

RESEARCH ARTICLE

Open Access

# Population genomics identifies patterns of genetic diversity and selection in chicken



Diyang Li<sup>1†</sup>, Yan Li<sup>1†</sup>, Miao Li<sup>1†</sup>, Tiandong Che<sup>1</sup>, Shilin Tian<sup>1,2</sup>, Binlong Chen<sup>1</sup>, Xuming Zhou<sup>3</sup>, Guolong Zhang<sup>4</sup>, Uma Gaur<sup>1</sup>, Majing Luo<sup>5</sup>, Kai Tian<sup>1</sup>, Mengnan He<sup>1</sup>, Shen He<sup>1</sup>, Zhongxian Xu<sup>1</sup>, Long Jin<sup>1</sup>, Qianzi Tang<sup>1</sup>, Yifei Dai<sup>2</sup>, Huailiang Xu<sup>1</sup>, Yaodong Hu<sup>1</sup>, Xiaoling Zhao<sup>1</sup>, Huadong Yin<sup>1</sup>, Yan Wang<sup>1</sup>, Rongjia Zhou<sup>5</sup>, Chaowu Yang<sup>6</sup>, Huarui Du<sup>6</sup>, Xiaosong Jiang<sup>6</sup>, Qing Zhu<sup>1\*</sup>  and Mingzhou Li<sup>1\*</sup>

## Abstract

**Background:** There are hundreds of phenotypically distinguishable domestic chicken breeds or lines with highly specialized traits worldwide, which provide a unique opportunity to illustrate how selection shapes patterns of genetic variation. There are many local chicken breeds in China.

**Results:** Here, we provide a population genome landscape of genetic variations in 86 domestic chickens representing 10 phenotypically diverse breeds. Genome-wide analysis indicated that sex chromosomes have less genetic diversity and are under stronger selection than autosomes during domestication and local adaptation. We found an evidence of admixture between Tibetan chickens and other domestic population. We further identified strong signatures of selection affecting genomic regions that harbor genes underlying economic traits (typically related to feathers, skin color, growth, reproduction and aggressiveness) and local adaptation (to high altitude). By comparing the genomes of the Tibetan and lowland fowls, we identified genes associated with high-altitude adaptation in Tibetan chickens were mainly involved in energy metabolism, body size maintenance and available food sources.

**Conclusions:** The work provides crucial insights into the distinct evolutionary scenarios occurring under artificial selection for agricultural production and under natural selection for success at high altitudes in chicken. Several genes were identified as candidates for chicken economic traits and other phenotypic traits.

**Keywords:** Population genomics, Genetic diversity, Selection, Chicken

## Background

Since the domestication of the red jungle fowl (*Gallus gallus*) (approximately 8000 to 5400 BC) in Asia [1], domestic chickens (*Gallus Gallus domesticus*) have been subject to the combined effects of natural and artificial selection. This has resulted in marked genetic diversity in a number of traits, leading to highly specialized chicken lines with unique traits, such as cockfighting fowls, bantams, meat and egg producing chickens. There are hundreds of phenotypically distinguishable domestic chicken breeds or lines worldwide [2], providing a unique opportunity to trace the history of domesticated poultry and define the signatures of selection resulting from both

domestication and the natural environment. For example, game fowl are a group of breeds selected specifically for cockfighting, and fighting cocks possess congenital aggression towards all males of the same species [3]. Chickens' wild ancestors are of great importance when considering the maintenance and improvement of domestic chicken breeds through introgression of genetic variation from wild-type genomes.

The Tibetan chicken is a breed endemic to China and is mainly distributed in Qinghai Province and the Tibetan Plateau. This population exhibits many phenotypic adaptations to the alpine, low-air-pressure environment [4]. In contrast, lowland chickens have experienced very strong selection for traits of biological and agricultural importance. The comparative analysis of Tibetan and lowland chicken genomes has the potential to shed light on the genetic

\* Correspondence: [zhuqingsicau@163.com](mailto:zhuqingsicau@163.com); [mingzhou.li@sicau.edu.cn](mailto:mingzhou.li@sicau.edu.cn)

<sup>†</sup>Diyang Li, Yan Li and Miao Li contributed equally to this work.

<sup>1</sup>Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu, China

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



components that are shaped by high-altitude adaptations in the Tibetan chicken.

To further investigate genomic variations underlying the domestication of chicken breeds and the high-altitude adaptation of Tibetan chickens, the whole genomes of 86 domestic chickens, including 50 lowland chickens from 9 phenotypically diverse breeds in China and 36 Tibetan chickens from 6 Qinghai-Tibetan Plateau localities (Additional file 1: Table S1), together with 5 red jungle fowls (RJFs), were used for a comparative population genomics analysis; we reported these data in a previous study [5].

## Results and discussion

### Genetic diversity, population structure and introgressions

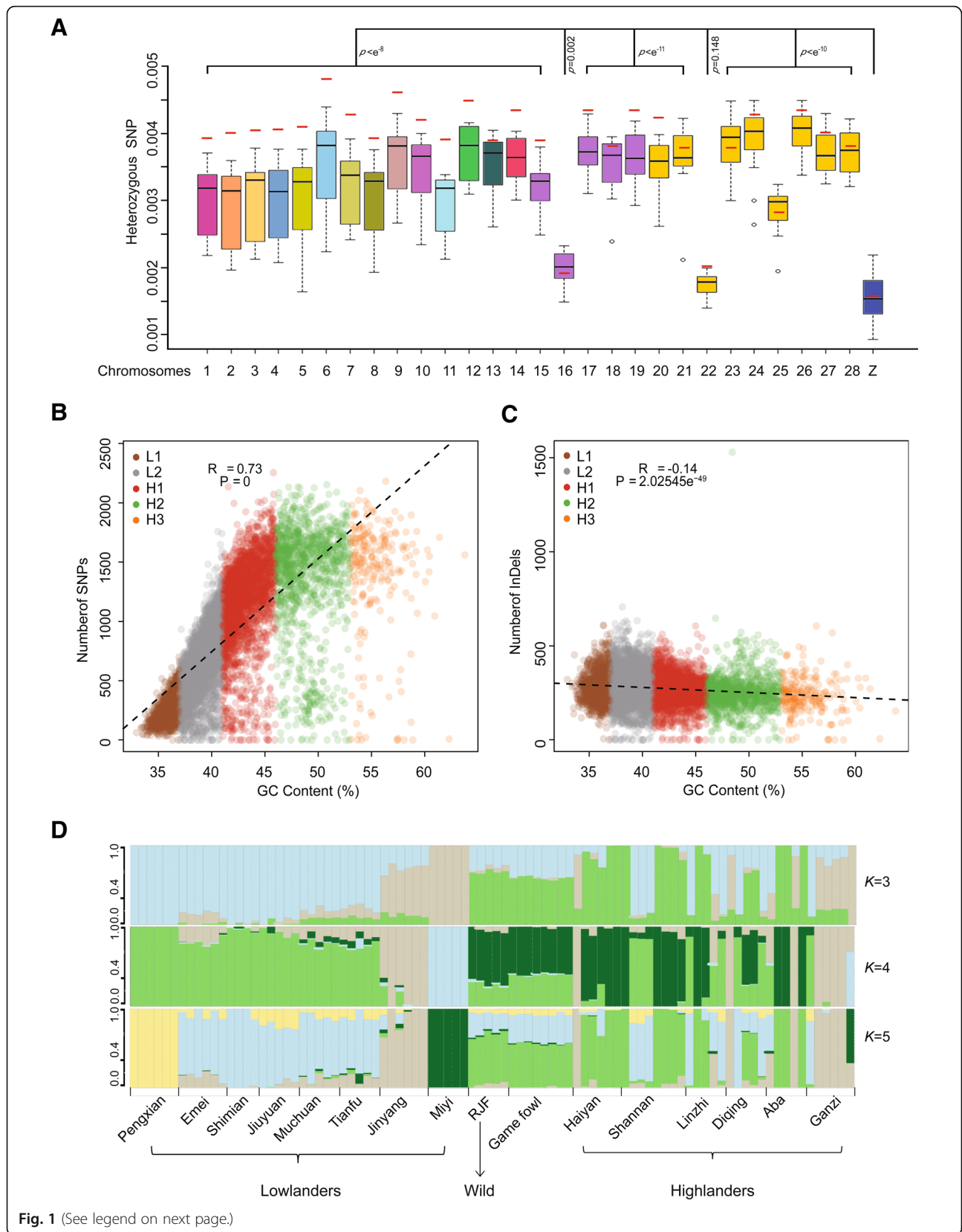
We analyzed the genetic diversity of 91 chicken genomes and identified a total of 5.27–9.59 million SNPs for each breed (Additional file 1: Figure S1). A total of 1398–4716 specific SNPs were detected for each breed/population (Additional file 1: Table S2). A small number of heterozygous breed-specific SNPs (7–89) were found for each breed. Distribution of heterozygous SNPs within each non-overlapping 500-kb window along both sex chromosomes and autosomes indicated that the Z chromosome had far fewer heterozygous SNPs than any autosome ( $p < 0.05$ ) except for chromosome 22 ( $p = 0.148$ ) (Fig. 1, Additional file 1: Figure S2A). A previous study also showed that genetic variation was significantly lower at Z-linked than at autosomal loci [6]. In addition, we performed pairwise comparisons of the genome-wide variation between the 15 domestic chicken populations. Once again, we found that higher population genetic differentiation was detected in the Z chromosome than in any of the autosomes (Additional file 1: Figure S3). The genomic level of variation between sex and autosomal chromosomes possibly help to reveal the long-term history of sexual selection in a species [7].

