ARTICLE

Check for updates

Analysis of genome and methylation changes in Chinese indigenous chickens over time provides insight into species conservation

Tao Zeng [1,10], Jianmei Yin[2,3,10], Peishi Feng[4], Feiran Han[5], Yong Tian[1], Yuntong Wang[6], Tiantian Gu[1], Yuhui Xu[5], Yali Liu[7], Guohui Li[2,3], Liang Qu[2,3], Li Chen[1], Lihong Gu[8], Wenwu Xu[1], Qian Xue[2,3], Qingyu Wei[9], Yongqing Cao[1], Peifeng Li[9], Huiyong Zhang[2,3], Guoqin Li[1], Lijun Liu[7], Chenghao Zhou[2,3], Zhengrong Tao[1], Junda Shen[1], Wei Han [2,3✉] & Lizhi Lu [1✉]

Conservation of natural resources is a vital and challenging task. Numerous animal genetic resources have been effectively conserved worldwide. However, the effectiveness of conservation programmes and the variation information of species have rarely been evaluated. Here, we performed whole-genome and whole-genome bisulfite sequencing of 90 Chinese indigenous chickens, which belonged to the Tibetan, Wenchang and Bian chicken breeds, and have been conserved under different conservation programmes. We observed that low genetic diversity and high DNA methylation variation occurs during ex situ in vivo conservation, while higher genetic diversity and differentiation occurs during in situ conservation. Further analyses revealed that most DNA methylation signatures are unique within ex situ in vivo conservation. Moreover, a high proportion of differentially methylated regions is found in genomic selection regions, suggesting a link between the effects of genomic variation and DNA methylation. Altogether our findings provide valuable information about genetic and DNA methylation variations during different conservation programmes, and hold practical relevance for species conservation.

[1] State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Institute of Animal Husbandry and Veterinary Science, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China. [2] National Chickens Genetic Resources, Jiangsu Institute of Poultry Science, Yangzhou 225125, China. [3] Technology Innovation Co., Ltd., Jiangsu Institute of Poultry Science, Yangzhou 211412, China. [4] Zhejiang University of Technology, Hangzhou 310014, China. [5] LC-Bio Technologies CO., LTD., Hangzhou 310000, China. [6] Technical Department, Biomarker Technologies Corporation, Beijing 101300, China. [7] Zhejiang Animal Husbandry Techniques Extension and Livestock and Poultry Monitoring Station, Hangzhou 310020, China. [8] Institute of Animal Science & Veterinary medicine, Hainan Academy of Agricultural Science, Haikou 571100, China. [9] College of Animal Science, Shanxi Agricultural University (Shanxi Academy of Agricultural Sciences), Taiyuan 030032, China. [10]These authors contributed equally: Tao Zeng, Jianmei Yin.
✉email: hanwei830@163.com; lulizhibox@163.com

Animal genetic resources are the foundation of sustainable development of animal production industry, and are vital to food security and livelihoods of millions of people. Chicken (*Gallus gallus domesticus* or *Gallus domesticus*) is the most common domestic animal worldwide[1]. Nearly 1,600 different indigenous chicken breeds are internationally recognised (FAO, 2020; http://www.fao.org/poultry-production-products/production/poultry-species/chickens/en/). Yet, the value of these resources is poorly understood, and the indigenous breeds are slowly getting replaced by commercial lines. Due to this a considerable proportion of indigenous chicken breeds are becoming extinct or are at a risk of extinction, which increased from 24.75% in 2014 to 30.65% in 2018 globally[2]. Therefore, conservation of indigenous chicken breeds is important and urgent for endangered animal protection and sustainable breeding.

Currently, the conservation programmes for chickens are commonly divided into two categories: in situ conservation and ex situ in vivo conservation. In China, 128 indigenous chicken breeds have been identified and conserved at the National Conservation Farm (NCF), where the breeds evolved or are now normally found and bred (in situ)[3]. These breeds play a crucial role as a source of meat and egg, and also provide a research model for understanding the adaptations of specific breeds to specific environmental challenges, due to their various phenotypic and physiological characteristics[4]. Moreover, 28 of them have been listed in the National Conservation Catalogue and are conserved in the National Chicken Genetics Resources Gene Bank (NCGR, Jiangsu) over the past three decades (ex situ in vivo). However, for certain indigenous chicken breeds, the environment and climate in places where ex situ in vivo conservation are markedly different from those of their origins. For instance, the Tibetan chicken is a unique breed native to the Qinghai-Tibet Plateau and shows distinctive genetic adaptation to high-altitude environments[4–7]. Evaluation of the effectiveness and identification of the natural variation related to environmental adaptation during ex situ in vivo conservation of indigenous chicken breeds will further refine animal resource conservation programmes.

Since the genome of the Red Junglefowl has been sequenced and published[8], genetic basis of domestication of chicken has been widely reported using population genomics[9,10]. Recent studies have indicated that epigenetic variation plays an important role in domestication[11–14]. DNA methylation is a central epigenetic modification that plays an essential regulatory role in cellular development and environmental responses[15,16]. Natural variation for DNA methylation can represent pure epigenetic variation that occurs independently, and is not associated with any genetic variation[17]. The relationship between DNA methylation and genetic variation has been reported on a genome-wide scale in *Arabidopsis*[18]. These results provide perspectives and ideas regarding the use of genetic variation combined with DNA methylation changes to assess the differentiation of indigenous chicken breeds in different conservation programmes.

Different geographical and environmental factors have resulted in the development of divergent phenotypic characteristics and environmental adaptability of these breeds. In the present study, we performed whole-genome resequencing and whole-genome bisulfite sequencing (WGBS) of samples obtained from three phenotypically and geographically diverse Chinese indigenous chicken breeds (TC: Tibetan chicken; WC: Wenchang chicken; and BC: Bian chicken). Each breed was divided into the following three groups according to the conservation programmes: Con – cryopreserved samples, as a control, in March 2000, kept in the NCF; In – in situ conservation after ~20 years, in March 2019, kept in the NCF; and Ex – ex situ in vivo conservation after approximately 20 years, in March 2019, kept in the NCGR

(Fig. 1a and Supplementary Table 1). The aim was to evaluate the effectiveness, and investigate the genetic and DNA methylation variations occurring during different conservation programmes of three phenotypically and geographically diverse Chinese indigenous chicken breeds.

Results and discussion

Sample collection and sequencing. A total of 985.74 Gb clean data were uniquely mapped to the 1.1 Gb chicken reference genome, resulting in an average of 98.44% coverage with ~9× depth for whole-genome resequencing (Supplementary Data 1 and Supplementary Fig. 1a). After SNP calling and quality-control criteria, we identified a total of 3.93 – 4.52 million high-quality SNPs and 0.64–0.75 million small insertions and deletions (InDel) for each individual. Functional annotation of the SNPs indicated that approximately 61.09% were located in introns, followed by 32.98% in intergenic regions, 3.19% in coding sequence (CDS), 2.12% in 3′-untranslated regions (3′-UTR) and 0.62% in 5'-UTR (Supplementary Data 2). For methylation, 357.20 Gb clean data were retained, which covered 71.07% of the reference genome with an average depth of 23× (Supplementary Table 2 and Supplementary Fig. 1b). After 5mC detection, a total of 91.28 % of all methyl cytosines were identified in the CG context, 1.47% in the CHG context and 7.25 % in the CHH context (Supplementary Data 3). The majority of the methylation was CG, indicating that mCG was the main type of methylation in these chicken breeds. These results were consistent with the previous studies on other chicken and pig breeds, which indicated that mCG is the major type of methylation[19,20].

Phylogenetic analysis. The neighbour-joining tree based on the chickens' SNPs and sequence data sets featured three groups, namely TC, WC and BC, with distinct phenotype and geographic location (Fig. 1b). These three groups were also supported by principal component analysis (PCA), which explained 7.34% of the total variation that divided TC, WC and BC into separate clusters (Fig. 1c and Supplementary Fig. 2)[21]. To estimate different ancestral proportions, we used clustering models with Admixture by assuming K ancestral populations (Fig. 1d)[22]. The cross-validation results also demonstrate clear division between TC, WC and BC at $K = 3$ (Supplementary Fig. 3). Notably, chickens under different conservation programmes basically formed two smaller branches. High consistency within each breed showed a closer genetic affinity between Ex and Con groups than that between In and Con groups (Fig. 1b), suggesting that visible genetic differentiation might occur among these three indigenous chicken breeds during in situ conservation over approximately 20 years. Additionally, two separated clades of In-TC were found in the phylogeny. Shared ancestry was also observed between few In-TC and other indigenous chickens (Fig. 1d). The results showed that the possibility that other populations could be introduced to In-TC, resulting in a visible genetic differentiation during in situ conservation of TC. Similar results were observed by Wang et al., who showed that TC may have two distinct groups, suggesting that TC could likely be traced back to two genetic sources[4].

