Conservation of genomic diversity and breed-specific characteristics of domestic chickens in China

Mengmeng Zhang[1], Yijun Liu[1], shiwei Wang[1], Han Zhang[1], mengxia Sun[1], junyan Wang[1], and Keliang Wu[2]

[1]China Agricultural University
[2]Affiliation not available

May 5, 2020

Abstract

Effective conservation and utilization of farm animals are fundamental for realizing sustainable increases in food production. In situ and ex situ conservation are the two main strategies currently used to protect domestic chicken in China. However, genomic diversity and population structure have not been compared in these conserved populations. One potential risk is that the use of genome-wide SNPs to optimize genomic diversity might not preserve particular alleles that are associated with breed-specific characteristics. Here, 361 individuals from three Chinese domestic chicken breeds were collected from populations conserved in situ and ex situ, and genotyped using GBS (genotyping-by-sequencing). We estimated the genomic diversity, analyzed population structures, and found that the small ex situ conserved populations that have been maintained in controlled environments retained less genetic diversity than the in situ's. In addition, genetic differentiation was detected between in situ and ex situ conserved populations within a single breed. We next analyzed selective signatures (FST, Pi, and XPEHH) to examine the genetic mechanisms underlying differentiation between in situ and ex situ conserved populations. We concluded that differentiation might be caused by genetic drift, or the differences were due to variants from the original populations. Finally, based on sequencing data obtained from the ex situ conserved populations, we used Di and Pi to identify "genomic conservation units" for breed-specific characteristics. Loci associated with the "genomic conservation unit" could be used to preserve breed-specific characteristics in the conservation program.

Introduction

Because of its long history of animal husbandry and diversified geographical conditions, China has a rich diversity of domestic chicken breeds. To date 107 Chinese breeds have been described (Rescources, 2011), among which are some with striking appearance and valuable traits. Genetic diversity provides the raw material for breed improvement and for the adaptation of livestock populations to changing environments and market demands. However, genetic diversity is at risk for many species throughout the world. Among domesticated avian species, chickens have by far the highest number of breeds at risk . In China alone 21 breeds are at risk, representing 1/5 of the total number of domestic chicken breeds (Rescources, 2011). Effective conservation and use of farm animals are necessary to obtain sustainable increases in food production. Conservation plans are commonly classified into three categories: *in situ* conservation; *ex situ in vivo*conservation; and *ex situ in vitro* conservation.*In vivo* methods are primarily used in China for the management of animal genetic resources, including both *in situ* and *ex situ*conservation. *In situ* conservation can best be described as the sustainable breeding of an endangered livestock breed in the normal adaptive production environment, or as close to it as practically possible, to conserve genetic diversity over a long period. *Ex situ* conservation is the preservation of endangered livestock outside of normal production systems . In China, two national gene banks (National Chicken Genetic Resources in Jiangsu and Zhejiang) and 23 national conservation farms have been established. However, few studies have compared the actual efficacy

of *in situ* and *ex situ*efforts to conserve chickens, although the FAO has recommended that livestock breed conservation status should be monitored regularly .

A comprehensive knowledge of genetic diversity within and between breed populations is required to manage animal genetic resources (Groeneveld et al., 2010). DNA markers are the most reliable molecular tools for the assessment of genetic diversity (Liu and Cordes, 2004b). RFLPs (Thurston et al., 2002), mtDNA (Avise et al., 1987; Avise et al., 1986; Harrison, 1989; Kocher et al., 1989; Zhang and Hewitt, 1996), RAPD (Ali et al., 2004; Dodgson et al., 1997; Koh et al., 1998; Levin et al., 1993), AFLPs (Parsons and Shaw, 2001; Savelkoul et al., 1999), Y-chromosome markers (Bruford et al., 2003; Zeder et al., 2006), VNTRs (Zane et al., 2002), and SNPs (Andersson and Georges, 2004; Liu and Cordes, 2004a; McMahon et al., 2014; Morin et al., 2004; Vignal et al., 2002) have been the most widely used marker systems. In addition, conservation programs have been based on pedigree information. The development of high-throughput genotyping techniques has made it possible to obtain large numbers of genomic markers that can be used to correct and reconstruct pedigrees. Genome-wide marker data is also regarded as a useful tool for the maintenance of genetic diversity (de Cara et al., 2011). However, conservation programs designed using genome-wide SNPs alone risk losing the genetic variations associated with some traits, resulting in reduced performance. In particular, domestic chicken breeds have specific performance traits that constitute an important genetic resource. The maintenance of genetic diversity must therefore include the preservation of these valuable phenotype variations. Because conservation populations are usually small, gene drift can also occur easily, and alleles that contribute to special traits might be lost. It is thus crucial to monitor and maintain the genetic diversity specifically responsible for breed characteristics as well as the general genetic diversity across the genome.

Here, we integrated genomic data from *in situ* and *ex situ*conserved chicken breeds in China, compared breeds to determine genomic diversity, and then used whole-genome SNP markers to assess the efficacy of ongoing *in situ* and *ex situ* conservation efforts. The data were examined to detect genomic signatures resulting from genetic differentiation between breeds managed using the two conservation practices, and selective signature analysis was also used to identify "genomic conservation units" to study the molecular breed-specific characteristics conservation. Our results provide insights into the genomic effects of ongoing conservation efforts, and establish a foundation for optimizing conservation programs for *in situ* and*ex situ* populations of Chinese domestic chickens.