To examine the potential impact of GC content on genetic variations, we characterized chicken genome into five isochore families defined by different GC levels, as described by numerous references [8, 9]. The chicken genomes were segmented into isochores by increasing GC levels from L1 to H3 (Additional file 1: Figure S2.B). The overall distribution of the lengths of isochores showed that more than 40% of the isochores had 37–41% GC content (Additional file 1: Figure S2.C). We used Pearson correlation coefficient to calculate the relationship between GC content and number of SNPs/indels. The number of SNPs and the GC level are positively correlated in the isochores of the chicken genome ( $r = 0.73$ ,  $p = 0$ ) (Fig. 1b and Additional file 1: Figure S4). Surprisingly, the number of indels and the GC level in the isochores are negatively correlated in the whole chicken population ( $r = -0.14$ ,  $p = 0$ ) (Fig. 1c) and each

chicken breed/population (Additional file 1: Figure S5). The opposite trends of SNP and indel density in the context of GC level are probably due to crossing over event, the influence of natural selection and environmental pressure. The isochore families consist of 15.09, 40.67, 29.27, 12.41, and 2.55% of the chicken genome for L1, L2, H1, H2, and H3 isochore, respectively. The H1 isochore family contains 40.46% of SNPs and 32.90% of indels (Additional file 1: Figure S6; Table S3). Herein, we show that H1 isochore is the dominant source of genetic variation in the chicken genome. Although, isochore families H2 (20.36% SNP, 13.08% indels) and H3 (4.27% SNP, 2.74% indels) also represents remarkable variation, while present in small amounts in the genome. An analysis of SNPs in 91 fowls shows that the single nucleotide mutation rate is highly dependent on the GC content of the genome, which is also the case in humans [10].

Linkage disequilibrium (LD) analysis showed that Tibetan chicken populations had a faster LD decay rate than other domestic chicken breeds, as did RJFs (Additional file 1: Figure S7A), reflecting a relatively high level of inbreeding under artificial breeding programs and, consequently, lower genomic diversity in the more heavily domesticated chickens. To consider the possible genetic admixture among the populations, we also performed population structure analysis with a full maximum-likelihood approach using *frappe* [11], which estimates individual ancestry and admixture proportions assuming  $K$  ancestral populations (Fig. 1d; Additional file 1: Figure S7B). At  $K = 3$ , we observed a division between the wild fowls and modern breeds as well as the populations living at high and low altitudes. However, the red jungle fowl share a similar genetic background with game fowl based on  $K = 3, 4, 5$ . This indicates that red jungle fowl interbreed with domestic chicken and believed to be a primary progenitor of game fowl. Mutual introgression was observed among chicken breeds, and the population structure also showed that the formation of phenotypically distinguishable domestic chicken breeds driven by artificial selection. For example, we found that chickens of highland such as Ganzhi, Shannan, Linzhi, and Diqing shared the genetic contribution with lowlanders at  $K = 5$ . Interestingly, when  $K = 4$  and  $K = 5$  population clusters were tested, the observed admixture patterns of the Miyi chicken and the Pengxian yellow chicken which is two distinct population living in lowland areas have a different genetic background to other indigenous chickens, respectively. This discrepancy may reflect the species ground dwelling, territorial characteristics etc.

We then used Japanese quail (*Coturnix japonica*) [12] as an outgroup to detect introgression between RJFs and 15 domestic chicken populations (Additional file 1: Table S4). Intensive crossbreeding has occurred in the past, and



(See figure on previous page.)

**Fig. 1** Heterozygous SNPs in each chicken chromosome; correlation between variations and GC content of the genome and population structures. **a** Boxplot showing the difference in the number of heterozygous SNP per base pair between the Z chromosome and all autosomes. The black lines inside the boxes represent the median of all 86 domestic chickens. The red lines represent the median of all five RJFs. A two-tailed *t*-test was conducted to detect significant differences between the Z chromosome and all autosomes. Correlation between the number of SNPs **(b)** or the number of indels **(c)** in 91 chickens and the GC level in the isochores of the chicken genome. **d** Population structures with the number of ancestral clusters, *K*, ranging from 3 to 5

intercrossing between divergent breeds and occasionally even with wild jungle fowls was widely practiced [13]. This crossing implies a history of introgression between domestic breeds and wild populations. Our population structure analyses indicate a recent history of mutual introgression between Tibetan chickens and other breeds, as well as between wild RJFs and domestic populations. Introgression signatures were widespread in the domestic chicken populations. Archaeological discoveries in the Indus Valley and in Hebei Province, China, suggest that chickens were probably domesticated from the red jungle fowl as early as 5400 BC [14]. During the domestication of these local Chinese chicken breeds, chickens were often allowed to roam in a free-range state, and recurrent admixture between wild RJFs and domesticated chickens was very common, especially among breeds raised in mountainous areas; More introgression from RJFs to Tibetan chickens (the average number of IBDs is 3144) were found compared to others (the average number of IBDs is 2150). Thus, the most likely explanation for the dispersion pattern seen in domestic individuals is a long history of genetic exchange between RJFs and chickens. We found evidence of admixture between Tibetan chickens and other domesticated individuals possibly due to the influence of increased human activity with *K* changing progressively from 2 to 7. These crosses imply a history of introgression between domestic breeds and wild populations. Pairwise *D*-statistics based on ABBA and BABA SNP frequency differences and tracts of identity by descent (IBD) revealed that RJFs are closer to Xishuangbanna game fowls and Jinyang silky fowls than to other lowland breeds (Additional file 1: Table S5).