Genetic diversity and differentiation. Genome-wide linkage disequilibrium (LD) in each population was estimated as the physical genomic distance at which the genotypic association (R^2) decays to less than half of its maximum value. Overall, the rate of LD decay was very fast in all three chicken breeds. Additionally, we observed a higher level of LD value in Ex groups of the three breeds compared with the In and Con groups (Fig. 2a and
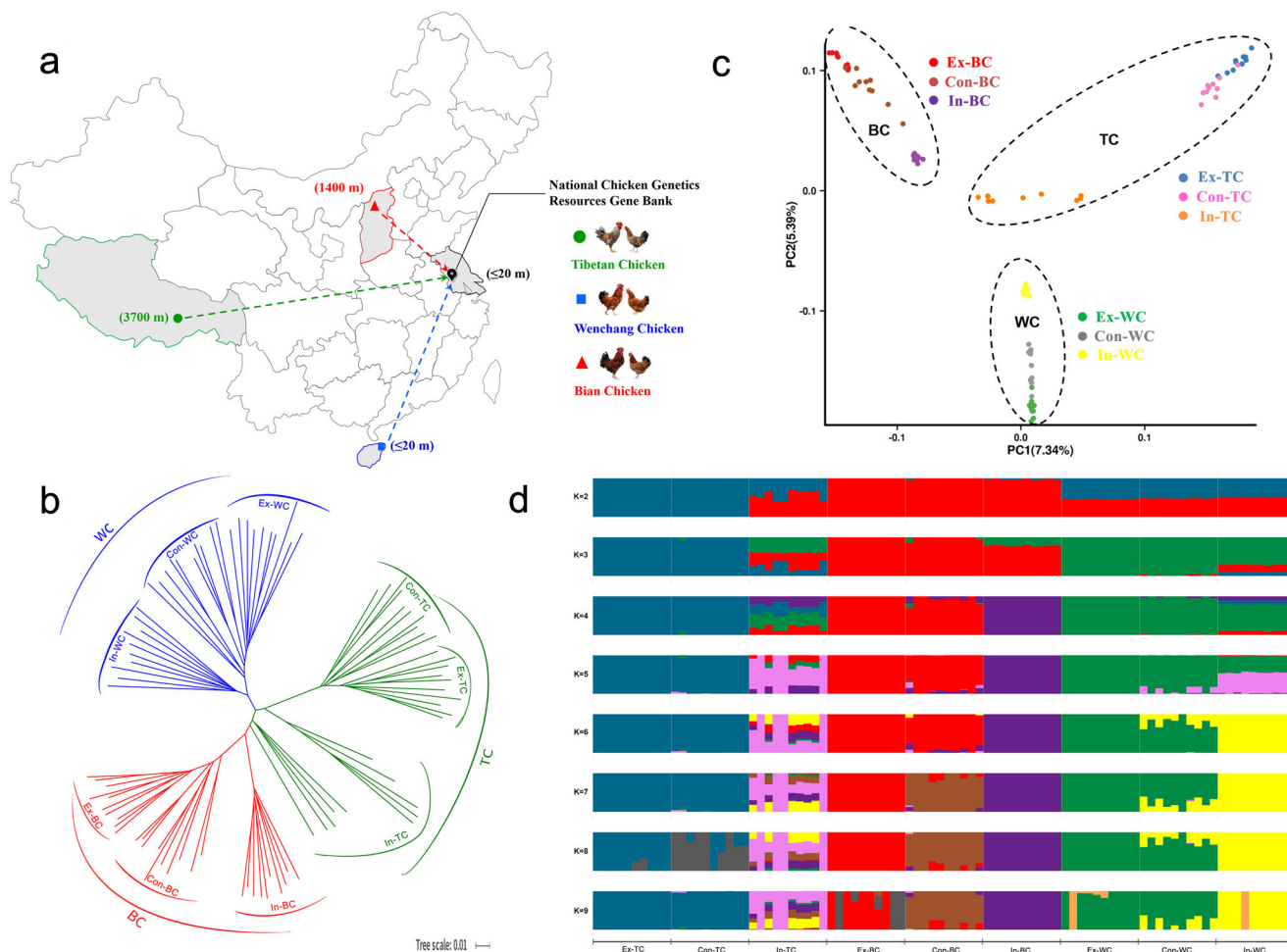
Fig. 1 Sample distribution, phylogenetics and population structure of indigenous chicken breeds during different conservation programmes.
a Geographic distribution of three selected indigenous chicken breeds. **b** The neighbour-joining tree of 90 indigenous chickens (TC: Tibetan chicken, WC: Wenchang chicken and BC: Bian chicken). Each breed was divided into three groups according to the conservation programmes (Con – cryopreserved samples, as a control, NCF; In – in situ conservation, NCF; and Ex – ex situ in vivo conservation, NCGR). **c** The principal component analysis (PCA) of 90 indigenous chickens. **d** Structure analysis of nine populations from three indigenous chicken breeds. Groupings of samples from 2–9 ancestral clusters (K) are shown.

Supplementary Fig. 4). For example, LD values for Ex-, Con- and In-TC were 246 bp, 210 bp and 216 bp, respectively. This result is consistent with a previous study, which reported that LD tends to enhanced as conservation continues under the ex situ in vivo chicken conservation[23]. The homozygosity/heterozygosity SNP ratio of different populations exhibited a high degree of consistency with the LD value, with an upward trend in the Ex groups and a decrease in the In groups when compared with Con groups (Fig. 2b).

Genetic diversity varies among species as well as within genomes, and has important implications for the conservation of species[24]. Thus, we assessed various indicators of genetic diversity under different conservation of the three indigenous chicken breeds. The average nucleotide diversity (π) and proportion of polymorphic markers (P_N) were lower in Ex groups of the three breeds than those in In groups and Con groups (Fig. 2c and Supplementary Table 3). However, strikingly, elevated levels of genetic diversity indicators were observed in the In groups of TC and WC breeds (Fig. 2c and Supplementary Table 3). Population genetic differentiation between the chicken breeds was measured by pairwise F_{ST} values. The result was in good agreement with the phylogenetic analysis results mentioned above, which indicated that the level of F_{ST} between In and Con groups was higher than

that between Ex and Con groups (Fig. 2d). Moreover, the highest F_{ST} was consistently observed between the In and Ex groups of the three indigenous chicken breeds.

