82	95.17	88.37	99.59	14.62	97.56
TCDQ1	25.99	96.64	96.31	92.80	99.37	21.83	97.75
TCDQ2	10.06	92.05	95.33	89.65	98.95	7.89	74.54

TCDQ3	26.42	96.20	96.20	92.58	99.50	23.01	97.82
TCDQ4	40.20	87.05	96.17	91.50	99.05	28.75	98.31
TCDQ5	14.31	93.26	95.93	90.84	99.10	11.26	92.51
TCDQ6	45.39	91.07	94.94	88.88	98.77	33.02	98.70
TCGZ1	10.44	92.00	95.50	90.00	99.00	8.26	86.81
TCGZ10	10.39	92.52	96.07	91.20	98.98	8.06	85.33
TCGZ3	22.17	95.64	95.98	92.20	99.12	19.38	97.67
TCGZ4	15.90	92.89	95.96	90.90	97.57	12.14	82.37
TCGZ5	34.58	86.39	96.50	92.09	98.81	24.12	98.20
TCGZ6	44.39	92.31	95.06	89.09	98.91	33.90	98.04
TCLZ1	11.24	84.48	94.99	89.20	98.27	8.03	81.04
TCLZ2	12.68	90.52	95.29	89.66	97.84	9.46	76.44
TCLZ3	10.22	87.47	93.84	86.89	97.81	7.50	93.90
TCLZ4	23.33	93.68	96.35	91.71	97.07	16.84	78.40
TCLZ5	21.37	91.22	94.93	88.88	99.02	16.44	96.97
TCQH1	9.15	93.82	95.54	90.65	99.30	7.53	95.76
TCQH10	19.06	95.43	96.01	90.19	99.55	15.47	97.63
TCQH11	19.06	95.58	96.10	90.35	99.54	15.60	97.63
TCQH5	13.07	77.23	95.07	89.51	97.99	8.40	89.50
TCQH8	18.95	95.87	96.23	90.71	99.50	15.50	97.67
TCQH9	18.66	96.03	96.29	90.83	99.50	15.34	97.64
TCSN1	27.98	96.13	96.16	92.52	99.39	24.22	98.01
TCSN3	32.23	86.34	96.53	92.16	99.27	23.08	98.20
TCSN4	23.99	96.39	96.33	92.83	99.48	21.05	97.76
TCSN5	24.60	87.73	95.85	90.76	98.92	17.87	94.28
TCSN6	13.12	91.95	95.85	90.76	98.95	10.24	96.77
TCSN7	14.76	94.84	96.69	91.67	99.66	11.74	97.28
TCSN8	16.91	95.80	96.44	91.03	99.73	14.01	97.33
TCSN9	15.18	94.82	96.52	91.22	99.50	12.54	97.38
RJF1	17.34	85.03	94.24	83.23	98.85	13.49	97.62
RJF2	28.60	95.77	93.80	88.26	99.57	24.34	98.66
RJF3	18.08	89.17	95.45	85.92	99.63	14.85	96.94
RJF4	21.71	91.40	95.90	87.03	99.56	18.18	97.12
RJF5	40.23	93.06	95.18	88.59	99.58	34.01	97.47

Table S3. Summary for SNPs in 67 chickens. Related to Table 1.

Category			Number	Ratio
Upstream			144,087	1.35%
Gene body	UTR5		129,631	1.21%
	UTR5/UTR3		5,702	0.05%
	CDS	Missense	55,480	0.52%
		Synonymous	123,589	1.16%
		Nonsyn/Syn ratio (ω)	0.453	-
		Stop gain	467	0.00%
		Stop loss	79	0.00%
		unknown	1,114	0.01%
	Intronic		5,626,745	52.63%
	Splicing		456	0.00%
	UTR3		236,288	2.21%
Downstream			141,610	1.32%
Upstream/Downstream			15,936	0.15%
Intergenic			4,200,109	39.28%
ncRNA		exonic	2,067	0.02%
		intronic	8,288	0.08%
		splicing	2	0.00%
ts			7,730,048	72.30%
tv			2,961,602	27.70%
ts/tv			2.61	-
Total			10,691,650	100.00%

Table S4. Summary for InDels in 67 chickens. Related to Table 1.

Category			**Number**	**Ratio**
Upstream			28,102	1.35%
Gene Body	UTR5		24,097	1.16%
	UTR5/UTR3		1,222	0.06%
	CDS	Frameshift deletion	4,169	0.20%
		Frameshift insertion	2,629	0.13%
		Non-frameshift deletion	2,571	0.12%
		Non-frameshift insertion	1,201	0.06%
		Stopgain	162	0.01%
		Stoploss	13	0.00%
		Unknown	156	0.01%
	Intronic		1,098,080	52.65%
	Splicing		667	0.03%
	UTR3		53,160	2.55%
Downstream			29,494	1.41%
Upstream/Downstream			3,656	0.18%
Intergenic			834,122	40.00%
ncRNA		exonic	395	0.02%
		intronic	1,612	0.08%
		splicing	1	0.00%
Total			2,085,509	100.00%

Table S5. Sample details for group abbreviations in the main text. Related to Table 1 and Figure 1.