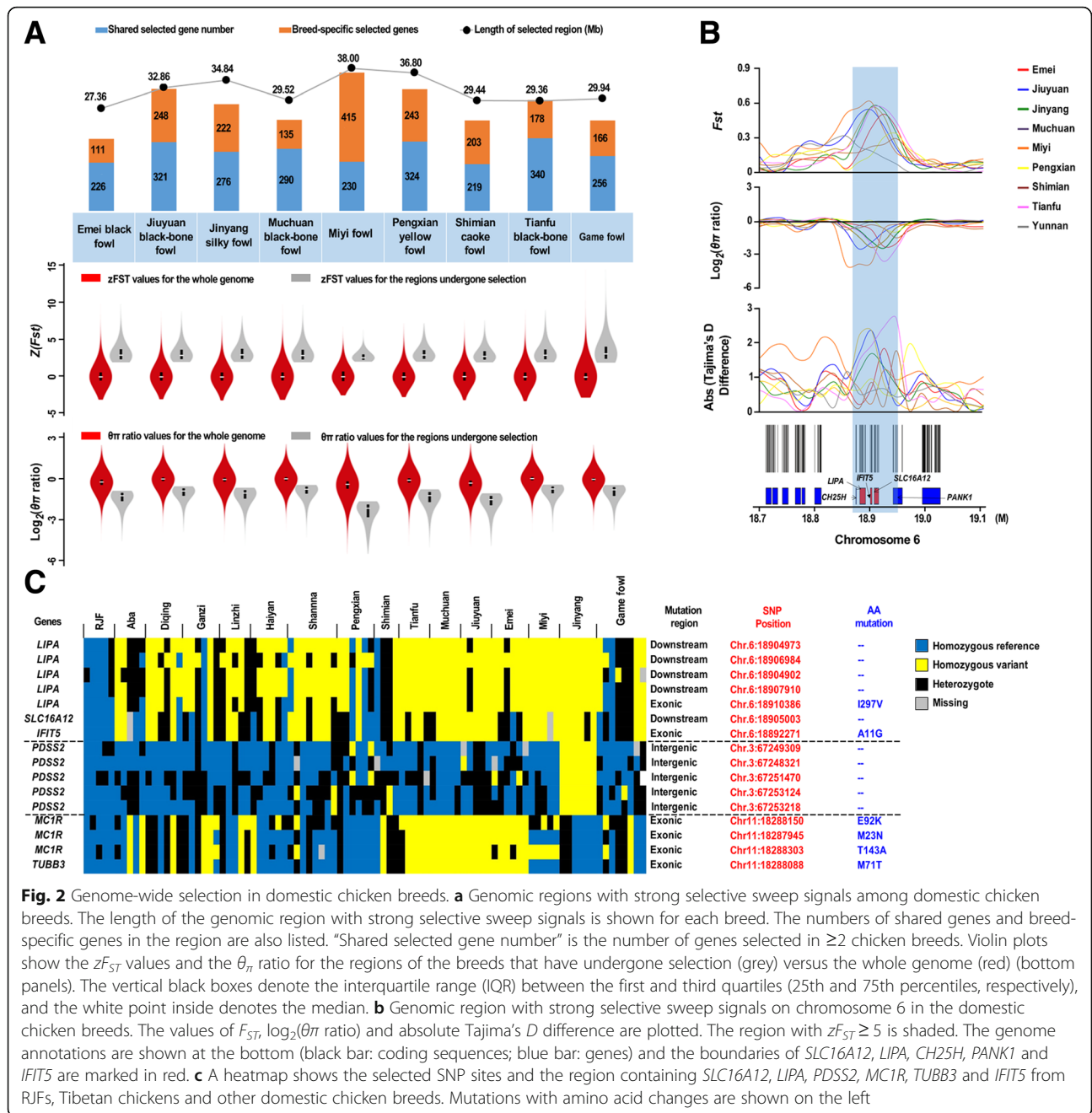
#### Genome-wide selective sweep signals in domestic fowls

To accurately detect the genomic footprints left by the intense artificial selection pressure, we performed pairwise comparisons of the genome-wide variation between the wild fowls and 10 geographically close but phenotypically diverse domestic breeds. After reviewing the distribution (Additional file 1: Figure S8), we concluded that 40 kb was the most appropriate window size because this size yielded few windows with  $\leq 20$  SNPs and retained a theoretically appropriate length to detect smaller sweeps. We identified genomic regions (with a total size of 27.36–38.00 Mb, corresponding to 2.61 to 3.63% of the genome and containing 337–645 genes per

breed) with strong selective sweep signals in each breed that also exhibited significant differences ( $p < 10^{-16}$ , Mann-Whitney *U* test) in the  $\log_2(\theta_\pi)$  ratio and  $F_{ST}$  value compared with the wild RJF genomic background (Fig. 2a; Additional file 1: Figures S9 and S10). Many genes in genomic regions with strong selective sweep signals are involved in growth, metabolism, immunity, behavior and reproduction (Additional file 1: Table S6) and may potentially contribute causally to phenotypic changes during selective breeding. We have identified a total of 74 genes that had strong selective sweep signals and were shared by more than five chicken breeds. Strikingly, five genes (*CH25H*, *PANK1*, *LIPA*, *SLC16A12* and *IFIT5*) in an 80-kb region (18.86–18.94 Mb) of chromosome 6 were under positive selection in at least five of the domestic populations, which had the highest values of  $F_{ST}$  at 0.54–0.61, and of  $zF_{ST}$  at 4.02–7.60 ( $p < 0.001$ ) (Fig. 2b and c, Additional file 2).

*LIPA* catalyzes the intracellular hydrolysis of cholesteryl esters and triglycerides in hepatocytes and macrophages, and a *LIPA* deficiency is associated with abnormal lipid deposition in multiple organs in humans [15]. *IFIT5* confers antiviral defense by disrupting protein-protein interactions in the host translation initiation machinery [16], and its expression is induced in ducks upon infection with influenza virus [17]. The selective sweep of genes involved in lipid and glucose metabolism and immune defense [18, 19] may be responsible for the dramatic phenotypic changes that are of economic value in domestic chickens, such as meat yield and disease resistance. *SLC16A12* is a member of the carboxylic acid transporter family, essential for the establishment and/or maintenance of homeostasis in the lenses and kidneys. Mutations in *SLC16A12* lead to deficiency in the transportation of metabolites, contributing to the development of cataracts and renal glycosuria in humans [20]. Recent studies have demonstrated that cholesterol 25-hydroxylase (Ch25h) is an interferon-inducible protein that can inhibit the replication of many enveloped viruses [21]. Pantothenate kinase (PanK), encoded by the gene *PANK1*, is the rate-determining enzyme in coenzyme A (CoA) biosynthesis [22]. Collectively, *LIPA*, *SLC16A12* and *IFIT5* may constitute the main genetic contributors to meat yield and disease resistance in modern chickens. This potential candidate region on chromosome 6 containing *LIPA*, *SLC16A12* and *IFIT5* could be considered a main genetic contributor to chicken domestication.