In conservation biology, inbreeding, or a change in allele frequencies over time, and change in different spatial scales (ranging from local to global) can cause loss of heterozygosity and genetic diversity[25,26]. Currently, ex situ in vivo conservation programme for indigenous chicken breeds are inherent to small-population paradigm, which may increase the possibility of genetic drift due to low effective population size and demographic isolation[27]. The occurrence of genetic drift can lead to a decrease in genetic diversity during ex situ in vivo conservation. We then computed Tajima's D in all populations (Supplementary Table 3). The values of Tajima's D in all groups were close to 1 (deviation from 0), indicating a possibility of weak balancing selection or a founder effect during the process of conservation. Additionally, ecological drivers, including climate change, pest outbreaks and human activities can also deeply affect genetic diversity[28–31]. A considerable number of studies have shown that species that evolved in a complex environment, such as the Qinghai–Tibetan Plateau, have comparatively low genetic diversity, for example, locusts[32], snub-nosed monkeys[33] and sheep[34]. In our study, the environment and climate of the regions of ex situ in vivo
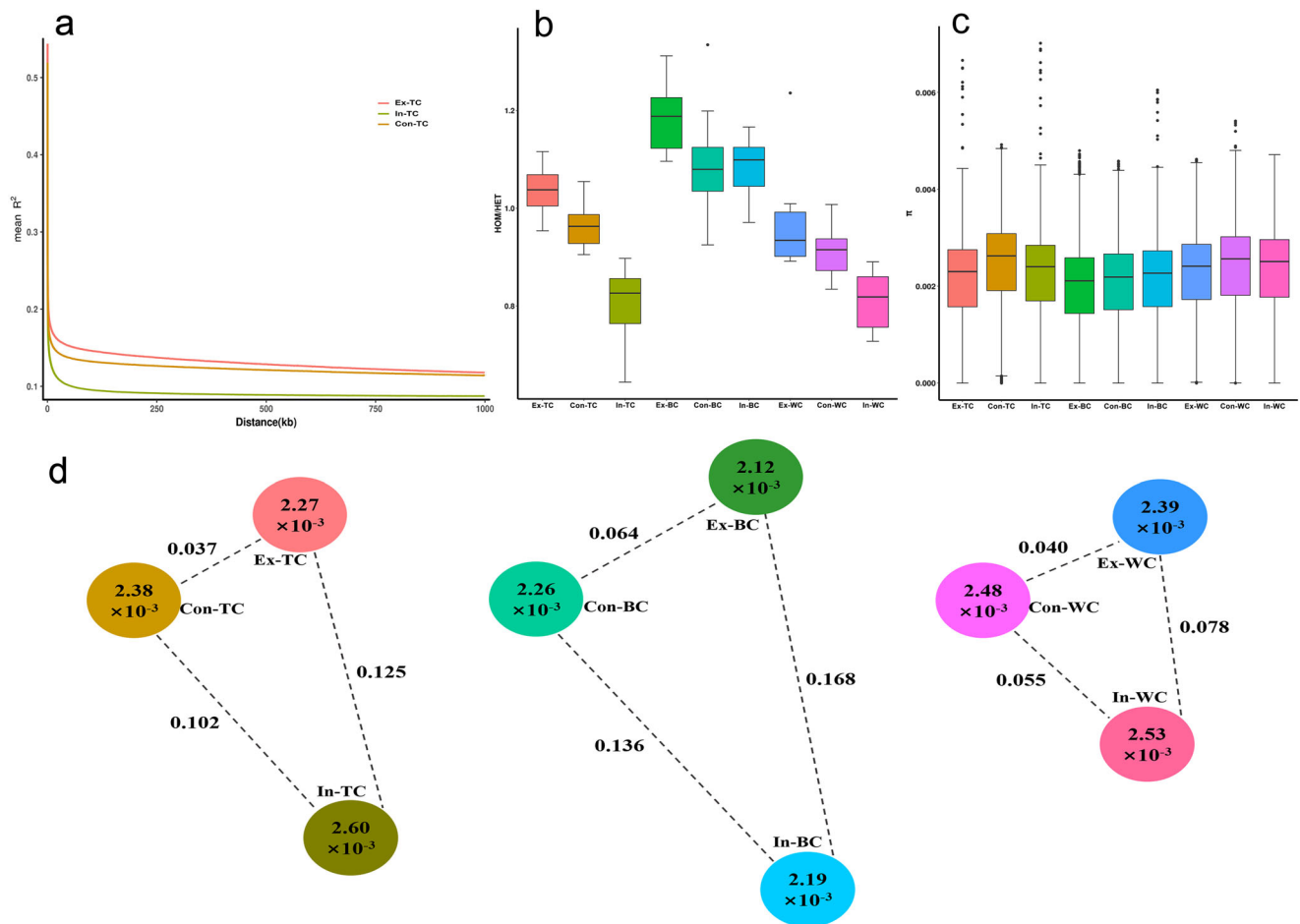
Fig. 2 Genetic diversity and differentiation of indigenous chicken breeds during different conservation programmes. a Patterns of linkage disequilibrium (LD) decay across the genome for TC during different conservation programmes (Con – cryopreserved samples, as a control, NCF; In – in situ conservation, NCF; and Ex – ex situ in vivo conservation, NCGR). LD decay in WC and BC are shown in Supplementary Fig. 4. X axis: physical distance between two SNPs marked in kb; Y axis: R^2, pearson's correlation coefficient, used to measure LD. **b** Heterozygosity across populations. Estimated as the ratio of homozygous SNPs (HOM) to heterozygous SNPs (HET) per individual. The centre line of boxplots represents median value of each population. **c** Nucleotide diversity (π) across populations. The centre line of boxplots represents median value of each population. **d** Genetic differentiation (F_{ST}) across populations in each breed of different conservation programmes. The value on each dotted line indicates F_{ST} between the two populations, and the value in each cycle represents π in this populations.

conservation were remarkably different from those of the origin of indigenous chicken breeds, especially for TC and WC (Supplementary Table 1). Therefore, short-term adaptability to environmental changes may result in the occurrence of adaptive selection, which can cause a decline in genetic diversity. In contrast to ex situ in vivo conservation, chickens that were subjected to in situ conservation exhibited higher genetic diversity and differentiation, especially in TC and WC. In fact, during in situ conservation, introduction of other populations of the same breed into the conservation population happens occasionally. However, low genetic diversity of BC was also observed during in situ conservation. This could be explained by the fact that BC has been considered as an endangered breed by the Chinese government, because of a rapid decrease in its population.

Overall, low genetic diversity occurred during ex situ in vivo conservation, while high genetic diversity and differentiation occurred during in situ conservation. Consequently, suitable conservation programmes should be formulated according to the physiological characteristics, living environment and risk status of indigenous chicken resources as follows: 1) In addition to NCF, the Local Genetics Resources Gene Bank (e.g., provincial level, in situ conservation) should be established to conserve indigenous

chicken resources whose living environment and climate are markedly different from the NCGR, such as Tibetan chicken. 2) Endangered chicken resources, such as Bian chicken, should be conserved in NCGR to fullest extent because of their limited population. 3) A combinative conservation programme of in situ and ex situ in vivo, and regular blood updates are recommended for normal chicken resources to maintain high levels of genetic diversity in the long term.

Selective signatures for high-altitude adaption in TC. Tibetans, as well as Tibetan wild and domesticated animals have revealed a highly hypoxic adaptation mechanisms to survive hypoxia occurring at high-altitudes[35–38]. To investigate the adaptive mechanisms of TC at high-altitude, we scanned the genome for regions with extreme divergence F_{ST} and the highest differences in π in 100 kb sliding windows relative to lowland populations (WC and BC). Previously, visible differentiation was observed in chickens under in situ conservation (Fig. 1b). To avoid this interference, Con group of each breed was used for selective sweep analysis. In total, we identified 228 candidate divergent regions (CDRs) ($F_{ST} \geq 0.35$) comprising 1022 candidate genes
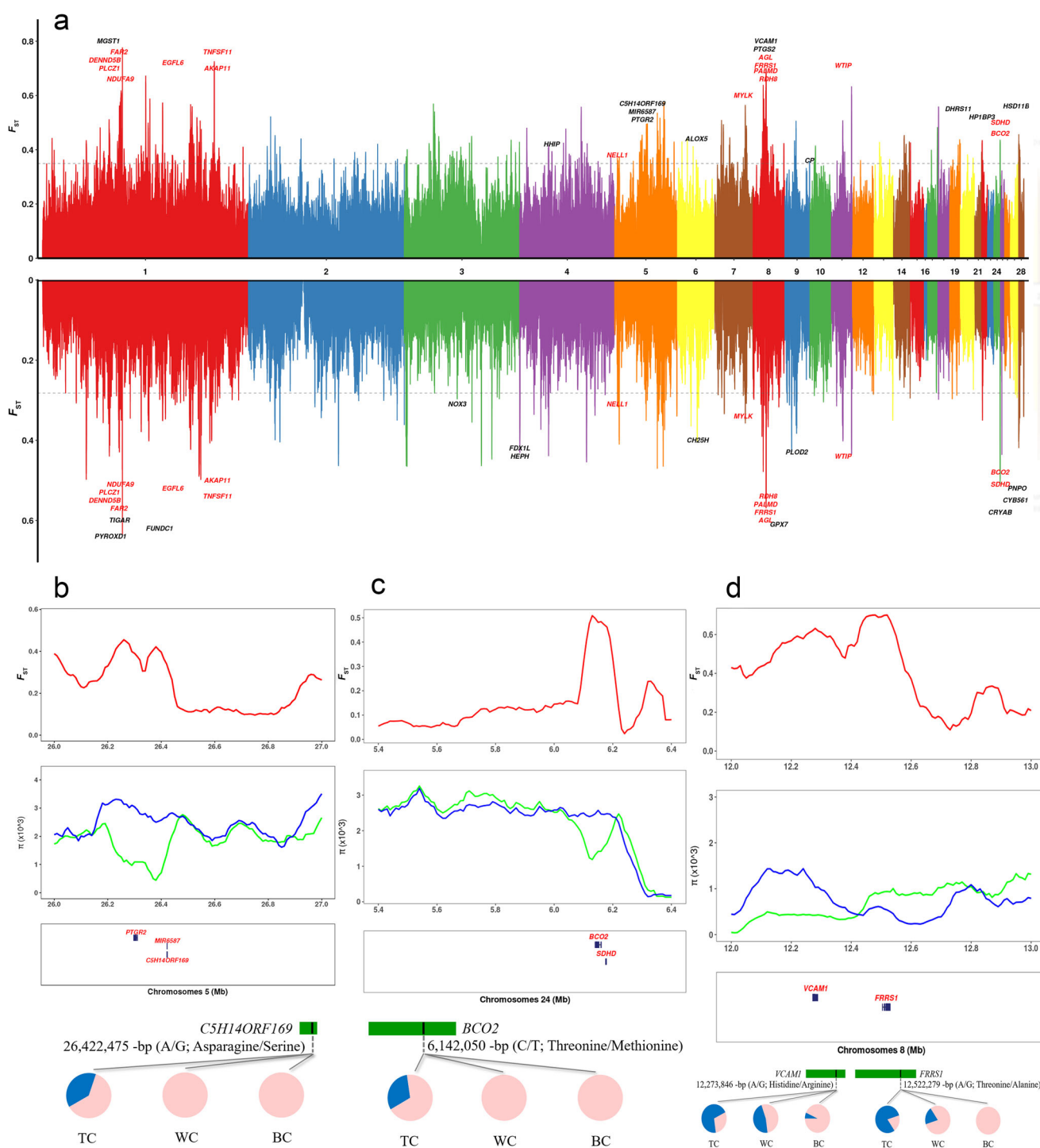
Fig. 3 Selection signals for adaptation to high-altitude hypoxia of TC. a Pairwise fixation index (F_{ST}) in 100-kb sliding windows across autosomes between TC and other chickens (upward coordinate axis: TC/BC; downward coordinate axis: TC/WC). The black horizontal dashed line corresponds to the genome-wide significance threshold ($F_{ST} = 0.35$ and 0.28 for TC/BC and TC/WC, respectively). Genes located in divergence regions and annotated by KEGG are indicated by their gene names. Overlapping candidate genes in two selected regions are marked in red. **b–d** Three extreme CDRs, which are found close to the genes in the pathways of "oxidation-reduction process" were enriched for sites on chromosomes 5 (**b**), 24 (**c**), and 8 (**d**). The pie charts represent the spectrum of allele frequencies at the non-synonymous loci of the focused genes C5H14ORF169, BCO2, VCAM1 and FRRS1 in TC and other chicken breeds. The variant allele is indicated in blue, whereas the reference allele in pink.