Sample ID	Groups	group abbreviations	Cluster	Cluster abbreviations
SL1	Sri Lankan	SL	Sri Lankan sub-group 1	SL1
SL3	Sri Lankan	SL	Sri Lankan sub-group 1	SL1
SL7	Sri Lankan	SL	Sri Lankan sub-group 1	SL1
SL9	Sri Lankan	SL	Sri Lankan sub-group 1	SL1
SL2	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL4	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL5	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL6	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL8	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL10	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SL11	Sri Lankan	SL	Sri Lankan sub-group 2	SL2
SA1	Saudi Arabian	SA	-	-
SA2	Saudi Arabian	SA	-	-
SA3	Saudi Arabian	SA	-	-
SA4	Saudi Arabian	SA	-	-
SA5	Saudi Arabian	SA	-	-
TC1	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1
TC7	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1
TCAB1	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1
TCAB3	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1
TCLZ2	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1
TCSN1	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 1 in TC1	TC1_sub1

TC2	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC3	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC4	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC5	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC8	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC9	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TC10	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCDQ3	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCDQ4	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCLZ3	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCQH5	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCQH8	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCQH9	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCQH10	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCQH11	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCSN6	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCSN7	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCSN8	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCSN9	Tibetan chicken group 1	TC1	Tibetan chicken sub-group 2 in TC1	TC1_sub2
TCAB5	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCDQ1	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCDQ5	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCGZ1	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCGZ3	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCGZ4	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1

TCGZ5	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TCGZ6	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 1 in TC2	TC2_sub1
TC6	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCAB6	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCAB7	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCDQ2	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCDQ6	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCGZ10	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCLZ1	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCLZ4	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCLZ5	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCQH1	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCSN3	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCSN4	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
TCSN5	Tibetan chicken group 2	TC2	Tibetan chicken sub-group 2 in TC2	TC2_sub2
RJF1	Red junglefowl	RJF	-	-
RJF2	Red junglefowl	RJF	-	-
RJF3	Red junglefowl	RJF	-	-
RJF4	Red junglefowl	RJF	-	-
RJF5	Red junglefowl	RJF	-	-

Table S6. D-statistic between Tibetan chickens and SL chickens. Related to Figure 1.

P1	P2	P3	P4	D value	SE	Z score
TC1_sub1	TC2_sub2	SL	RJF	-0.0589	0.0015	-40.4315
TC2_sub1	TC2_sub2	SL	RJF	-0.0542	0.0014	-38.2555
TC1_sub2	TC2_sub2	SL	RJF	-0.0484	0.0011	-42.6085
TC1_sub1	TC1_sub2	SL	RJF	-0.0120	0.0010	-11.9055
TC2_sub1	TC1_sub2	SL	RJF	-0.0049	0.0013	-3.7505

Table S7. D-statistic between Tibetan chickens and SA chickens. Related to Figure 1.

P1	P2	P3	P4	D value	SE	Z score
TC2_sub1	TC1_sub1	SA	RJF	-0.05	0.00203783	-24.59
TC2_sub1	TC2_sub2	SA	RJF	-0.046	0.00167179	-27.64
TC2_sub1	TC1_sub2	SA	RJF	-0.04	0.00155033	-25.8
TC1_sub2	TC1_sub1	SA	RJF	-0.011	0.00156117	-7.046
TC1_sub2	TC2_sub2	SA	RJF	-0.006	0.00155738	-3.66

Table S8. KEGG enrichment of PSGs in SA breed from the comparison between SA and TC2_Sub2. Related to Figures 2-6. KEGG pathway enrichment analysis was implemented by KOBAS 3.0 software. Ontologies with a cutoff value of 0.05 for *P*-values as being biologically significant.

#Term	ID	PSGs number	P-Value	PSGs
Metabolic pathways	gga01100	78	1.46E-08	*MSMO1, IDO2, G6PD, APIP, NDUFB8, ATP5J2, NDUFB5, GNPDA1, ALG6, PIGB, NDUFS3, NOS2, TPK1, P4HA1, ME1, RRM2, PGS1, GADL1, TBXAS1, PDHX, B3GNT5, HSD17B12, CYP2J21, CYP2J23, CYP2J22, AASS, PLA2G12B, ACSM3, AFMID, GCLC, HPGDS, DGKI, LPCAT2, GBE1, ETNK1, SYNJ2, ATP6V0D2, ST3GAL2, SUCLA2, GATB, TAT, PLCB2, SIIL, AMY2A, ACSS1, EXTL3, PTDSS1, AK9, ATP6V1H, PIGP, TPH1, MGAT2, RIMKLB, PIK3C2G, TK1, PRPS2, CDS2, NDUFA9, PAFAH1B1, MAN1A1, MAN1A2, FTCD, GALNT12, LPIN1, B3GALT5, SC5D, HEXA, NDUFA4, LOC101747660, MTHFD1L, AKR1D1, SDHD, CYP2J19, PPAT, POLE2, GFPT1, HLCS, PIGK*
Melanogenesis	gga04916	13	3.36E-05	*PLCB2, ADCY7, ADCY1, CAMK2D, GNAO1, ADCY9, MAP2K1, KIT, WNT8B, LEF1, RAF1, CREBBP, WNT7B*
Salmonella infection	gga05132	11	4.94E-05	*IL18, NOS2, TJP1, DYNC1I1, ARPC1B, ARPC1A, IFNGR2, LOC112529922, FLNB, ACTB, LOC112532987*
Gap junction	gga04540	12	5.24E-05	*PLCB2, TJP1, TUBA3E, ADCY1, TUBB4B, LOC100858386, ADCY9,*