**Fig. 2** Genome-wide selection in domestic chicken breeds. **a** Genomic regions with strong selective sweep signals among domestic chicken breeds. The length of the genomic region with strong selective sweep signals is shown for each breed. The numbers of shared genes and breed-specific genes in the region are also listed. “Shared selected gene number” is the number of genes selected in  $\geq 2$  chicken breeds. Violin plots show the  $zF_{ST}$  values and the  $\theta_{\pi}$  ratio for the regions of the breeds that have undergone selection (grey) versus the whole genome (red) (bottom panels). The vertical black boxes denote the interquartile range (IQR) between the first and third quartiles (25th and 75th percentiles, respectively), and the white point inside denotes the median. **b** Genomic region with strong selective sweep signals on chromosome 6 in the domestic chicken breeds. The values of  $F_{ST}$ ,  $\log_2(\theta_{\pi}$  ratio) and absolute Tajima’s  $D$  difference are plotted. The region with  $zF_{ST} \geq 5$  is shaded. The genome annotations are shown at the bottom (black bar: coding sequences; blue bar: genes) and the boundaries of *SLC16A12*, *LIPA*, *CH25H*, *PANK1* and *IFIT5* are marked in red. **c** A heatmap shows the selected SNP sites and the region containing *SLC16A12*, *LIPA*, *PDSS2*, *MC1R*, *TUBB3* and *IFIT5* from RJFs, Tibetan chickens and other domestic chicken breeds. Mutations with amino acid changes are shown on the left

**Phenotypic traits analysis**

**Silky feathers**

We observed 204 unique selected genes in a 20.7-Mb region on the chromosomes of Jinyang silky fowls, with the highest  $F_{ST}$ -window occurring at 10.02–10.38 Mb ( $F_{ST} = 0.69$ ,  $zF_{ST} = 7.68$ ,  $p < 0.001$ ). The top ten selected genes included *CACNA2D1*, *GRM8*, *PDE7B*, *NECAB2*, *CAMK2D*, *SLC38A8*, *PDSS2*, *ZBTB43*, *NFATC3* and *ASCC3*. It should be noted that *PDSS2* ( $F_{ST} = 0.60$ ,  $zF_{ST} = 6.44$ ,  $p < 0.001$ ) is a gene previously established as the causal gene for the silky-feather phenotype [23]. The

causative mutation is located 103 base pairs upstream of the coding sequence of *PDSS2*. In addition, in this study, we found a SNP (chr3:67,304,974) located in the coding region of *PDSS2* that is specific for this breed, and the function of this mutation has not been reported yet. Further studies are needed to explore the function of this mutation.

**Body weight**

We detected the strongest selective sweep, reaching a  $zF_{ST}$  score of 9.67 ( $F_{ST} = 0.70$ ,  $p < 0.001$ ), in Muchuan

black-boned fowl; the sweep was located in the upstream non-coding region of the *SMPD3* gene. *SMPD3* controls postnatal growth and development, and inactivation of *SMPD3* is associated with skeletal deformities [24]. The represented genes encoding myofibrillar proteins (*MYH1E*, *MYH1C* and *MYPN*) also had high  $zF_{ST}$  scores. The highest differentiation peak in the Shimian Caoke fowl occurred in a region containing the *GLI3* gene ( $F_{ST}=0.66$ ,  $zF_{ST}=7.04$ ,  $p < 0.001$ ), which plays a role in the patterning of the brain and limbs [25]. These genes are associated with growth, consistent with the fact that these two chicken breeds have the greatest body weight at 180 days of age among all the included breeds (Additional file 1: Table S7).

### Egg production

Among all 10 chicken breeds we examined, Tianfu black-boned fowls lay the greatest number of eggs ( $n = 102.46$ ) at 300 days of age (Additional file 1: Table S7). We detected a selected region for laying traits that harbored the *KIF18A* gene and had the highest differentiation peak ( $F_{ST}=0.57$ ,  $zF_{ST}=7.76$ ,  $p < 0.001$ ). This is not surprising, given a recent report that *KIF18A* is required for mitotic progression during germline development [26]. We also found that *KIF18A* was under selection in Pengxian yellow fowl ( $F_{ST}=0.514$ ,  $zF_{ST}=4.668$ ) which ranked 2nd on egg numbers. Additionally, a highly differentiated peak occurred in a region containing *LIN28* ( $F_{ST}=0.44$ ,  $zF_{ST}=5.87$ ,  $p < 0.001$ ), which encodes a microRNA-binding protein [27] that is also known as a stem cell factor [28]. Testis-expressed gene 11 (*Tex11*) ( $F_{ST}=0.23$ ,  $zF_{ST}=2.59$ ,  $p < 0.001$ ), which affects testicular size in Australian Brahman cattle [29], also exhibited a selection signal in our analysis.

### Black skin and bones

Three breeds in this study (the Muchuan black-boned fowl, the Tianfu black-boned fowl and the Jiuyuan black-boned fowl) have black skin and bones, whereas the other chicken populations have white skin and bones. The inferred genome variation that is subject to a selective sweep consists of a set of 22 genes that are mainly located on chromosome 11 (Additional file 1: Table S8). One of these candidate genes, melanocortin-1 receptor (*MC1R*), plays a crucial role in melanocyte development, proliferation and differentiation and is associated with coat color in many mammals, including coloration in Holstein dairy cattle and eumelanin production in rats [30]. In chicken, *BMP7* was found to be upregulated in hyperpigmentation of the visceral peritoneum (HVP) affected chickens [31]. By comparing black-boned chickens with non-black-boned domestic fowls, we found that the genes detected by Dorshorst et al., such as *EDN3* ( $F_{ST}=0.153$ ,  $zF_{ST}=5.23$ ,  $p < 0.001$ ) and *TUBB1* ( $F_{ST}=0.156$ ,  $zF_{ST}=5.33$ ,  $p < 0.001$ ),

also showed high differentiation peaks. They found that a complex genomic rearrangement involving the endothelin 3 locus causes dermal hyperpigmentation in the Chicken [32]. In addition to these genes, our results suggested that *AEBP2*, *CHMP2B*, *PLEKHA5*, *PIT-1*, *MC1R* and *TUBB3* might also be associated with the trait of black bones, feathers and skin in birds (Fig. 2c).

### Aggressiveness

The Xishuangbanna game fowl is a famous game breed originating in Yunnan, China. However, the genes underlying the trait of aggressiveness remain unknown. It seems that positive selection of game fowls leads to convergence of neural activity, and defects in relevant genes have been implicated in the pathogenesis of Alzheimer's and Parkinson's diseases. For example, *APP*, *SNCA* and *PPT1* are involved in neural activity, and defects in those genes have been implicated in the pathogenesis of Alzheimer's disease, Parkinson's disease, and infantile neuronal ceroid lipofuscinosis, respectively [33–35]. In addition, many genes related to muscle development and cardiovascular activity, such as *FGF14* [36], *VCL* [37], *MYH11* and *SYNE1* [38] were also found to be altered in the game breeds. The protein encoded by *MYH11* is a smooth muscle myosin belonging to the myosin heavy chain family. It functions as a major contractile protein, converting chemical energy into mechanical energy through the hydrolysis of ATP [39]. Effective production of mechanical energy would be advantageous to game fowls in combat. A network of the proteins encoded by specific genes (Fig. 3a, b) shows their involvement in pathways that are probably responsible for aggressiveness, including pathways related to aromatic compound biosynthesis, muscle cell development, muscle contraction, locomotor behavior, microtubule-based processes, calmodulin binding, actin binding, GnRH (Gonadotropin-releasing hormone) signaling and MAPK (mitogen-activated protein kinase) signaling.