between TC and BC (Fig. 3a and Supplementary Data 4) and 168 CDRs ($F_{ST} \geq 0.28$) comprising 700 candidate genes between TC and WC (Fig. 3a and Supplementary Data 5). Gene Ontology (GO) analysis revealed that several categories were significantly enriched for signals of selection (Supplementary Data 6–8), such

as protein kinase A signalling and motor activity, which is related to metabolic regulation. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that a large number of genes were enriched into environmental information processing (Supplementary Fig. 5).

Table 1 Summary of candidate genes that may have a role in the high-altitude adaptation of TC.

Symbol	F_{ST} value		GO_Term	Biological category
	Con-TC/Con-BC	Con-TC/Con-WC		
FAR2	0.46	0.42	GO:0055114	Oxidation-reduction process
NDUFA9	0.62	0.43	GO:0055114	Oxidation-reduction process
FRRS1	0.70	0.57	GO:0055114	Oxidation-reduction process
RDH8	0.40	0.35	GO:0055114	Oxidation-reduction process
BCO2	0.44	0.51	GO:0055114	Oxidation-reduction process
AKAP11	0.36	0.32	GO:0005777	Peroxisome
EGFL6	0.38	0.28	GO:0005509	Calcium ion binding
NELL1	0.37	0.31	GO:0005509	Calcium ion binding
TNFSF11	0.36	0.32	GO:0055074	Calcium ion homeostasis
DENND5B	0.37	0.30	GO:0005262	Calcium channel activity
PLCZ1	0.51	0.44	GO:0006816	Calcium ion transport
MYLK	0.37	0.29	GO:0051928	Positive regulation of calcium ion transport
WTIP	0.51	0.40	GO:0001666	Response to hypoxia
SDHD	0.44	0.51	GO:0071456	Cellular response to hypoxia

Among the selected genes, 308 genes were under selection in all TCs. Further examination of these genes revealed that 14 genes, including *FAR2*, *NDUFA9*, *FRRS1*, *RDH8*, *BCO2*, *AKAP11*, *EGFL6*, *NELL1*, *TNFSF11*, *DENND5B*, *PLCZ1*, *MYLK*, *WTIP* and *SDHD*, were functionally associated with oxidation-reduction process, calcium ion binding/transport and response to hypoxia, which indicates a role in high-altitude adaptation of the TC (Table 1 and Fig. 3a)[39–52]. Moreover, three extreme CDRs were found close to the genes associated with the pathways of "oxidation-reduction process" (Fig. 3b–d). We then identified non-synonymous SNP mutations in selected genes of extreme CDRs and found that the variant allele frequencies of non-synonymous SNPs in four genes (*C5H14ORF169*, *BCO2*, *VCAM1* and *FRRS1*) were significantly different ($P < 0.01$) between TC and other chicken breeds (Fig. 3b–d).

Differentially methylated regions in chickens between in situ and ex situ in vivo conservation programmes. To elucidate the DNA methylation variations occurring during different conservation programmes of the three indigenous chicken breeds, we identified differentially methylated regions (DMRs) between different populations according to a previously described method[18,53]. Compared with Con group, 4987 DMRs were detected in Ex group of TC, including 4582 CG-DMRs, 1 CHG-DMR and 404 CHH-DMRs (Fig. 4a and Supplementary Data 9). While, compared with Con group, 3833 DMRs were identified in In group, which included 3362 CG-DMRs, 1 CHG-DMR and 470 CHH-DMRs (Supplementary Data 10). Similar trend was observed in WC with a total of 6394 DMRs (6143 CG-DMRs, 1 CHG-DMR and 250 CHH-DMRs) in Ex group and 3957 DMRs (3638 CG-DMRs, 1 CHG-DMR and 318 CHH-DMRs) in In group (Fig. 4a and Supplementary Data 11–12). In contrast, compared with Ex group of BC (3864 CG-DMRs, 69 CHG-DMRs and 223 CHH-DMRs), more DMRs were identified in In group (5360 CG-DMRs, 3 CHG-DMRs and 450 CHH-DMRs; Supplementary Data 13–14).

Then, we identified DMRs between Ex groups and In groups of the three indigenous chicken breeds. The results showed that DMRs between Ex and In groups of TC, WC and BC increased by 39.47%, 13.57% and 2.32%, respectively, when compared to DMRs between In and Con groups (Fig. 4b). The results further indicated that methylation difference between Ex and In groups were higher in TC than other breeds. Subsequently, we compared the characteristics of DMRs in different contexts, and found that CHH-DMRs were significantly different from CG-DMRs. A higher proportion of CHH-DMRs was found in genic regions, whereas CG-DMRs were mainly located in intron regions

(Supplementary Fig. 6). In addition, we found that the average length of CHH-DMRs was less than that of CG-DMRs. However, no significant difference in the length of CHH-DMRs and CG-DMRs were observed between In and Ex groups (Fig. 4c). To investigate the potential functions of the differentially methylated genes between In and Ex groups in chickens, functional enrichment analyses of the genes with DMRs were performed based on their annotation in GO and KEGG databases. The results showed that many DMRs were enriched in signal transduction, metal ion binding and transcription regulation activity. KEGG pathway analysis showed that the differentially methylated genes were mainly enriched in the neuroactive ligand-receptor interaction, MAPK signalling pathway and calcium signalling pathway (Supplementary Figs. 7–9).

DNA methylation provides a mechanism for organisms to adapt, within and between generations[54]. Numerous studies have indicated that DNA methylation can be responsive to climate change and plays an important role in certain developmental processes[13,55]. In our analysis, more DMRs were found in Ex group of TC and WC. Given the variety of DMR length, we identified overlapping bases (bp) and found that only a small proportion of these were found in the DMRs between different conservation programmes (Fig. 4d). This indicates that different DNA methylation-mediated adaptation mechanisms existed in the three chicken breeds. Further investigation showed that relatively few bases (7.41% and 7.83%) in the DMRs showed overlapping in Ex group in comparison with In group (9.89% and 13.15%) in TC and WC, respectively. However, we observed an opposite trend for different conservation programmes in BC (Fig. 4d). DNA methylation variation levels were higher in Ex group of TC and WC than their respective In group. This could be explained by the differences in climatic and geographic features of the two conservation types. Interestingly, combined with previous analyses, we found an opposite trend in DNA methylation variation levels and genetic diversity (Fig. 2c), suggesting that an association might exist between these two variations. Furthermore, we found that a few bases in the DMRs were overlapped in the Ex group of the three chicken breeds (Fig. 4e).