				MAP2K1, ADCY7, PRKG1, ITPR2, RAF1
MAPK signaling pathway	gga04010	21	9.00E-05	*CACNA1I, RPS6KA2, TGFB3, FLNB, TNFRSF1A, CACNB2, CREB5, LOC100858386, MAP3K7, CACNA1G, CACNA1C, MAP2K1, LOC101750513, NLK, NF1, LOC112529922, FGF12, MAPKAPK5, DUSP16, RAF1, LOC112532987*
GnRH signaling pathway	gga04912	11	2.23E-04	*PLCB2, ADCY7, ADCY1, CAMK2D, MMP2, CACNA1C, ADCY9, MAP2K1, LOC100858386, ITPR2, RAF1*
Vascular smooth muscle contraction	gga04270	12	4.62E-04	*PRKCH, PLCB2, ADCY7, ADCY1, PLA2G12B, CACNA1C, KCNMA1, ADCY9, MAP2K1, PRKG1, ITPR2, RAF1*
Influenza A	gga05164	14	6.29E-04	*IL18, SOCS3, TNFRSF1A, CREB5, IL12B, MAP2K1, IFNAR1, TLR7, IFNGR2, ACTB, CPSF4, RAF1, CREBBP, AGFG1*
Wnt signaling pathway	gga04310	13	8.32E-04	*PLCB2, LRP6, TBL1XR1, PPARD, MAP3K7, DKK1, CAMK2D, WNT8B, NFATC2, LEF1, NLK, CREBBP, WNT7B*
Adrenergic signaling in cardiomyocytes	gga04261	12	2.05E-03	*PLCB2, ADCY7, ADCY1, LOC112533460, CAMK2D, CACNB2, CREB5, CACNA1C, ADCY9, ATP1B3, KCNQ1, PPP2R5A*
Arachidonic acid metabolism	gga00590	7	2.54E-03	*CYP2J21, TBXAS1, CYP2J23, CYP2J22, PLA2G12B, CYP2J19, HPGDS*
Linoleic acid metabolism	gga00591	5	3.76E-03	*PLA2G12B, CYP2J19, CYP2J21, CYP2J23, CYP2J22*
Adherens junction	gga04520	8	4.38E-03	*TJP1, NLK, MAP3K7, PTPRM, LEF1, ACTB, CREBBP, CTNNA3*
Calcium signaling pathway	gga04020	13	6.56E-03	*CACNA1I, NOS2, ADCY7, ADCY1, ERBB4, CAMK2D, CACNA1C, PLCB2, ADCY9, ITPR2, CACNA1G, LOC112533510, PHKG1*
Oocyte meiosis	gga04114	9	6.95E-03	*RPS6KA2, ADCY7, ADCY1, CAMK2D, ADCY9, MAP2K1, AR, ITPR2, PPP2R5A*
Cytokine-cytokine receptor interaction	gga04060	13	1.05E-02	*IFNAR1, IL18, EDA, ACVR1, IFNGR2, TNFRSF1A, EDA2R, IL12B, KIT, TGFB3, CCR6, CCR9, CCL20*
Glycosphingolipid biosynthesis	gga00603	3	1.20E-02	*HEXA, ST3GAL2, B3GALT5*

- globo series				
Glutathione metabolism	gga00480	5	1.74E-02	*HPGDS, GSTAL2, RRM2, GSTA2, GCLC*
Glycerophospholipid metabolism	gga00564	8	1.88E-02	*PTDSS1, LPIN1, PLA2G12B, CDS2, DGKI, LPCAT2, ETNK1, PGS1*
Progesterone-mediated oocyte maturation	gga04914	7	2.31E-02	*RPS6KA2, ADCY7, ADCY1, ADCY9, PDE3B, MAP2K1, RAF1*
Purine metabolism	gga00230	11	2.61E-02	*AK9, ADCY7, PRPS2, ADCY1, PDE3A, ADCY9, PDE3B, PPAT, RRM2, PDE4B, POLE2*
Base excision repair	gga03410	4	2.69E-02	*PCNA, POLE2, LIG1, LIG3*
ErbB signaling pathway	gga04012	7	2.76E-02	*AREG, PTK2, ERBB4, CAMK2D, MAP2K1, PAK6, RAF1*
Oxidative phosphorylation	gga00190	9	2.80E-02	*NDUFA4, NDUFB8, ATP5J2, NDUFB5, ATP6V1H, NDUFA9, NDUFS3, ATP6V0D2, SDHD*
Endocytosis	gga04144	15	2.84E-02	*RAB11FIP3, AMPH, TGFB3, ERBB4, SNX4, GRK6, GRK7, ARPC1A, KIT, ZFYVE9, DNM3, ARPC1B, CHMP1A, AP2B1, AGAP1*
mTOR signaling pathway	gga04150	10	3.56E-02	*RPS6KA2, RRAGA, TNFRSF1A, GRB10, ATP6V1H, MAP2K1, WNT8B, LRP6, RAF1, WNT7B*
Apoptosis	gga04210	9	3.98E-02	*TUBA3E, TNFRSF1A, MAP2K1, TRADD, XIAP, ITPR2, ACTB, RAF1, HTRA2*
Phagosome	gga04145	9	4.87E-02	*TUBA3E, DYNC1I1, TUBB4B, THBS2, ATP6V1H, NOX3, LAMP2, ATP6V0D2, ACTB*

Table S10. KEGG enrichment of PSGs in SL breed from the comparison between SL and TC2_Sub2. Related to Figures 2-4. KEGG pathway enrichment analysis was implemented by KOBAS 3.0 software. Ontologies with a cutoff value of 0.05 for *P*-values as being biologically significant.