### High-altitude adaptation in Tibetan chickens

To determine whether the Tibetan chicken has been subject to natural selective pressure for altitude adaptation, we explored the genomic landscape of population differentiation to identify candidate genes that potentially control high-altitude-specific adaptive traits present in Tibetan chickens. Our previous study found that there are at least three distinct clusters among the six geographically representative populations of Tibetan chickens: the chicken inhabiting Tibet and Qinghai (in cluster 3, TC1) were genetically closer to the RJF than to other domestic chickens, while the Tibetan chickens inhabiting Yunnan and Sichuan (clusters 1 and 2, TC2) were closer to the domestic populations and were split into two clusters [5].

We measured genome-wide variations and the frequency spectrum separately in Tibetan and lowland chickens



(See figure on previous page.)

**Fig. 3** Aggressiveness-related genes and associated pathways. **a** Whole-genome comparison between eight game fowls and five RJFs reveals the enrichment of 46 selected genes involved in nine biological processes and pathways that contribute to the aggressiveness behavior. **b** Ingenuity Pathway Analysis (IPA) networks of aggressiveness-associated pathways. Aggressiveness proteins are represented on the outer circle with the corresponding pathways. Predicted regulators are represented on the inner circle. Different shapes denote the nature of molecules

(Additional file 1: Figure S11). A total of 158 and 117 candidate genes with selection signals were identified using sliding-window analysis in TC1 and TC2 respectively (Additional file 2). These genes were involved in 25 categories (Additional file 1: Table S9). We discovered five genes under selection in all Tibetan chickens (*TRIT1*, *HPCAL4*, *NT5CIA*, *LOC419677*, and *HEYL*) in a small region (40 kb) of chromosome 23 that showed high  $zF_{ST}$  and low  $\log_2(\theta\pi)$  ratio values (Additional file 2); these genes could help explain the local adaptation of Tibetan fowls. *TRIT1* is a tumor suppressor, and its expression can decrease cell growth [40], which could be responsible for the dwarf phenotype of Tibetan chickens. In addition, the protein encoded by the *HPCAL4* gene is a  $Ca^{2+}$ -binding protein highly similar to human hippocalcin protein and hippocalcin-like 1 protein. Human cytosolic 5'-nucleotidase is expressed at high levels in skeletal and heart muscle [41] and is involved in purine, pyrimidine, nicotinate and nicotinamide metabolism. Cytosolic 5'-nucleotidase 1A (*NT5CIA*) is also important in regulation of body fluid levels. This gene has been identified among a set of transcripts that are differentially expressed during high-altitude acclimatization in the Indian Air Force and Army Aviation Corps populations and is postulated to allow for more efficient oxygen utilization [42]. *HEYL* is a member of the Hairy-related transcription factor (HRT) family. HRT proteins may regulate specific sets of cardiac genes by modulating the function of GATA proteins and other cardiac transcriptional activators in a signal-dependent manner [43]. Tibetan chickens inhabiting high-altitude locales have increased red blood cell counts, blood oxygen affinity and hemoglobin concentrations, as well as decreased mean corpuscular volume, to overcome the harsh extremes of their environment [44]. Therefore, *NT5CIA* and *HEYL* may be involved in the high-altitude adaption of oxygen delivery in Tibetan chickens. Several candidate genes in the calcium-signaling pathway may be involved in the hypoxia adaptation experienced by Tibetan chickens [45]. Additionally, few of the positively selected genes were involved in the hypoxia-inducible factor 1 signaling pathway, which is primarily utilized in response to hypoxia in mammals such as humans [46, 47], Tibetan Mastiffs [48, 49] and horses [50]. In this study, these genes were mainly associated with energy metabolism, immune response, growth regulation and skeletal muscle development. These biological processes in particular are involved in GTPase regulation activities, suggesting the importance of energy metabolism for maintenance

of proper body temperature in Tibetan chicken populations.

We next searched for mutations in coding sequences that have become fixed or nearly fixed in Tibetan chickens. We first searched for nonsense mutations as obvious candidates of functional significance that may have contributed to rapid evolution in Tibetan chickens. However, none of the candidates occurred at a high frequency in these Tibetan chicken lineages. Then, we screened the individual sequence data for the presence of mutations that showed a marked allele frequency difference between Tibetan chickens and other domestic chickens (greater than 80% in one group, less than 20% in the other). Differences in allele frequencies between the two groups were analyzed using Pearson's chi-squared test. The  $p$  values from Pearson's chi-squared test were  $< 0.001$  for all alleles. Eleven SNPs located in coding regions were found to have extreme differences in allele frequencies between Tibetan chickens and other domestic chickens, and only three genes, namely, *PKD2L1*, *EVI5* and *ZDHHC9*, consisted of nonsynonymous mutations (Additional file 1: Figure S12A).

*PKD2L1* encodes a member of the polycystin protein family that is an integral membrane protein involved in cell-cell/matrix interactions. Animals lacking the *PKD2L1* gene are completely devoid of taste responses to sour stimuli [51]. On the Tibetan plateau, wet hulless-barley distillers' grains (WHDG) that contain a large amount of acid (such as acetic acid, propionic acid, butyric acid and lactic acid) are produced annually and commonly used in combination with drier feedstuffs to feed Tibetan domesticated animals [52]. WHDG feed is potentially part of why *PKD2L1* has been under selection in Tibetan chickens. *EVI5* is involved in the regulation of GTPase activity. Through genetic adaptations, Tibetan chickens may have developed effective strategies to cope with the acidity of their feed or to cope with the alpine environment in high-altitude regions. Mutations in the coding region of *EVI5* aligned with the orthologous protein sequences from 4 other vertebrates were shown in Additional file 1: Figure S12B. *ZDHHC9* encodes a palmitoyltransferase of NRAS and HRAS, and its mutation is associated with X-linked mental retardation [53]. Further studies are needed to confirm whether *ZDHHC9* is involved in high-altitude adaption.

In summary, we found that genes associated with high-altitude adaptation in Tibetan chickens were mainly involved in energy metabolism, body size maintenance and digestion, which could be related the chilly climate of the highlands, the high body temperature characteristic of



birds, the dwarf phenotype of Tibetan chickens, and adaptation to available food sources. Additionally, we believe that birds can adapt to hypoxic stimuli by changing the expression levels of genes. Transcriptome data from lowland and Tibetan chickens are needed to further explore the high-altitude hypoxia adaptation of birds.

## Conclusion

By analyzing multiple individual genomes representing various chicken populations and breeds, this study reveals a large and complex landscape of genetic diversity and selective sweeps that occurred during the domestication of chickens. Comparison of the genome sequences of red jungle fowls, nine lowland domestic chicken breeds and Tibetan chickens provided insights into the distinct evolutionary scenarios occurring under artificial selection for agricultural production and under natural selection for success at a high altitude. Several genes were identified as candidates for chicken economic traits and other phenotypic traits. Our results indicate that Tibetan chickens evolved in adaptation to a high-altitude alpine environment during their long history of living at high altitudes.