DNA methylation variations contribute to genomic selection signatures during Ex conservation programme of TC. DNA methylation variations can lead to extensive phenotypic variations, such as environmental adaptation[18,56], energy use efficiency[57], and disease resistance[58]. It has been proposed that methylation plays a crucial role in cellular response to hypoxia[59]. Meanwhile, recent studies have indicated that local genetic variations can influence DNA methylation[18,60,61]. To identify the
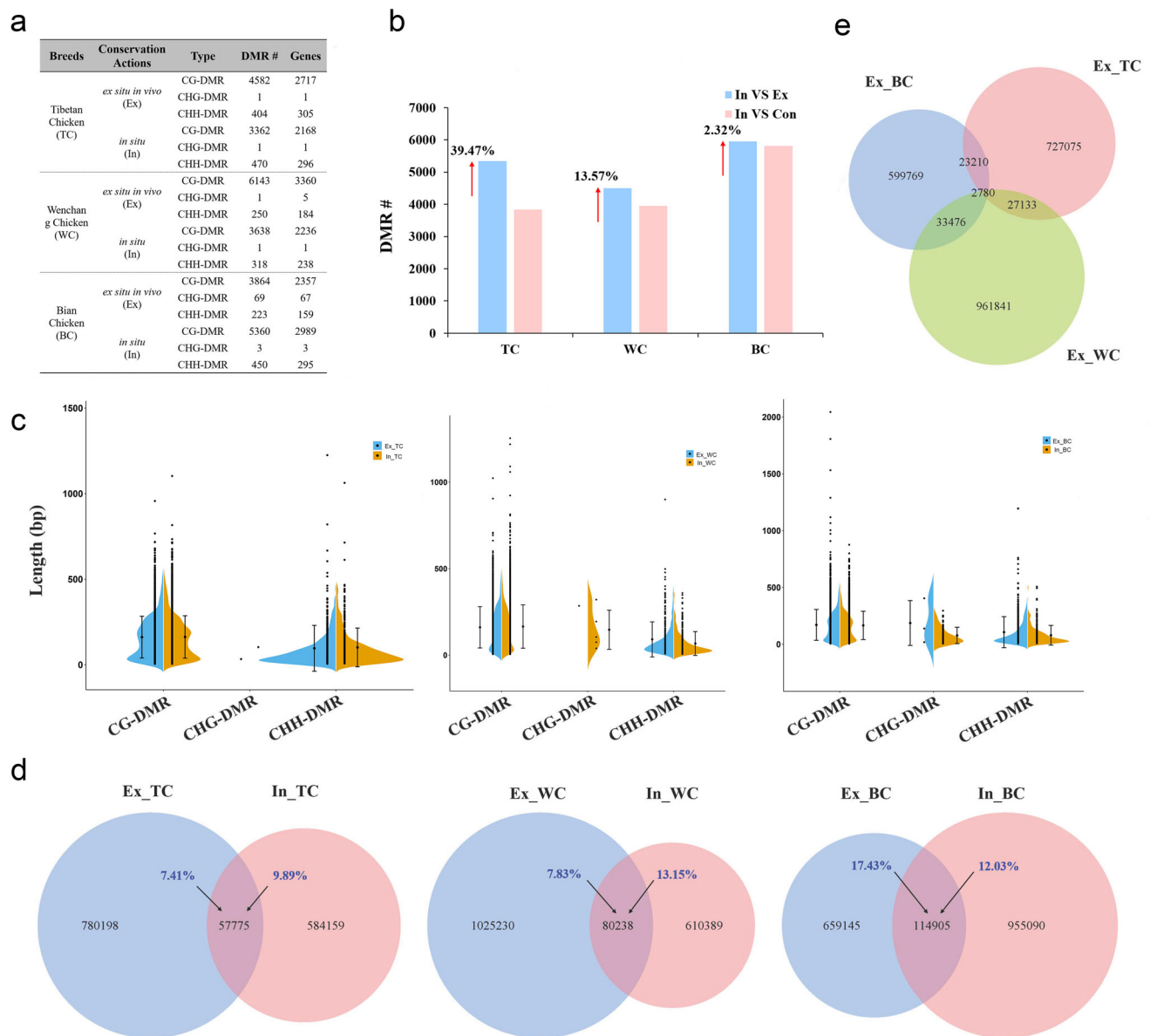
Fig. 4 Differentially methylated regions (DMRs) during in situ (In) and ex situ in vivo (Ex) conservation programmes of chickens. a Number of DMRs detected in different conservation programmes of three indigenous chicken breeds. **b** Change in numbers of DMRs detected between Ex and In of three indigenous chicken breeds. **c** Length of DMRs during different conservation programmes of three indigenous chicken breeds. **d** Venn diagram of the numbers of overlapping bases in DMRs during different conservation programmes of three indigenous chicken breeds. Overlapping areas are presented as percentages. **e** Overlapping bases (bp) in DMRs during Ex conservation of three indigenous chicken breeds.

DMRs that might be associated with local genetic variations, we investigated the percentage of DMRs in F_{ST} selection regions and other regions via whole genome scanning. Overall, a high proportion of DMRs was found in both the F_{ST} and non-F_{ST} regions, during Ex conservation programme of the three chicken breeds (Fig. 5a). Meanwhile, we found that the proportion of DMRs was higher in F_{ST} selection regions than in non-F_{ST} regions in the Ex of TC, whereas approximately the same in group of TC, and Ex and In groups of WC (Fig. 5a). In contrast, the proportion of DMRs was higher in non-F_{ST} regions in Ex of BC. Taken together, these results indicate that the variation patterns, both genetic and DNA methylation, were different during in situ and ex situ in vivo conservation programmes in chickens, which could relate to the features of the environment of origin, such as altitude and temperature, risk status and so on. Remarkably, the high proportion of DMRs in the F_{ST} regions of Ex group of TC suggests a

link between these DMRs and local genetic variations. It is well known that the climatic and geographic characteristics of ex situ in vivo conservation of TCs are markedly different from those of in situ conservation. Therefore, local genetic variations due to environmental differences in the Ex group of TC may influence DNA methylation level, and ultimately lead to changes in gene expression. In our analyses, the DMRs that were found in the F_{ST} regions were considered as "sweep DMRs (s-DMRs)".

In total, 334 s-DMRs comprising 323 s-CG-DMRs and 11 s-CHH-DMRs were identified in the Ex group of TC (Supplementary Data 15). The s-DMRs contained 124 candidate genes. Further examination of these genes identified eight genes, including *HTR2A*, *EYS*, *DST*, *RAP1GDS1*, *HTT*, *HTR7L*, *FKBP7*, and *MYLK*, that have functional associations with calcium ion binding/calcium signalling pathway, suggesting potential roles for these genes in high-altitude adaptation of TC (Fig. 5b and
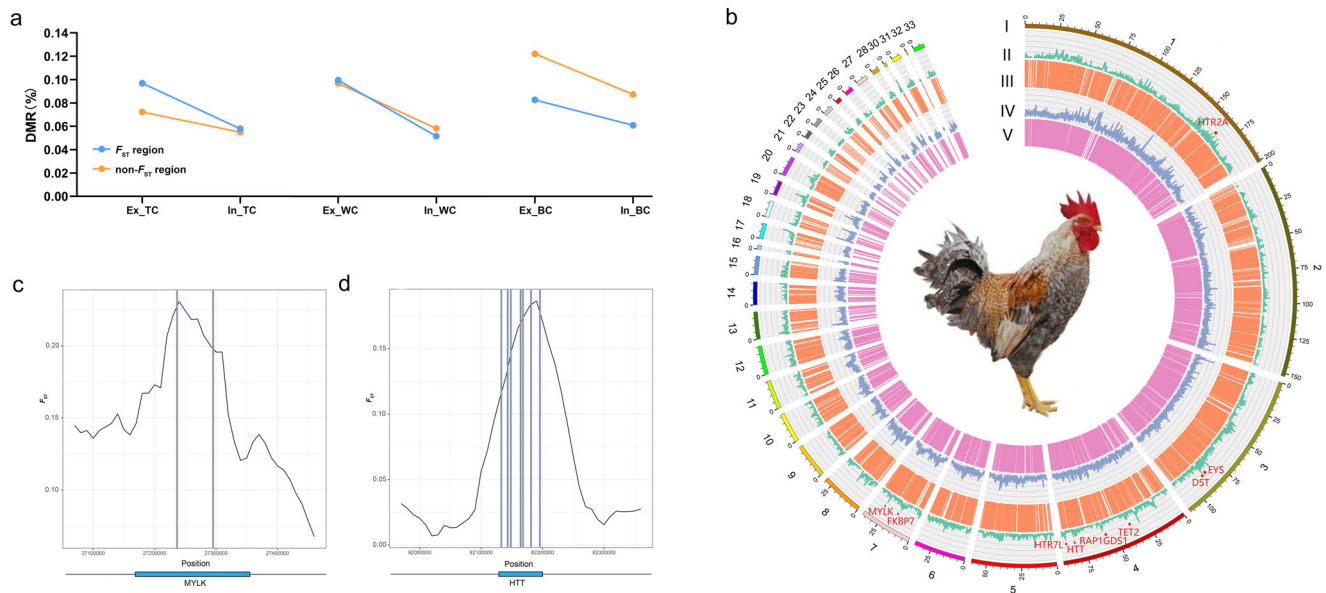
Fig. 5 DNA methylation variations contribute to genomic selection signatures (s-DMRs) during Ex conservation programme of TC. a Percentage of DMRs in F_{ST} selection regions and other regions of indigenous chicken breeds (TC: Tibetan chicken, WC: Wenchang chicken and BC: Bian chicken) during different conservation programmes (In – in situ conservation and Ex – ex situ in vivo conservation). **b** Genome and methylation landscapes during different conservation programmes of TC. From outer to inner circles are as follows: chromosome scheme (I), selective sweeps detected during Ex programme (II), DMRs during Ex programme (III), selective sweeps detected during In programme (IV), and DMRs during In programme (V). Red dots in circle II denote genes located in s-DMRs. **c**, **d** Two prominent genes *MYLK* (**c**) and *HTT* (**d**), which were located in s-DMRs during Ex programme of Tibetan chickens. Grey shadow depicts the location of DMRs.