Term	ID	PSG number	*P*-value	PSGs
Adrenergic signaling in cardiomyocytes	gga04261	18	5.99E-07	*PPP2R3A, ADCY1, ATP2B2, PLCB4, MYH15, ADCY2, CAMK2D, CACNA1C, ADCY8, MAPK14, CACNA1S, AKT3, CREM, MAPK13, CACNA2D3, PPP2R5C, ATP2B4, LOC101750560*
GnRH signaling pathway	gga04912	12	3.88E-05	*CACNA1C, PLCB4, ADCY1, ADCY2, CAMK2D, ADCY8, LOC100858386, EGFR, MAPK14, GNRH1, MAPK13, CACNA1S*
Metabolic pathways	gga01100	64	5.23E-05	*MSMO1, AOC3, TRAK2, CHPF2, MBOAT2, P4HA1, SGMS2, TPK1, PLCB4, BCO1, URAH, PAPSS1, SGPL1, UROC1, GCNT3, CES1L1, HKDC1, NDST2, DNMT1, DGKI, PFKP, PLCH2, PAICS, ITPKB, GALNTL6, SPTLC3, NT5C1A, MRI1, GOT2, BCKDHB, EPRS, PTS, DGKE, LSS, EXTL3, ACACA, TBXAS1, LOC107051134, ACSS3, IDH2, SEPHS1, HK1, COMT, ALOX15B, GALNT9, PIGV, LOC112533497, LAMA3, NAT8L, MCEE, MAN1A2, CMBL, TRIT1, BPNT1, CDS2, G6PC, AGPAT5, PPAT, POLE3, ANPEP, MGAT4C, PRIM2, SDHD, GCSH*
Calcium signaling pathway	gga04020	17	7.50E-05	*CACNA1B, ATP2B2, PLCB4, ADCY1, ADCY2, CAMK2D, ADCY8, SLC25A4, CACNA1C, EGFR, LOC112533510, CACNA1S, TACR2, ITPKB, PHKA2, TACR3, ATP2B4*

Neuroactive ligand-receptor interaction	gga04080	23	9.80E-05	*GABRQ, GABRG1, GABBR2, UTS2R, GABRA3, S1PR2, S1PR1, ADRA2C, GRM8, GLP1R, MC1R, GRID2, MC3R, MTNR1B, MTNR1A, TACR3, TACR2, PARD3, GLRA2, GRIK2, GRIK4, SSTR1, CRHR1*
MAPK signaling pathway	gga04010	20	1.62E-04	*TGFB3, LOC101750560, CACNA2D3, CACNA1S, MAP4K4, DUSP6, CACNA1B, LOC100858386, EGFR, PTPRR, MAPK14, CACNA1C, FLNC, AKT3, MAPK13, FGF12, FGF14, FGFR2, PTPN5, CASP3*
Insulin signaling pathway	gga04910	12	1.38E-03	*PPP1R3E, ACACA, LOC101750560, HKDC1, RPS6KB1, HK1, G6PC, PPARGC1A, GSK3B, AKT3, SORBS1, PHKA2*
Gap junction	gga04540	9	2.52E-03	*TJP1, PLCB4, ADCY1, ADCY2, ADCY8, LOC100858386, EGFR, TUBB3, GUCY1A2*
Adherens junction	gga04520	8	3.63E-03	*TJP1, EGFR, SORBS1, PARD3, PTPRB, CDH1, CTNNA3, IQGAP3*
AGE-RAGE signaling pathway in diabetic complications	gga04933	9	4.63E-03	*TGFB3, PLCB4, CASP3, PIM1, MAPK14, AKT3, MAPK13, COL3A1, LOC101750560*
Melanogenesis	gga04916	9	4.97E-03	*MC1R, PLCB4, ADCY1, ADCY2, CAMK2D, ADCY8, GSK3B, ASIP, WNT7B*
Progesterone-mediated oocyte maturation	gga04914	8	6.31E-03	*LOC101750560, ADCY2, MAD1L1, ADCY8, MAPK14, AKT3, MAPK13, ADCY1*
Purine metabolism	gga00230	12	9.03E-03	*PAPSS1, ADCY1, ADCY2, ADCY8, GUCY1A2, PAICS, PPAT, POLE3, PDE5A, NT5C1A, PRIM2, URAH*
Vascular smooth muscle contraction	gga04270	9	1.13E-02	*PRKCH, PLCB4, ADCY1, ADCY2, CACNA1C, ADCY8, RAMP2, GUCY1A2, CACNA1S*
Butirosin and neomycin	gga00524	2	1.36E-02	*HKDC1, HK1*

biosynthesis				
Propanoate metabolism	gga00640	4	1.96E-02	*ACSS3, BCKDHB, MCEE, ACACA*
Galactose metabolism	gga00052	4	2.43E-02	*HKDC1, HK1, G6PC, PFKP*
Fructose and mannose metabolism	gga00051	4	2.68E-02	*HKDC1, PFKFB3, HK1, PFKP*
Ubiquitin mediated proteolysis	gga04120	9	3.37E-02	*WWP2, UBE2N, UBE2E3, PARK2, HERC1, PPIL2, LOC107051050, UBE2S, UBE2F*
Salmonella infection	gga05132	6	3.37E-02	*IL18, TJP1, DYNC1H1, MAPK14, MAPK13, FLNC*
FoxO signaling pathway	gga04068	9	3.66E-02	*SIRT1, TGFB3, LOC101750560, EGFR, S1PR1, G6PC, MAPK14, AKT3, MAPK13*
Glycerolipid metabolism	gga00561	5	4.25E-02	*MBOAT2, AGPAT5, DGKE, DGKI, LOC112533497*
Dorso-ventral axis formation	gga04320	3	4.76E-02	*ETS1, EGFR, SPIRE2*
Wnt signaling pathway	gga04310	8	4.78E-02	*PLCB4, LOC107051757, LRP5, CAMK2D, INVS, DKK2, GSK3B, WNT7B*

Table S12. The list of chickens living in tropical climates. Related to Figures 4 and 5.