## Methods

### Genome sequencing, analysis of the population structure and evolutionary history

The whole genomes of 86 domestic chickens including 50 lowland chickens from 9 phenotypically diverse breeds ([Emei black fowl, Emei, 400 m altitude], [Jiuyuan black fowl, Jiuyuan, 900 m altitude], [Jinyang silky fowl, Jinyang, 460 m altitude], [Muchuan black-boned fowl, Muchuan, 500 m altitude], [Miyi fowl, Miyi, 1400 m altitude], [Pengxian yellow fowl, Pengxian, 800 m altitude], [Tianfu black-boned fowl, Chengdu, 540 m altitude], [Shimian Caoke fowl, Shimian, 790 m altitude] and Xishuangbanna game fowl) in China and 36 Tibetan chickens from 6 Qinghai-Tibetan Plateau localities ([Aba, Sichuan Province, 3300 m altitude], [Ganzi, Sichuan Province, 3390 m altitude], [Diqing, Yunnan Province, 3280 m altitude], [Haiyan, Qinghai Province, 3260 m altitude], [Linzhi, Tibet, 3100 m altitude] and [Shannan, Tibet, 3700 m altitude]) (Additional file 1: Table S1)), together with 5 red jungle fowls (RJFs), were used for comparative population genomics. Among them, the genomes of Xishuangbanna game fowl and RJFs were downloaded from NCBI (GenBank accession number PRJNA241474). Description and quantification of the phenotypic diversity of each breed are provided in Additional file 1: Table S7. The whole genomes of chickens were sequenced on the Illumina HiSeq 2000 platform. We generated sequencing libraries using a TruSeq Nano DNA HT Sample Preparation Kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions, and index codes were added to attribute

sequences to each sample. Briefly, the DNA sample was fragmented to a size of 350 bp by sonication, and the DNA fragments were then end polished, A-tailed, and ligated with the full-length adapter for Illumina sequencing with further polymerase chain reaction (PCR) amplification. The sequencing data have been submitted to NCBI Sequence Read Archive (SRA) with the accession number SRP067615. Approximately 96.30% of high-quality paired-end reads were mapped to the chicken reference genome with an average coverage depth of 18.06-fold for each individual using BWA software [54]. We first examined individual SNPs independently confirmed by both SAMtools [14] and GATK [55]. The methods for genome sequencing, sequence filtering, data analysis, SNP calling, and insertion and deletion (indel) calling can be found in our previous study [5]. To evaluate LD decay, we used Haploview [56] to calculate the coefficient of determination ( $r^2$ ) between each pair of loci. The average  $r^2$  was calculated for the pairwise markers in a 40-kb window and averaged across the whole genome. To determine the most appropriate window size, capable of yielding few windows with  $\leq 20$  SNPs while retaining a theoretically appropriate length to detect smaller sweeps, we computed the number of bins with  $\geq 20$  SNPs in 1-Mb, 500-kb, 200-kb, 100-kb, 40-kb, 20-kb and 10-kb windows (Additional file 1: Table S10). The window size with the highest proportion of bins with  $\geq 20$  SNPs was considered the best one to use. Population structure analysis with a full maximum-likelihood approach was performed using *frappe* [11].

### Segmenting the chicken genome into isochores

The chicken genome (*Gallus gallus* 4) was first split into single chromosomes. isoSegmenter (Version 1.5.1) was used to identify isochores on each chromosome, with all parameters on default settings except window size (<https://github.com/bunop/isoSegmenter>). More specifically, we calculated the standard deviation (SD) of GC for fixed window sizes ( $w$ ) ranging from 10 kb to 750 kb for all isochores comprising at least four windows. The smaller window size was selected when a plateau began to emerge across increasing window sizes [57].

### Introgression analysis – D-statistics (ABBA-BABA tests) and identity by descent (IBD)

To detect introgression between RJFs and the 10 domestic chicken breeds, we computed D-statistics based on ABBA and BABA SNP frequency differences using the following expression [58]:

$$D(P_1, P_2, P_3, \text{Japanese quail}) = \frac{\sum_{i=1}^n (1 - \hat{p}_{i1}) \hat{p}_{i2} \hat{p}_{i3} - \hat{p}_{i1} (1 - \hat{p}_{i2}) \hat{p}_{i3}}{\sum_{i=1}^n (1 - \hat{p}_{i1}) \hat{p}_{i2} \hat{p}_{i3} + \hat{p}_{i1} (1 - \hat{p}_{i2}) \hat{p}_{i3}}$$

where P1, P2, P3 and Japanese quail (P4) are the four

different populations under comparison: P1 (one domestic chicken population) and P2 (another domestic chicken population), are sister taxa, P3 comprises RJFs and P4 (Japanese quail) is outgroup. The sequence-based genotypes were phased from the 91 samples, and tracts of identity by descent (IBD) between the two chicken populations were inferred with BEAGLE 4.1 [59].

### Sweep analysis

The selective sweep screen was performed using a method that includes the sequence diversity statistic ( $\theta_\pi$ ) [60], the population differentiation statistic ( $F_{ST}$ ) [61] and Tajima's  $D$  value [62] using a 40-kb window with a 20-kb step. To detect characteristics that have undergone selection associated with high altitude or domestication, we measured the genome-wide variation between the highland (Qinghai-Tibetan Plateau) and lowland groups (Sichuan Basin) and the genome-wide variation between the domesticated breeds and the RJFs. Because sex chromosomes and autosomes are subjected to different selective pressures and have different effective population sizes, the  $Z$ -transformed population differentiation ( $zF_{ST}$ ) [63] was estimated only for the autosomes. The formula used to calculate  $zF_{ST}$  was as follows:  $z(F_{ST}) = (F_{ST} - \mu(F_{ST})) / \sigma(F_{ST})$ , where  $\mu$  is the mean of  $F_{ST}$  and  $\sigma$  is the standard deviation of  $F_{ST}$ . The genome signature with significantly high  $F_{ST}$  (corresponding to the top 5% level) and  $\theta_\pi$  ratio ( $\theta_\pi$ , lowland population/ $\theta_\pi$ , Tibetan population; top 5% level presented) values were identified as extensively diversified. The gene annotation is performed in the selected genomic regions according to the reference chicken genome. Genes were submitted to DAVID (<https://david.ncifcrf.gov/>) for enrichment analysis of the significant overrepresentation of GO biological process (GO-BP) and molecular function (GO-MF) terms, as well as InterPro domains and KEGG pathways. In all tests, the whole set of known chicken genes was defined as the background, and  $p$  values (i.e., EASE scores), indicating the significance of the overlap between various gene sets, were calculated using a modified Fisher's exact test corrected by the Benjamini-Hochberg procedure. Only GO-BP, GO-MF, KEGG pathway or InterPro domain terms with a  $p$  value less than 0.05 were considered significant and included on the list. For aggressiveness of game fowls, we first detected genes selected for game fowls, and then these candidate genes were subjected to Ingenuity Pathway Analysis (IPA) network analysis [64].

### Additional files

**Additional file 1:** This file includes Figures S1 to S12 and Tables S1 to S10. (DOCX 17545 kb)

**Additional file 2:** Lists of information containing all the positively selected genes in chicken breeds. (XLSX 1018 kb)

### Abbreviations

ATP: Adenosine triphosphate; bp: Base pairs; DNA: Deoxyribonucleic acid; GO: Gene ontology; IBD: Identity by descent; indels: Insertion/Deletions; IPA: Ingenuity Pathway Analysis; LD: Linkage disequilibrium; m: Meter; NCBI: National Center for Biotechnology Information; PCR: Polymerase chain reaction; RJF: Red jungle fowl; SNP: Single nucleotide polymorphism

### Acknowledgements

The authors gratefully acknowledge Lin Ye and Yao Zhang who helped and supported this research. We thank Shailendra Kumar Mishra who helped to revise the manuscript.