Supplementary Data 16). Ca^{2+} is one of the most important regulators of pulmonary vascular function. Calcium signalling stimulates the translation of HIF-1, a principal regulator of the transcriptional response to hypoxia[62–65]. One prominent candidate gene, *MYLK*, which encodes a Ca^{2+}/calmodulin dependent enzyme, was identified in s-CG-DMRs (Fig. 5c) and as well as in CDRs between TC and other chickens (Fig. 3a and Table 1). Previous studies have reported that MYLK protein levels are upregulated through the mediation of HIF-1α with concomitant vascular endothelial barrier dysfunction when the cells are exposed to hypoxic stress[66]. *MYLK* upregulation is associated with augmented vasoconstriction, a marker of high-altitude pulmonary oedema[67]. Interestingly, recent studies have suggested that hypoxia upregulates the expression of *MYLK*, and thereby increases the potential to transport oxygen by promoting smooth muscle contraction in TC embryos[68]. We also highlighted two candidate genes, *HTT* and *EYS*, which contained seven DMRs and one DMR in the selection region, respectively (Fig. 5d and Supplementary Fig. 10a). Although the role of *HTT* gene in high-altitude hypoxia adaptation is unknown, previous studies have found an association between *HTT* and InsP$_3$R1-mediated neuronal Ca^{2+} signalling, and have provided an explanation for the derangement of cytosolic Ca^{2+} signalling in Huntington's disease patients and mouse models[69]. Furthermore, *TET2*, which is a key factor in DNA demethylation, and serves as a methyl cytosine dioxygenase, was found in s-CG-DMRs (Supplementary Fig. 10b). Hypoxia has been shown to influence DNA methylation either by transcriptionally activating *TET* or by reducing *TET* activity[70,71]. Recently, *TET* has also been shown to affect hypoxia signalling through its interaction with HIF-1α[72,73]. Overall, these genes may be associated with the high-altitude adaptation in TC.

Conclusions

Conservation of animal genetic resources plays an important role in biodiversity protection. In the present study, we performed whole-genome sequence analysis and WGBS analysis of three indigenous chicken breeds that have been conserved in different conservation programmes. Our results demonstrated that a higher genetic diversity and differentiation occurs during in situ conservation, while reduced genetic diversity occurs during ex situ in vivo conservation programme of three indigenous chicken breeds. DNA methylation variation levels were higher during ex situ in vivo conservation were higher than during in situ conservation. Moreover, a high proportion of DMRs in genomic selection regions was found during ex situ in vivo conservation of Tibetan chickens, suggesting an association between these DMRs and local genetic variations. Additionally, we discovered several candidate genes in selective sweep regions and DMRs, which correlated with calcium signalling pathway, suggesting a potential regulatory mechanism underlying high-altitude adaptation in Tibetan chickens. Collectively, our study provides valuable information about the genetic variations and DNA methylation variations during different conservation programmes of indigenous chicken breeds, contributes to a profound understanding of high-altitude adaptation and will facilitate future endangered species conservation.

Methods

Sampling, DNA sequencing and WGBS. Three Chinese indigenous chicken breeds that were conserved in different conservation programmes (in situ and ex situ in vivo) were used in this study (TC: Tibetan chicken; WC: Wenchang chicken; and BC: Bian chicken). Each breed was divided into the following three groups according to the conservation programmes (Con – cryopreserved samples, as a control, in March 2000, kept in the NCF; In – in situ conservation after approximately 20 years, in March 2019, kept in the NCF; and Ex – ex situ in vivo conservation after ~20 years, in March 2019, kept in the NCGR) (Fig. 1a and Supplementary Table 1). Blood samples were collected from a total of 90 individuals, of which 30 for each breed and 10 for each group in the same breed. Animals used in this study were raised in accordance with the national standard of Laboratory Animal Guidelines for ethical review of animal welfare. All experiment procedures were approved by the Animal Use Committee of Zhejiang Academy of Agricultural Sciences (No. 20-022).

Genome DNA was extracted using a DNA isolation kit (Tiangen, Beijing, China) according to the manufacturer's instructions. For genome sequencing, a

minimum of 0.5 µg of genomic DNA from each sample was used to construct a library with an insert size of ~350 bp. Paired-end (PE) sequencing libraries were constructed according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA) and sequenced using Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA). WGBS libraries were prepared according to the protocol described in a previous report[74]. The DNA samples were fragmented with sonication and subjected to bisulfite conversion. Ultra-high-throughput paired-end sequencing was carried out using Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA) according to the manufacturer instructions. Raw HiSeq sequencing data were processed by Illumina base-calling pipeline (SolexaPipeline-1.0).

Reads mapping and variant calling. To obtain high-quality data, we implemented quality-control procedures to remove the following types of reads: (1) read pairs containing adapters, (2) reads with ≥ 10% unidentified nucleotides (N), and (3) >50% of the read bases with a Phred quality score (i.e., Q-score) <10. A total of 985.74 Gb of high-quality paired-end reads were generated with an average quality of 96.98% for Q20 and 93.66% for Q30. The filtered high-quality data were mapped to the chicken reference genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/315/GCF_000002315.5_GRCg6a/GCF_000002315.5_GRCg6a_genomic.fna.gz) using Burrows Wheeler Aligner (v0.7.8) with default parameters[75]. Mapping results were then converted into the BAM format and sorted using SAMtools (v1.2). On an average, 98.44% of the reads were mapped, resulting in a final average sequencing coverage of 9× per individual. The mapped reads were further filtered by removing duplicate reads using the Picard MarkDuplicates (http://sourceforge.net/projects/picard/). After alignment, we performed SNP calling using GATK (v3.7), and the output was further filtered using VCFtools[76,77]. Moreover, to exclude SNP callings errors, variant sites with QD < 2.0, QUAL < 30, MQ < 40, FS > 60.0 were discarded. The process of indel calling was the same as described for the SNPs. Heterozygosity and homozygosity for each sample were calculated by the total number of callable sites across the whole genome. Finally, SNP variations were annotated using the SnpEff software (v4.3)[78]. The whole-genome sequencing data and distribution of SNPs and InDels in the three Chinese indigenous chicken breeds are shown in Supplementary Data 1–2.

Phylogenetic tree, population structure, principal component analysis and Admixture. To determine the phylogenetic relationships, we constructed a neighbour-joining phylogenetic tree based on MEGA X software (v10.1.5) with Kimura 2-parameter model and 1000 bootstrap replicates[79]. Population structure was then inferred using the software ADMIXTURE (v1.3.0) to quantify genome-wide admixture between three Chinese indigenous chicken breeds populations[22]. The number of genetic clusters K ranged from 2 to 9 with 10,000 maximum iterations. In order to understand the relationships between different geographic populations and same breed with different conservation actions, principal component analysis (PCA) of the 90 samples was performed based on all SNPs using EIGENSOFT software (v7.2.1)[21].

Genetic diversity, differentiation and linkage disequilibrium. The inbreeding coefficient (F), observed heterozygosity (Ho), expected heterozygosity (He) and proportion of polymorphic markers (Pn) were calculated using PLINK (v1.07)[80]. Effective population size (Ne) for each breed was calculated using default parameters in N$_e$ESTIMATOR software (v1.3)[81]. The nucleotide diversity (π) and Tajima's D (θ_w) were calculated based on the list of high-quality SNPs using VCFtools (v0.1.14) with a sliding window approach (100-kb windows with 10-kb steps)[82]. Population differentiation was measured by pairwise F_{ST} using the unbiased estimator of Weir and Cockerham (1984) with default parameters[83]. To evaluate linkage disequilibrium (LD) for different groups, we calculated the correlation coefficient (r^2) between alleles at two separate SNP loci using Haploview (v4.2) with parameters "-maxdistance 500 -dprime -minMAF 0.05 -hwcutoff 0.001"[84].