Sample ID	Sampling location
India1	India
India11	India
India13	India
India15	India
India17	India
India19	India
India21	India
India23	India
India25	India
India27	India
India29	India
india3	India
India31	India
India33	India
India35	India
India37	India
India39	India
India41	India
India43	India
India45	India
India47	India
india49	India
India5	India
India51	India
India55	India
India57	India
India59	India
India61	India
India7	India
India9	India
Xcelris_10A	India
Xcelris_10B	India
ypt2082	Pakistan
ypt2086	Pakistan
ypt2087	Pakistan
ypt2091	Pakistan
ypt2092	Pakistan

ypt2093	Pakistan
ypt2094	Pakistan
ypt2095	Pakistan
ypt2098	Pakistan
ypt2100	Pakistan
ypt2101	Pakistan
ypt2102	Pakistan
ypt2103	Pakistan
ypt2105	Pakistan
ypt2106	Pakistan
ypt2107	Pakistan
ypt2108	Pakistan
ypt2109	Pakistan
ypt2110	Pakistan
ypt2111	Pakistan
ypt2112	Pakistan
ypt2114	Pakistan
ypt2115	Pakistan
ypt2116	Pakistan
ypt2117	Pakistan
ypt2118	Pakistan
ypt2119	Pakistan
ypt2120	Pakistan
ypt2121	Pakistan
ypt2122	Pakistan
ypt2124_L4_I052	Srilanka
ypt2126_L3_I053	Srilanka
ypt2127_L1_I036	Srilanka
ypt2128_L6_I054	Srilanka
ypt2144_L6_I055	Srilanka
ypt2145_L6_I056	Srilanka
ypt2146_L6_I001	Srilanka
Xcelris_97	Bihar(hybrid), India
Xcelris_5A	border area of AP, TN & KNT (VELLORE), India
Xcelris_5B	border area of AP, TN & KNT (VELLORE), India
67S_L1_I011	Esfahan, Iran
69S	Esfahan, Iran
70S	Esfahan, Iran
71S_L3_I023	Esfahan, Iran
72S_L1_I024	Esfahan, Iran
73S_L1_I025	Esfahan, Iran

74S_L4_I026	Esfahan, Iran
ypt957	Esfahan, Iran
ypt958	Esfahan, Iran
ypt959	Esfahan, Iran
ypt960	Esfahan, Iran
ypt961	Esfahan, Iran
Xcelris_12A	kashmir, India
35S_L5_I041	Mahhad, Iran
ypt912	Mahhad, Iran
ypt913	Mahhad, Iran
ypt915	Mahhad, Iran
ypt916	Mahhad, Iran
ypt917	Mahhad, Iran
Xcelris_6A	punjab, India
Xcelris_6B	punjab, India
127S_L2_I045	Shiraz, Iran
128S_L1_I042	Shiraz, Iran
129S_L3_I043	Shiraz, Iran
130S_L4_I044	Shiraz, Iran
17S_L3_I024	Shiraz, Iran
18S_L7_I025	Shiraz, Iran
19S_L4_I026	Shiraz, Iran
20S_L7_I027	Shiraz, Iran
82S	Shiraz, Iran
83S_L3_I033	Shiraz, Iran
ypt909_L5_I004	Shiraz, Iran
ypt910_L3_I005	Shiraz, Iran
ypt911_L6_I006	Shiraz, Iran
ypt924_L3_I021	Shiraz, Iran
ypt925_L2_I022	Shiraz, Iran
ypt926_L2_I023	Shiraz, Iran

Table S13. Population frequencies at the focal *EPAS1* SNP in 67 chickens. Related to Figure 2.

Allele/genotype	SA frequency (5)	SL frequency (11)	RJF (5)	TC (46)
A	0.700	0.046	0.1	0
T	0.300	0.955	0.9	0
AA	2	0	0	0
TT	0	10	4	46
AT	3	1	1	0
Missing	0	0	0	0
Sample Number	5	11	5	46

Table S14. Population frequencies at the focal EPAS1 SNP in 845 chickens. Related to Figure 2.

Site / Mutation	Intron	Allele / genotype	Domesticated Chickens		TCs (112)	RJFs (149)	ALL (845)
			TRCs (109)	NTRCs (475)			
Chr3:26726127 T->C	6th	C	0.130	0.021	0.000	0.023	0.032
		T	0.870	0.980	1.000	0.977	0.968
		CC	9	2	0	0	11
		TT	80	424	106	127	737
		CT	7	14	0	6	27
		Missing	13	35	6	16	70
Chr3:26731748 T->A	6th	A	0.167	0.055	0.005	0.065	0.065
		T	0.846	0.938	0.995	0.952	0.921
		AA	8	7	0	5	20
		TT	80	428	103	125	736
		AT	20	38	1	8	67
		Missing	1	2	8	11	22
Chr:26731832 A->G	6th	G	0.182	0.057	0.005	0.039	0.064
		A	0.818	0.943	0.995	0.961	0.936
		GG	9	8	0	2	19
		AA	77	425	103	133	738
		AG	21	38	1	7	67
		Missing	2	4	8	7	21

Table S15. Comparisons of genotypes of TRCs and NTRCs with the background of RJFs and all chickens using Chi-square tests. Related to Figure 2.