### Funding

The design and initiation of the study were supported by grants from the Sichuan Provincial Department of Science & Technology Program (2019JDTD0009). Sample collection and main body of the work were funded by the National Natural Science Foundation of China (31402063 and 31522055), the China Agricultural Research System (CARS-41), the 13th Five-Year Plan for breeding programs in Sichuan – Selective breeding of new breeds and synthetic strains of broiler (2016NYZ0025 and 2016NYZ0043). Genome sequencing and publication costs of this article were supported by the National Program for Support of Top-notch Young Professionals and the Young Scholars of the Yangtze River.

### Availability of data and materials

All data generated or analyzed in the present study are included in this article and can be found in the Additional files 1 and 2. Genomic data are available at NCBI with the accession number SRP067615 (<https://www.ncbi.nlm.nih.gov/sra/SRP067615>).

### Authors' contributions

DL, YL, QZ and MingzhouL conceived and designed the experiment; BC and MiaoL collected samples; YL, MiaoL, HX, BC, XZhou, HY, XZhao, YW, TC, ST, YH, UG, KT, MH, SH, ZX, GZ, CY, HD, XJ, MLuo, RZ, YD, LJ and QT performed the experiment and statistical analysis; DL and YL wrote the manuscript. All the authors have read and approved the final manuscript.

### Ethics approval and consent to participate

The animal handling experiments were approved by the Institutional Animal Care and Use Committee of Sichuan Agricultural University under permit number YCS-B20100804.

### Consent for publication

Not applicable

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu, China. <sup>2</sup>Novogene Bioinformatics Institute, Beijing, China. <sup>3</sup>Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. <sup>4</sup>Department of Animal Science, Oklahoma State University, Stillwater, OK, USA. <sup>5</sup>Hubei Key Laboratory of Cell Homeostasis, Laboratory of Molecular and Developmental Genetics, College of Life Sciences, Wuhan University, Wuhan, China. <sup>6</sup>Sichuan Animal Science Academy, Chengdu, China.

Received: 12 June 2018 Accepted: 19 March 2019

Published online: 02 April 2019

## References

- Xiang H, Gao J, Yu B, Zhou H, Cai D, Zhang Y, Chen X, Wang X, Hofreiter M, Zhao X. Early Holocene chicken domestication in northern China. *Proc Natl Acad Sci U S A*. 2014;111(49):17564–9.
- Ekarius C. Storey's illustrated guide to poultry breeds. *Countryside Small Stock J*. 2008;92:288.
- Guo X, Fang Q, Ma C, Zhou B, Wan Y, Jiang R. Whole-genome resequencing of Xishuangbanna fighting chicken to identify signatures of selection. *Genet Sel Evol*. 2016;48(1):62.
- Li M, Zhao C. Study on Tibetan chicken embryonic adaptability to chronic hypoxia by revealing differential gene expression in heart tissue. *Sci China C Life Sci*. 2009;52(3):284–95.
- Li D, Che T, Chen B, Tian S, Zhou X, Zhang G, Li M, Gaur U, Li Y, Luo M, et al. Genomic data for 78 chickens from 14 populations. *GigaScience*. 2017;6:1.
- Sundstrom H, Webster MT, Ellegren H. Reduced variation on the chicken Z chromosome. *Genetics*. 2004;167(1):377–85.
- Corl A, Ellegren H. The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. *Evolution*. 2012;66(7):2138–49.
- Bernardi G. The neoselectionist theory of genome evolution. *Proc Natl Acad Sci U S A*. 2007;104(20):8385–90.
- Bernardi G, Olofsson B, Filipki J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F. The mosaic genome of warm-blooded vertebrates. *Science*. 1985;228(4702):953–8.
- Fryxell KJ, Moon WJ. CpG mutation rates in the human genome are highly dependent on local GC content. *Mol Biol Evol*. 2005;22(3):650–8.
- Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol*. 2005;28(4):289–301.
- Kawahara-Miki R, Sano S, Nunome M, Shimmura T, Kuwayama T, Takahashi S, Kawashima T, Matsuda Y, Yoshimura T, Kono T. Next-generation sequencing reveals genomic features in the Japanese quail. *Genomics*. 2013;101(6):345–53.
- Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, Baig M, Beja-Pereira A, Ding ZL, Palanichamy MG, Zhang YP. Multiple maternal origins of chickens: out of the Asian jungles. *Mol Phylogenet Evol*. 2006;38(1):12–9.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
- Bernstein DL, Hukova H, Bialer MG, Desnick RJ. Cholesteryl ester storage disease: review of the findings in 135 reported patients with an underdiagnosed disease. *J Hepatol*. 2013;58(6):1230–43.
- Abbas YM, Pichlmair A, Gorna MW, Superti-Furga G, Nagar B. Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature*. 2013;494(7435):60–4.
- Vandervlen HA, Petkau K, Ryan-Jean KE, Aldridge JR Jr, Webster RG, Magor KE. Avian influenza rapidly induces antiviral genes in duck lung and intestine. *Mol Immunol*. 2012;51(3–4):316–24.
- Zhang L, Chen D, Yu L, Wei Y, Li J, Zhou C. Genome-wide analysis of the ovodefensin gene family: monophyletic origin, independent gene duplication and presence of different selection patterns. *Infect Genet Evol*. 2019;68:265–72.
- Yu LT, Xiao YP, Li JJ, Ran JS, Yin LQ, Liu YP, Zhang L. Molecular characterization of a novel ovodefensin gene in chickens. *Gene*. 2018;678:233–40.
- Kloekener-Gruissem B, Vandekerckhove K, Nurnberg G, Neidhardt J, Zeitz C, Nurnberg P, Schipper I, Berger W. Mutation of solute carrier SLC16A12 associates with a syndrome combining juvenile cataract with microcornea and renal glucosuria. *Am J Hum Genet*. 2008;82(3):772–9.
- You H, Yuan H, Fu W, Su C, Wang W, Cheng T, Zheng C. Herpes simplex virus type 1 abrogates the antiviral activity of Ch25h via its virion host shutoff protein. *Antivir Res*. 2017;143:69–73.
- Rock CO, Karim MA, Zhang YM, Jackowski S. The murine pantothenate kinase (Pank1) gene encodes two differentially regulated pantothenate kinase isozymes. *Gene*. 2002;291(1–2):35–43.
- Feng C, Gao Y, Dorshorst B, Song C, Gu X, Li Q, Li J, Liu T, Rubin CJ, Zhao Y, et al. A cis-regulatory mutation of PDSS2 causes silky-feather in chickens. *PLoS Genet*. 2014;10(8):e1004576.
- Stoffel W, Jenke B, Block B, Zumbansen M, Koebeke J. Neutral sphingomyelinase 2 (smpd3) in the control of postnatal growth and development. *Proc Natl Acad Sci U S A*. 2005;102(12):4554–9.
- Kraus P, Fraidenaich D, Loomis CA. Some distal limb structures develop in mice lacking sonic hedgehog signaling. *Mech Dev*. 2001;100(1):45–58.
- Czechanski A, Kim H, Byers C, Greenstein I, Stumpff J, Reinholdt LG. Kif18a is specifically required for mitotic progression during germ line development. *Dev Biol*. 2015;402(2):253–62.
- Tsialikas J, Romer-Seibert J. LIN28: roles and regulation in development and beyond. *Development*. 2015;142(14):2397–404.
- Yang DH, Moss EG. Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. *Gene Expr Patterns*. 2003;3(6):719–26.
- Lyons RE, Loan NT, Dierens L, Fortes MR, Kelly M, McWilliam SS, Li Y, Bunch RJ, Harrison BE, Barendse W, et al. Evidence for positive selection of taurine genes within a QTL region on chromosome X associated with testicular size in Australian Brahman cattle. *BMC Genet*. 2014;15:6.
- Ni Q, Shao Y, Wang YZ, Jing YH, Zhang YC. Impact of high altitude on the hepatic fatty acid oxidation and synthesis in rats. *Biochem Biophys Res Commun*. 2014;446(2):574–9.
- Luo C, Qu H, Wang J, Wang Y, Ma J, Li C, Yang C, Hu X, Li N, Shu D. Genetic parameters and genome-wide association study of hyperpigmentation of the visceral peritoneum in chickens. *BMC Genomics*. 2013;14:334.
- Dorshorst B, Molin AM, Rubin CJ, Johansson AM, Stromstedt L, Pham MH, Chen CF, Hallbook F, Ashwell C, Andersson L. A complex genomic rearrangement involving the endothelin 3 locus causes dermal hyperpigmentation in the chicken. *PLoS Genet*. 2011;7(12):e1002412.
- Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, Dumanchin C, Feuillette S, Brice A, Vercelletto M, et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet*. 2006;38(1):24–6.
- Linnertz C, Lutz MW, Ervin JF, Allen J, Miller NR, Welsh-Bohmer KA, Roses AD, Chiba-Falek O. The genetic contributions of SNCA and LRRK2 genes to Lewy body pathology in Alzheimer's disease. *Hum Mol Genet*. 2014;23(18):4814–21.
- Niida Y, Yokoi A, Kuroda M, Mitani Y, Nakagawa H, Ozaki M. A girl with infantile neuronal ceroid lipofuscinosis caused by a novel PPT1 mutation and paternal uniparental isodisomy of chromosome 1. *Brain Dev*. 2016;38(7):674–7.
- van Swieten JC, Brusse E, de Graaf BM, Krieger E, van de Graaf R, de Koning I, Maat-Kievit A, Leegwater P, Dooijes D, Oostra BA, et al. A mutation in the fibroblast growth factor 14 gene is associated with autosomal dominant cerebellar ataxia [corrected]. *Am J Hum Genet*. 2003;72(1):191–9.
- Taylor MR, Carniel E, Mestroni L. Cardiomyopathy, familial dilated. *Orphanet J Rare Dis*. 2006;1:27.
- Zhang Q, Skepper JN, Yang F, Davies JD, Hegyi L, Roberts RG, Weissberg PL, Ellis JA, Shanahan CM. Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. *J Cell Sci*. 2001;114(Pt 24):4485–98.
- Stern-Straeter J, Bonaterra GA, Juritz S, Birk R, Goessler UR, Bieback K, Bugert P, Schultz J, Hormann K, Kinscherf R, et al. Evaluation of the effects of different culture media on the myogenic differentiation potential of adipose tissue- or bone marrow-derived human mesenchymal stem cells. *Int J Mol Med*. 2014;33(1):160–70.
- Smaldino PJ, Read DF, Pratt-Hyatt M, Hopper AK, Engelke DR. The cytoplasmic and nuclear populations of the eukaryote tRNA-isopentenyl transferase have distinct functions with implications in human cancer. *Gene*. 2015;556(1):13–8.
- Hunsucker SA, Spychala J, Mitchell BS. Human cytosolic 5'-nucleotidase I: characterization and role in nucleoside analog resistance. *J Biol Chem*. 2001;276(13):10498–504.
- Soma S. Hypoxic signature of high altitude acclimatization: a gene expression study. *Ind J Aerospace Med*. 2012;56:1.
- Kathiriyai S, King IN, Murakami M, Nakagawa M, Astle JM, Gardner KA, Gerard RD, Olson EN, Srivastava D, Nakagawa O. Hairy-related transcription factors inhibit GATA-dependent cardiac gene expression through a signal-responsive mechanism. *J Biol Chem*. 2004;279(52):54937–43.
- Zhang H, Wu CX, Chamba Y, Ling Y. Blood characteristics for high altitude adaptation in Tibetan chickens. *Poult Sci*. 2007;86(7):1384–9.
- Wang MS, Li Y, Peng MS, Zhong L, Wang ZJ, Li QY, Tu XL, Dong Y, Zhu CL, Wang L, et al. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Mol Biol Evol*. 2015;32(7):1880–9.
- Huerta-Sanchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, et al. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*. 2014;512(7513):194–7.
- Lorenzo FR, Huff C, Myllymaki M, Olenchok B, Swierczek S, Tashi T, Gordeuk V, Wuren T, Ri-Li G, McClain DA, et al. A genetic mechanism for Tibetan high-altitude adaptation. *Nat Genet*. 2014;46(9):951–6.

48. Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z, et al. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Res.* 2014;24(8):1308–15.
49. Li Y, Wu DD, Boyko AR, Wang GD, Wu SF, Irwin DM, Zhang YP. Population variation revealed high-altitude adaptation of Tibetan mastiffs. *Mol Biol Evol.* 2014;31(5):1200–5.
50. Hendrickson SL. A genome wide study of genetic adaptation to high altitude in feral Andean horses of the paramo. *BMC Evol Biol.* 2013;13:273.
51. Huang AL, Chen X, Hoon MA, Chandrashekar J, Guo W, Trankner D, Ryba NJ, Zuker CS. The cells and logic for mammalian sour taste detection. *Nature.* 2006;442(7105):934–8.
52. Yuan X, Yu C, Shimojo M, Shao T. Improvement of fermentation and nutritive quality of straw-grass silage by inclusion of wet Hulled-barley Distillers' grains in Tibet. *Asian-Australas J Anim Sci.* 2012; 25(4):479–85.
53. Maak S, Boettcher D, Tetens J, Swalve HH, Wimmers K, Thaller G. Expression of microRNAs is not related to increased expression of ZDHHC9 in hind leg muscles of splay leg piglets. *Mol Cell Probes.* 2010;24(1):32–7.
54. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics.* 2010;26(5):589–95.
55. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–303.
56. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21(2):263–5.
57. Cozzi P, Milanese L, Bernardi G. Segmenting the human genome into Isochores. *Evol Bioinforma.* 2015;11:253–61.
58. Durand EY, Patterson N, Reich D, Slatkin M. Testing for ancient admixture between closely related populations. *Mol Biol Evol.* 2011;28(8):2239–52.
59. Browning BL, Browning SR. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics.* 2013;194(2):459–71.
60. Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A.* 1979;76(10):5269–73.
61. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution.* 1984;38:1358–70.
62. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 1989;123(3):585–95.
63. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar A, Lindblad-Toh K. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature.* 2013;495(7441):360–4.
64. Ghosh S, Loffredo CA, Mitra PS, Trnovec T, Palkovicova Murinova L, Sovcikova E, Hoffman EP, Makambi KH, Dutta SK. PCB exposure and potential future cancer incidence in Slovak children: an assessment from molecular fingerprinting by ingenuity pathway analysis (IPA(R)) derived from experimental and epidemiological investigations. *Environ Sci Pollut Res Int.* 2018;25(17):16493–507.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