Genome scanning of selective signatures. To identify genome-wide selective sweeps associated with high-altitude adaptation. We screened the chicken genomes with a sliding window approach (100-kb windows with 10-kb step length), and estimated F_{ST} and π ln ratio values for each window using VCFtools (v0.1.14) and in-house scripts for comparison between TC and BC and between TC and WC[77]. The significance threshold was set to the top 5% for F_{ST} and π ln ratio. We then considered the windows with top 5% simultaneously as candidate selective regions under strong selective sweep and subsequently examined for potential candidate genes. Finally, we identified 228 candidate divergent regions (CDRs) ($F_{ST} \geq 0.35$) comprising 1022 candidate genes between TC and BC (Fig. 3a and Supplementary Data 4) and 168 CDRs ($F_{ST} \geq 0.28$) comprising 700 candidate genes between TC and WC (Fig. 3a, Supplementary Data 5). Functional enrichment terms, including Gene Ontology (GO)[85] and KEGG pathways[86] were retrieved for these genes. GO terms and pathways with enrichment q-values < 0.05 were considered to be significantly enriched.

WGBS analysis. Adapters and low-quality bases in the WGBS reads were removed using in-house Perl scripts. The sequence quality was then verified with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). A total of 357.20 Gb of clean reads was generated with an average quality of 91.93% for Q30. Reads that passed quality control were mapped to the chicken reference genome by Bismark (v0.21.0)[87]. After filtering the duplicate reads, the methylation information for each cytosine site was extracted. Methylation states were evaluated based on the binomial distribution test principle followed by coverage ≥ 4X and false discovery rate (FDR < 0.05) correction. The methylation level (CG, CHG and CHH) was defined as the ratio between the number of reads of a methylated C to the total number of reads of that site[88].

DMR detection and functional enrichment analysis. Genome coverage of CG, CHG and CHH sites under different sequencing depths, and distribution of clean reads in different CG density regions were analysed. The genome coverage of the CG, CHG and CHH sites in every chromosome and different genome components including promoter, gene body and downstream was also analysed. DMRs were detected according to the cytosine (C) mapping of uniquely mapped reads into the reference genome. MOABS (v1.3.7.7) was used to screen DMRs, and the coverage depth was required to be no less than 10X, with at least three different methylation sites[89]. DNA methylation levels of different groups were compared pairwise through Fisher's exact test and the p-values were adjusted for multiple comparisons using the Benjamini-Hochberg method. CG-DMR candidate regions, CHG-DMR candidate regions, and CHH-DMR candidate regions were required to have an average methylation level differences of >0.3, >0.2 and >0.2, respectively, between the corresponding populations. Finally, the regions with an FDR less than 0.05 were identified as DMRs. We identified DMRs between Con and Ex, Con and In groups and between Ex and In groups of the three indigenous chicken breeds. To explore the function of the genes containing DMRs, GO enrichment and KEGG pathway analyses were conducted for differentially methylated genes using Blast2GO (v5.2.5)[90]. GO terms and pathways with enrichment q-values <0.05 were considered to be significantly enriched.

Statistics and reproducibility. Details about experimental design and statistics used in different data analyses are given in the respective sections of results and methods. Unless stated otherwise, statistical analysis was performed using base and nlme in R(4.2.0). Details allowing the reproducibility of all analyses are provided in the methods section. For genome resequencing, we used 90 chickens of different conservation programmes of three Chinese indigenous chicken breeds. For whole-genome bisulphite sequencing, we used at least three chickens samples for pool from nine subgroup of three Chinese indigenous chicken breeds.

Data availability

Code availability

References

1. Eda, M. Origin of the domestic chicken from modern biological and zooarchaeological approaches. *Anim. Front.* **11**, 52–61 (2021).
2. FAO. *Status And Trends Of Animal Genetic Resources* (Commission on genetic resources for food agriculture, Rome, 2019).
3. Chen, G., Wang, K., Wang, J., Ding, C. & Yang, N. *Poultry Genetic Resources in China* (Shanghai Scientific and Technological Press, Shanghai, 2004).
4. Wang, M. S. et al. Genomic analyses reveal potential independent adaptation to high altitude in tibetan chickens. *Mol. Biol. Evol.* **32**, 1880–1889 (2015).
5. Zhang, Q. et al. 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 (2016).

6. Zhang, H., Wang, X. T., Chamba, Y., Ling, Y. & Wu, C. X. Influences of hypoxia on hatching performance in chickens with different genetic adaptation to high altitude. *Poult. Sci.* **87**, 2112–2116 (2008).

7. Jia, C. L., He, L. J., Li, P. C., Liu, H. Y. & Wei, Z. H. Effect of egg composition and oxidoreductase on adaptation of Tibetan chicken to high altitude. *Poult. Sci.* **95**, 1660–1665 (2016).

8. Hillier, L. W. et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695–716 (2004).

9. Rubin, C. J. et al. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–591 (2010).

10. Wang, M. S. et al. An evolutionary genomic perspective on the breeding of dwarf chickens. *Mol. Biol. Evol.* **34**, 3081–3088 (2017).

11. Shen, Y. et al. DNA methylation footprints during soybean domestication and improvement. *Genome Biol.* **19**, 128 (2018).

12. Springer, N. M. & Schmitz, R. J. Exploiting induced and natural epigenetic variation for crop improvement. *Nat. Rev. Genet* **18**, 563–575 (2017).

13. Xu, J. et al. Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. *Genome Biol.* **20**, 243 (2019).

14. Hoglund, A. et al. The methylation landscape and its role in domestication and gene regulation in the chicken. *Nat. Ecol. Evol.* **4**, 1713–1724 (2020).

15. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).

16. Zhang, Y., Gou, W., Ma, J. & Zhang, H. Genome methylation and regulatory functions for hypoxic adaptation in Tibetan chicken embryos. *PeerJ* **5**, e3891 (2017).

17. Richards, E. J. Inherited epigenetic variation–revisiting soft inheritance. *Nat. Rev. Genet.* **7**, 395–401 (2006).

18. Kawakatsu, T. et al. Epigenomic diversity in a global collection of arabidopsis thaliana accessions. *Cell* **166**, 492–505 (2016).

19. Han, W. et al. Genome-wide analysis of the role of DNA methylation in inbreeding depression of reproduction in Langshan chicken. *Genomics* **112**, 2677–2687 (2020).

20. Wang, M. et al. Whole-genome methylation analysis reveals epigenetic variation in cloned and donor pigs. *Front. Genet.* **11**, 23 (2020).

21. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

22. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

23. Zhang, M. et al. Genomic diversity dynamics in conserved chicken populations are revealed by genome-wide SNPs. *BMC Genomics.* **19**, 598 (2018).

24. Ellegren, H. & Galtier, N. Determinants of genetic diversity. *Nat. Rev. Genet.* **17**, 422–433 (2016).

25. Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K. & Allendorf, F. W. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv. Genet.* **11**, 355–373 (2010).

26. Palstra, F. P. & Fraser, D. J. Effective/census population size ratio estimation: a compendium and appraisal. *Ecol. Evol.* **2**, 2357–2365 (2012).

27. Mussmann, S. M. et al. Genetic rescue, the greater prairie chicken and the problem of conservation reliance in the Anthropocene. *R. Soc. Open Sci.* **4**, 160736 (2017).

28. Banks, S. C. et al. How does ecological disturbance influence genetic diversity? *Trends Ecol. Evol.* **28**, 670–679 (2013).

29. Bjornstad, O. N. & Grenfell, B. T. Noisy clockwork: time series analysis of population fluctuations in animals. *Science* **293**, 638–643 (2001).

30. Zhou, Z. et al. An intercross population study reveals genes associated with body size and plumage color in ducks. *Nat. Commun.* **9**, 2648 (2018).

31. Forcada, J. & Hoffman, J. I. Climate change selects for heterozygosity in a declining fur seal population. *Nature* **511**, 462–465 (2014).

32. Ding, D. et al. Genetic variation in PTPN1 contributes to metabolic adaptation to high-altitude hypoxia in Tibetan migratory locusts. *Nat. Commun.* **9**, 4991 (2018).

33. Zhou, X. et al. Population genomics reveals low genetic diversity and adaptation to hypoxia in snub-nosed monkeys. *Mol. Biol. Evol.* **33**, 2670–2681 (2016).

34. Hu, X. J. et al. The genome landscape of tibetan sheep reveals adaptive introgression from argali and the history of early human settlements on the qinghai-tibetan plateau. *Mol. Biol. Evol.* **36**, 283–303 (2019).