Site	Groups	TRCs (109)	NTRCs (475)	TCs (112)	RJFs (149)	ALL (845)
Chr3 : 26726127	TRCs (109)	-	2.19E-12	5.81E-08	6.03E-06	1.66E-10
	NTRCs (475)	2.19E-12	-	3.57E-02	8.34E-01	1.06E-01
Chr3 : 26731748	TRCs (109)	-	2.16E-08	3.56E-09	3.54E-04	1.32E-07
	NTRCs (475)	2.16E-08	-	1.75E-03	5.19E-01	3.05E-01
Chr3 : 26731832	TRCs (109)	-	1.32E-09	4.93E-10	1.33E-07	1.02E-09
	NTRCs (475)	1.32E-09	-	1.32E-03	2.20E-01	5.15E-01

Transparent Methods

Whole-genome sequence quality filtering

We used the genome sequences for 36 high-altitude Tibetan chickens (NCBI accession: SRP067615) that we had previously generated, and downloaded 32 additional accessions from the NCBI database, including accession PRJNA241474, containing genome sequences for 6 red junglefowl and 10 Tibetan chickens, and accession PRJNA453469, containing sequences for 11 Sri Lankan indigenous chickens and 5 Saudi Arabian indigenous chickens. Approximately 1.67 Tb of whole-genome sequences based on Illumina paired-end technology was used in this study. To avoid reads with artificial bias that mainly resulted from base-calling duplicates and adapter contamination, fastp software (Chen et al., 2018) was used to perform quality control with the default parameters. Consequently, we obtained 1.56 Tb (~20.25-fold per individual) of high-quality paired-end reads, including 95.13% and 88.98% nucleotides with phred quality ≥ Q20 (with an accuracy of 99.0%) and ≥Q30 (with an accuracy of 99.9%), respectively.

Read mapping and genomic variant calling

After sequence quality filtering, we first mapped the remaining high-quality sequences to the chicken reference genome (GRCg6a, GCF_000002315.6) using BWA software (v. 0.7.17-r1188) (Abasht et al., 2009) with the command 'mem -t 10 -k 32 -M'. Second, we converted SAM format to BAM format using the package SAMtools (v.1.9) (Wahlberg et al., 2009). Third, we sorted BAM files using the package Sambamba (v. 0.6.8) (Tarasov et al., 2015). Finally, the sorted bam file was marked in duplicate using the command "MarkDuplicates" in the package picard (v. 2.18.15, http://broadinstitute.github.io/picard). Subsequently, we performed individual gVCF calling according to the best practices using the Genome Analysis Toolkit (GATK, version v4.1.2.0) (McKenna et al., 2010) with the HaplotypeCaller-based method and then population SNP calling by merging all gVCFs with the commands

“GenomicsDBImport” and “GenotypeGVCFs”. To obtain credible population SNP and InDel sets, we performed a screening process as follows:

(a) For filtering SNPs, the hard filter command ‘VariantFiltration’ was applied to exclude potential false-positive variant calls with the parameter ‘--filterExpression "QD < 2.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0"’.

(b) For filtering InDels, the hard filter command ‘VariantFiltration’ was applied to exclude potential false-positive variant calls with the parameter ‘"QD < 2.0 || FS > 200.0 || SOR > 10.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0"’.

(c) Screening of biallelic variants was performed with a Hardy–Weinberg equilibrium p-value >= 0.01. After this filtering step, we obtained 23,746,923 high-quality SNPs and 2,854,845 high-quality InDels.

(d) Variants were filtered out when the proportion of samples within the population lacking the variant was > 10% and the minor allele frequency (MAF) was < 0.01.

Consequently, a total of 10,691,650 biallelic high-quality SNPs and 2,085,509 biallelic high-quality InDels were retained for subsequent analyses.

Annotation of genomic variants

Genomic variant annotation was performed according to the chicken genome using the package ANNOVAR (Wang et al., 2010). Based on the genome annotation, genomic variants were categorized as being in exonic regions (overlapping with a coding exon), intronic regions (overlapping with an intron), splice sites (within 2 bp of a splicing junction), upstream and downstream regions (within a 1-kb region upstream or downstream from the transcription start site), and intergenic regions. The functional consequences of the variants in coding regions were further grouped into synonymous, nonsynonymous, frameshift, nonframeshift, stop-gain, stop-loss and unknown. Unknown function indicates various errors in the gene structure definition in the database file.

Phylogenetic tree and population structure

To investigate the genetic relationships between 67 chickens, we constructed a phylogenetic tree using the maximum likelihood (ML) algorithm with the GTRGAMMA model and 100 bootstrap iterations based on PGVs using the RAxML program (version: 8.2.12) (Stamatakis, 2014). The population genetic structure was examined via an expectation maximization algorithm, as implemented in the program FRAPPEv1.170 (Tang et al., 2005). The number of assumed genetic clusters K ranged from 2 to 9, with 10,000 iterations for each run. We also conducted PCA to evaluate genetic structure using GCTA software (Yang et al., 2011).

Genome-wide selective sweep scanning

Based on chicken population SNP data sets, four methods were applied to identify genome-wide selection signals.

Firstly, by comparing the 10 pairwise genetic differentiation (F_{ST}) (Weir and Cockerham, 1984) between these populations including breeds/cluster of SA, SL, TC1_Sub2, TC2_Sub2 and RJFs with a sliding window (40-kb windows sliding in 20-kb steps), we employed a modified population branch statistic (PBS) approach (Yi et al., 2010; Zhan et al., 2014), which showed very powerful to detect incomplete selective sweeps over short divergence times. Our approach was designed to take advantage of three outgroup and used to identify SA/SL breeds specific selection genomic regions. Based on the original PBS algorithm, we specifically modified the PBS formula to detect SA breed as follows:

$$(T^{\text{SA-TC2\_2}} + T^{\text{SA-TC1\_2}} + T^{\text{SA-RJF}} - T^{\text{TC1\_2-TC2\_2}} - T^{\text{TC1\_2-RJF}} - T^{\text{TC2\_2-RJF}}) / 3$$

Where T represents the population divergence time in units scaled by the population size, which is the negative log transformed (1 - F_{ST}) between two populations. The same formula was also used to identify selection sweep of the SL breed. We considered as the candidate selection regions when PBS values of the comparative

sliding windows at a significance of $P < 0.01$ (Z-test). Therefore, we identified 24.18 Mb (2.30% of the genome) and 15.58 Mb (1.48% of the genome), encompassing 723 and 464 genes, in SA and SL breeds, respectively.

Second, nucleotide diversity ($\theta\pi$) (Nei and Li, 1979) were calculated based on a sliding window (40-kb windows sliding in 20-kb steps) in two populations, A and B. The statistic $\log_2(\theta\pi^A / \theta\pi^B)$ was then calculated with respect to A and B populations. A negative value (1% outlier) suggests selection in population A, a positive value (1% outlier) suggests selection in population B.

Third, the test of cross-population composite likelihood ratio (XP-CLR) (Chen et al., 2010) was performed with the following parameters: sliding window size, 0.1 cM; grid size, 10 k; maximum number of SNPs within a window, 300; and correlation value for 2 SNPs weighted with a cutoff of 0.99. The general genetic maps for chickens used in this study were 2.8 cM/Mb for chr1-9 and 6.4 cM/Mb for chr10-28 and chr32 (Axelsson et al., 2005). Using the TC2_Sub2 cluster as the control population, we performed the XP-CLR test for SA and SL breeds, respectively. The thresholds for identifying candidate selection sweep regions were top 1% of XP-CLR score.

Fourth, the sign of cross-population extended haplotype homozygosity (XP-EHH) indicated the direction of selection for the two alleles in two populations with default parameters (Sabeti et al., 2007). For the XP-EHH score between two populations of A and B, a positive value (top 1% of all) suggests selection in population A, whereas a negative value (bottom 1% of all) suggests selection in population B.

By integrating the four methods, we identified candidate regions supported by at least one method. Subsequently, all candidate selection regions were assigned to corresponding SNPs and genes.

Genotype frequency analysis

To infer the patterns of genotype frequency changes, we set a priority order with

TRCs (tropical chickens), NTRCs (non-tropical chickens), and Tibetan chickens. Genotype frequencies for SNPs were first calculated for each population, and then a linear regression was performed to examine the patterns with the priority order. The equation for the linear regression was as follows: $Y = aX + b$, where Y is the estimated value of genotype frequency, X was the observed value of genotype frequency, a was the slope, and b was the intercept.

Functional enrichment analysis for PSGs

We submitted the PSGs to PANTHER for Gene Ontology (GO) enrichment analysis of GO biological process (GO-BP), molecular function (GO-MF) and cellular component (GO-CC) terms (Gene Ontology et al., 2013) with the background of all known chicken genes using the binomial distribution test **(Mi et al., 2013)**. KEGG is a database resource for understanding the high-level functions and utilities of biological systems, such as cells, organisms and ecosystems, based on molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (http://www.genome.jp/kegg/). KEGG pathway enrichment analysis of PSGs was implemented by KOBAS 3.0 software (Kanehisa et al., 2008; Mao et al., 2005) based on the statistical significance of the binomial distribution test results. For ontologies, we considered a cutoff value of 0.05 for P-values as being biologically significant (Mi et al., 2013).

Heat stress assay

The C2C12 cell line, an immortalized mouse myoblast cell line, was purchased from Shanghai Life Science Research Institute, Chinese Academy of Science. Cells were grown in high-glucose Dulbecco's modified Eagle's medium (DMEM) (Gibco, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS) (Gibco, Carlsbad, CA), 100 μg/mL streptomycin and 100 U/mL penicillin. Cells were maintained in a

humidified atmosphere containing 5% CO_2 at 37 ℃ for normal growth. For the heat stress assay, we cultured the cells under two temperature conditions: 42 ℃ for heat stress and 37 ℃ for the control. Heat stress treatment at 42 ℃ was performed for 3, 6 and 12 hours. Then, the cells were harvested at each time point for further qRT-PCR to investigate gene expression.

RNA extraction and real-time PCR

Cellular total RNA was extracted using RNAiso Plus reagent (Takara, Japan) according to the manufacturer's protocols and reverse transcribed to cDNA using EasyScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Transgen Biotech, Beijing, China) following the manufacturer's recommendations. The expression levels of each gene were normalized to that of *ACTB*, which was used as an endogenous control. We performed quantitative real-time polymerase chain reaction (qRT-PCR) on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA) using TransStart Top Green qPCR SuperMix (Transgen Biotech, Beijing, China) following the manufacturer's protocols. The qRT-PCR measurements were evaluated in triplicate. The fold change in expression levels was determined using the $2^{-\triangle\triangle Ct}$ method. Data were analyzed by GraphPad Prism software 7.0 (GraphPad, USA), and $p < 0.05$ was considered statistically significant.