35. Wang, G. D. et al. Selection and environmental adaptation along a path to speciation in the Tibetan frog Nanorana parkeri. *Proc. Natl Acad. Sci. USA* **115**, E5056–E5065 (2018).

36. Xin, J. et al. Chromatin accessibility landscape and regulatory network of high-altitude hypoxia adaptation. *Nat. Commun.* **11**, 4928 (2020).

37. Wu, D. D. et al. Pervasive introgression facilitated domestication and adaptation in the Bos species complex. *Nat. Ecol. Evol.* **2**, 1139–1145 (2018).

38. Wang, M. S. et al. Ancient Hybridization with an unknown population facilitated high-altitude adaptation of Canids. *Mol. Biol. Evol.* **37**, 2616–2629 (2020).

39. Cheng, J. B. & Russell, D. W. Mammalian wax biosynthesis. I. Identification of two fatty acyl-Coenzyme A reductases with different substrate specificities and tissue distributions. *J. Biol. Chem.* **279**, 37789–37797 (2004).

40. Stroud, D. A. et al. Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* **538**, 123–126 (2016).

41. Vargas, J. D. et al. Stromal cell-derived receptor 2 and cytochrome b561 are functional ferric reductases. *Biochim Biophys. Acta* **1651**, 116–123 (2003).

42. Rattner, A., Smallwood, P. M. & Nathans, J. Identification and characterization of all-trans-retinol dehydrogenase from photoreceptor outer segments, the visual cycle enzyme that reduces all-trans-retinal to all-trans-retinol. *J. Biol. Chem.* **275**, 11034–11043 (2000).

43. Amengual, J. et al. A mitochondrial enzyme degrades carotenoids and protects against oxidative stress. *FASEB J.* **25**, 948–959 (2011).

44. Tanji, C. et al. A-kinase anchoring protein AKAP220 binds to glycogen synthase kinase-3beta (GSK-3beta) and mediates protein kinase A-dependent inhibition of GSK-3beta. *J. Biol. Chem.* **277**, 36955–36961 (2002).

45. Manabe, R. et al. Transcriptome-based systematic identification of extracellular matrix proteins. *Proc. Natl Acad. Sci. USA* **105**, 12849–12854 (2008).

46. Zou, X. et al. NELL-1 binds to APR3 affecting human osteoblast proliferation and differentiation. *FEBS Lett.* **585**, 2410–2418 (2011).

47. Luan, X. et al. Crystal structure of human RANKL complexed with its decoy receptor osteoprotegerin. *J. Immunol.* **189**, 245–252 (2012).

48. Yoshimura, S-i., Gerondopoulo, A., Linford, A., Rigden, D. J. & Barr, F. A. Family-wide characterization of the DENN domain Rab GDP-GTP exchange factors. *J. Cell Biol.* **191**, 367–381 (2010).

49. Cox, L. J. et al. Sperm phospholipase Czeta from humans and cynomolgus monkeys triggers Ca^{2+} oscillations, activation and development of mouse oocytes. *Reproduction* **124**, 611–623 (2002).

50. Shimizu, S. et al. Ca^{2+}-calmodulin-dependent myosin light chain kinase is essential for activation of TRPC5 channels expressed in HEK293 cells. *J. Physiol.* **570**, 219–235 (2006).

51. Foxler, D. E. et al. The LIMD1 protein bridges an association between the prolyl hydroxylases and VHL to repress HIF-1 activity. *Nat. Cell Biol.* **14**, 201–208 (2012).

52. Sun, F. et al. Crystal structure of mitochondrial respiratory membrane protein complex II. *Cell* **121**, 1043–1057 (2005).

53. Dubin, M. J. et al. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *Elife* **4**, e05255–e05255 (2015).

54. Johannes, F., Colot, V. & Jansen, R. C. Epigenome dynamics: a quantitative genetics perspective. *Nat. Rev. Genet.* **9**, 883–890 (2008).

55. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476 (2008).

56. Schmitz, R. J. et al. Patterns of population epigenomic diversity. *Nature* **495**, 193–198 (2013).

57. Hauben, M. et al. Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proc. Natl Acad. Sci. USA* **106**, 20109 (2009).

58. Reinders, J. et al. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes Dev.* **23**, 939–950 (2009).

59. Watson, J. A., Watson, C. J., McCann, A. & Baugh, J. Epigenetics, the epicenter of the hypoxic response. *Epigenetics* **5**, 293–296 (2010).

60. Guerrero-Bosagna, C. From epigenotype to new genotypes: relevance of epigenetic mechanisms in the emergence of genomic evolutionary novelty. *Semin. Cell Dev. Biol.* **97**, 86–92 (2020).

61. Furey, T. S. & Sethupathy, P. Genetics. Genetics driving epigenetics. *Science* **342**, 705–706 (2013).

62. Shimoda, L. A. & Undem, C. Interactions between calcium and reactive oxygen species in pulmonary arterial smooth muscle responses to hypoxia. *Respir. Physiol. Neurobiol.* **174**, 221–229 (2010).

63. Rodman, D. M. et al. Low-voltage-activated (T-type) calcium channels control proliferation of human pulmonary artery myocytes. *Circ. Res.* **96**, 864–872 (2005).

64. Hui, A. S., Bauer, A. L., Striet, J. B., Schnell, P. O. & Czyzyk-Krzeska, M. F. Calcium signaling stimulates translation of HIF-alpha during hypoxia. *FASEB J.* **20**, 466–475 (2006).

65. Mottet, D. et al. Role of ERK and calcium in the hypoxia-induced activation of HIF-1. *J. Cell Physiol.* **194**, 30–44 (2003).

66. Qi, H. et al. Involvement of HIF-1α in MLCK-dependent endothelial barrier dysfunction in hypoxia. *Cell Physiol. Biochem.* **27**, 251–262 (2011).

67. Pandey, P., Mohammad, G., Singh, Y. & Qadar Pasha, M. A. ROCK2 and MYLK variants under hypobaric hypoxic environment of high altitude associate with high altitude pulmonary edema and adaptation. *Appl. Clin. Genet.* **8**, 257–267 (2015).

68. Zhang, Y., Gou, W., Zhang, Y., Zhang, H. & Wu, C. Insights into hypoxic adaptation in Tibetan chicken embryos from comparative proteomics. *Comp. Biochem Physiol. Part D. Genomics Proteomics* **31**, 100602 (2019).

69. Tang, T. S. et al. Huntingtin and huntingtin-associated protein 1 influence neuronal calcium signaling mediated by inositol-(1,4,5) triphosphate receptor type 1. *Neuron* **39**, 227–239 (2003).

70. Thienpont, B. et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**, 63–68 (2016).

71. Mariani, C. J. et al. TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. *Cell Rep.* **7**, 1343–1352 (2014).

72. Tsai, Y. P. et al. TET1 regulates hypoxia-induced epithelial-mesenchymal transition by acting as a co-activator. *Genome Biol.* **15**, 513 (2014).

73. Fan, S. et al. TET is targeted for proteasomal degradation by the PHD-pVHL pathway to reduce DNA hydroxymethylation. *J. Biol. Chem.* **295**, 16299–16313 (2020).

74. Urich, M. A., Nery, J. R., Lister, R., Schmitz, R. J. & Ecker, J. R. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nat. Protoc.* **10**, 475–483 (2015).

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

76. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

77. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

78. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).

79. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

80. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

81. Do, C. et al. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol. Ecol. Resour.* **14**, 209–214 (2014).

82. Decker, J. E. et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet.* **10**, e1004254 (2014).

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

84. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).

85. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, D277–D280 (2004).

86. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).

87. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

88. Schultz, M. D., Schmitz, R. J. & Ecker, J. R. 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends Genet.* **28**, 583–585 (2012).

89. Sun, D. et al. MOABS: model based analysis of bisulfite sequencing data. *Genome Biol.* **15**, R38 (2014).

90. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

Author contributions
T.Z., L.Z.L. and W.H. planned the project. L.Z.L., W.H., T.Z., J.M.Y., P.S.F. and Y.T. designed the research. T.Z., J.M.Y., T.T.G., Y.L.L., G.H.L., L.Q., L.C., L.H.G., Q.X., Q.Y.W., Y.Q.C., P.F.L., H.Y.Z., G.Q.L., L.J.L., C.H.Z., Z.R.T., J.D.S. and W.H. prepared the research materials. T.Z., P.S.F., F.R.H., Y.T.W., Y.H.X., Y.T. and W.W.X. performed the experiments for sequencing and the data analyses. T.Z. drafted the manuscript. T.Z., L.Z.L. and W.H. revised the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03907-7.

Correspondence and requests for materials should be addressed to Wei Han or Lizhi Lu.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: George Inglis. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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