Identification of molecular convergence

Convergent sites included both "parallel" and "convergent" sites that were previously defined (Zhang and Kumar, 1997). We used 11 species, which included six birds and five mammals, to identify convergent amino acid substitutions for adaptation to tropical climate. Of these species, three also experience desert stress, namely, the southern ostrich (*S. camelus*) (Zhang et al., 2015), which is abundant in dry areas with low rainfall in Africa; the yellow-throated sandgrouse (*P. gutturalis,* NCBI accession

GCF_000699245.1), which was collected from the Sharjah Breeding Center in the United Arab Emirates (Zhang et al., 2014); and Bactrian camels (Jirimutu et al., 2012), which live in desert regions. Other species were used as control branches, such as the White throated tinamou (*Tinamus guttatus,* GCA_002215935.1) (Zhang et al., 2014), the domestic pigeon (*Columba livia,* GCF_000337935.1) (Shapiro et al., 2013), turkey (*Meleagris gallopavo,* GCF_000146605.3) (Dalloul et al., 2010), human (*Homo sapiens,* GCF_000001405.39), the mouse (*Mus musculus,* GCF_000001635.26) and cattle (*Bos taurus,* GCF_002263795.1). Amino acid sequences for each node of the 4,844 single-copy orthologs were deduced under the "aaml" model using CODEML in PAML (Yang, 1997) based on the phylogenetic relationship **(Fig. 6D)**. Convergent substitutions between the two divergent lineages of species must meet the following criteria: $\mathbf{S_1 \neq S_2}$; $\mathbf{S_3 = S_4}$; $\mathbf{S_3 \neq S_1}$ and $\mathbf{S_4 \neq S_2}$. The criteria for parallel substitutions are $\mathbf{S_1 = S_2}$; $\mathbf{S_3 = S_4}$; $\mathbf{S_3 \neq S_1}$ and $\mathbf{S_4 \neq S_2}$, where $\mathbf{S_1}$, $\mathbf{S_2}$, $\mathbf{S_3}$, and $\mathbf{S_4}$ represent ancestral amino acids for the interior nodes in the tree **(Fig. 6E)**. Comparisons of observed convergent sites versus those generated under random expectation were performed under the JTT-f_{gene} amino acid substitution models for each gene, using scripts provided in a previous study (Zou and Zhang, 2015).

References

Abasht, B., Sandford, E., Arango, J., Settar, P., Fulton, J.E., O'Sullivan, N.P., Hassen, A., Habier, D., Fernando, R.L., Dekkers, J.C., *et al.* (2009). Extent and consistency of linkage disequilibrium and identification of DNA markers for production and egg quality traits in commercial layer chicken populations. BMC genomics *10 Suppl 2*, S2.

Axelsson, E., Webster, M.T., Smith, N.G., Burt, D.W., and Ellegren, H. (2005). Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Res *15*, 120-125.

Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. Genome Res *20*, 393-402.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884-i890.

Dalloul, R.A., Long, J.A., Zimin, A.V., Aslam, L., Beal, K., Blomberg Le, A., Bouffard, P., Burt, D.W., Crasta, O., Crooijmans, R.P., *et al.* (2010). Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. PLoS Biol *8*, e1000475.

Gene Ontology, C., Blake, J.A., Dolan, M., Drabkin, H., Hill, D.P., Li, N., Sitnikov, D., Bridges, S., Burgess, S., Buza, T., *et al.* (2013). Gene Ontology annotations and resources. Nucleic Acids Res *41*, D530-535.

Jirimutu, Wang, Z., Ding, G., Chen, G., Sun, Y., Sun, Z., Zhang, H., Wang, L., Hasi, S., Zhang, Y., *et al.* (2012). Genome sequences of wild and domestic bactrian camels. Nat Commun *3*, 1202.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., *et al.* (2008). KEGG for linking genomes to life and the environment. Nucleic Acids Res *36*, D480-484.

Mao, X., Cai, T., Olyarchuk, J.G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics *21*, 3787-3793.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat Protoc *8*, 1551-1566.

Nei, M., and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA *76*, 5269-5273.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., *et al.* (2007). Genome-wide detection and characterization of positive selection in human populations. Nature *449*, 913-918.

Shapiro, M.D., Kronenberg, Z., Li, C., Domyan, E.T., Pan, H., Campbell, M., Tan, H., Huff, C.D., Hu, H., Vickrey, A.I., *et al.* (2013). Genomic diversity and evolution of the head crest in the rock pigeon. Science *339*, 1063-1067.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.

Tang, H., Peng, J., Wang, P., and Risch, N.J. (2005). Estimation of individual admixture: analytical and study design considerations. Genet Epidemiol *28*, 289-301.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032-2034.

Wahlberg, P., Carlborg, O., Foglio, M., Tordoir, X., Syvanen, A.C., Lathrop, M., Gut, I.G., Siegel, P.B., and Andersson, L. (2009). Genetic analysis of an F(2)

intercross between two chicken lines divergently selected for body-weight. BMC genomics *10*, 248.

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

Weir, B.S., and Cockerham, C.C. (1984). Estimating F-Statistics for the Analysis of Population Structure. Evolution *38*, 1358-1370.

Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet *88*, 76-82.

Yang, Z.H. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci *13*, 555-556.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S.*, et al.* (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. Science *329*, 75-78.

Zhan, S., Zhang, W., Niitepold, K., Hsu, J., Haeger, J.F., Zalucki, M.P., Altizer, S., de Roode, J.C., Reppert, S.M., and Kronforst, M.R. (2014). The genetics of monarch butterfly migration and warning colouration. Nature *514*, 317-321.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., Storz, J.F., Antunes, A., Greenwold, M.J., and Meredith, R.W. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. Science *346*, 1311-1320.

Zhang, J., Li, C., Zhou, Q., and Zhang, G. (2015). Improving the ostrich genome assembly using optical mapping data. GigaScience *4*, s13742-13015-10062-13749.

Zhang, J.Z., and Kumar, S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. Mol Biol Evol *14*, 527-536.

Zou, Z.T., and Zhang, J.Z. (2015). Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? Mol Biol Evol *32*, 2085-2096.