Material and Methods

Ethics statement

All experimental procedures were approved by the Animal Welfare Committee of China Agricultural University (Approval Number: XK257).

Populations

Individuals from three Chinese indigenous breeds were selected from*in situ* and *ex situ* conserved populations (361 total, consisting of 120 Beijing You chickens, 120 Baier Yellow chickens, and 121 Langshan chickens). Of these, 270 chickens (representing three successive conserved generations from an *ex situ* conserved population) had been used in our previous study (Zhang et al., 2018). The three breeds originate from three different regions within China (see sampling information in Table 1 and Figure 1). The mating systems used in the two conservation systems differ (R:R for *in situ* and R:F for *ex situ* (Zhang et al., 2018)). In addition, the *in situ* conserved chickens have been subjected to conservation for a longer time, and have a larger population size relative to the *ex situ* chickens.

Blood samples were collected from the wing vein and stored at -20°C. We used the Qiagen DNeasy Tissue kit (Qiagen, Germany) to extract genomic DNA from blood, and verified the integrity and purity of DNA by agarose gel electrophoresis and optical density (A_{260}/A_{280} ratio). 3 μg of high-quality DNA was used to construct sequencing libraries for each sample.

Genotyping

To obtain genomic data consistent with our previous study, we processed the samples as described in Zhang et al.(Zhang et al., 2018). After double digestion with restriction enzymes *MseI* and *HaeIII* , all DNA samples were genotyped by high-throughput sequencing using an Illumina HiSeq 4000 sequencer (Illumina, San Diego, CA, USA) and the protocol provided by the manufacturer. To improve mapping, in-house scripts were used to remove low-quality reads from the data set. Reads were excluded if they (i) contained adapter sequences, (ii) if [?]10% of nucleotides were unidentified (N), or (iii) if¿50% bases had low phred quality scores (¡5). The remaining high quality paired-end reads were mapped to the *Gallus gallus* 5.0 reference genome using the Burrows-Wheeler Alignment tool (BWA) (v0.7.8) (Li and Durbin, 2009) with default parameters. PCR duplicates were removed using SAMtools rmdup (v1.3.1) (Li et al., 2009).

The aligned BAM files for the 361 chickens were used to detect variants at the population scale using the SAMtools suite (v1.3.1), including BCFtools, with parameters as described in our previous study (Zhang et al., 2018). Single nucleotide variants (SNVs) within 5 bp of an indel were removed. SNPs and INDELs were annotated with ANNOVAR v2013– 08-23 (ANNOVAR, RRID:SCR 012821) (Wang et al., 2010), using gene annotations obtained from the Ensembl database . In the annotation step, SNPs and INDELs were classified into eight categories based on genomic locations, including exonic regions (synonymous, nonsynonymous, stop gain, and stop loss), splicing sites, intronic regions, 5' and 3' UTRs, upstream and downstream regions, and intergenic regions. The dbSNP database was used to identify novel genetic variations.

SNPs located on non-chromosomes were removed. Data were also excluded if they met any of the following four criteria: (1) individuals with missing genotype data for more than 5% of the typed SNPs (call rate [?] 0.95). (2) variants with missing call rates [?] 0.01. (3) SNPs with very low minor allele frequencies (MAF [?] 0.01). (4) SNPs with frequencies that deviate significantly from Hardy-Weinberg Equilibrium (P-value > 10e-6). Removal of low-quality SNPs helped to avoid false-positives and also enhanced the ability to identify loci associated with traits and estimate effective genomic diversity.

Population structure analysis The neighbor-joining (NJ) tree was constructed using the neighbor-joining method in MEGA v7.0 (Kumar et al., 2016) and was visualized in FigTree v1.4.4 (Rambaut, 2018). Population stratification was analyzed by complete linkage clustering of individuals using genome-wide SNP data in PLINK (Purcell et al., 2007). A principal component analysis (PCA) (Price et al., 2006) was conducted using PLINK, and scatter plots were generated using R v3.5.3 . Population structure was analyzed using ADMIXTURE v1.3.0 (Alexander et al., 2009), which applies a likelihood model-to large whole-genome SNP genotype datasets. The number of populations (K) varied from K = 2 to 9 to obtain the maximum likelihood estimates for inference of population structure. Cross-validation was performed to provide a low cross-validation error, which made the optimal K value more apparent. The parameter standard errors were estimated using 1000 bootstrap replicates. The cross-validation plot was generated using R v3.5.3 .

Genomic diversity assessment within populations Allelic richness (Ar), proportion of polymorphic markers (Pn), expected heterozygosity (He), and observed heterozygosity (Ho) were used to investigate genome-wide genomic diversity within 12 subpopulations. Ar was calculated using ADZE v.1.0 (Szpiech et al., 2008). Pn, He, and Ho were calculated using PLINK v1.9 (Purcell et al., 2007).

Evaluation of inbreeding coefficient (F) Two metrics were used to estimate levels of inbreeding in the conserved chicken populations.

The F_{ES} inbreeding coefficient is based on the mating system. The relative change in average inbreeding ([?]F) was obtained by linear regression of the average annual inbreeding coefficient over time, Ft = 1-(1-[?]F)^{t}, where t represents the generation. The increment of hypothetical inbreeding ([?]F) is different for different conservation retention modes. For random mating, random selection$F = \frac{1}{8Nm} + \frac{1}{8Nf}$, and for random mating within families, $F = \frac{3}{32Nm} + \frac{1}{32Nf}$, where Nf and Nm represents number of dams and sires, respectively.

The F_{ROH} inbreeding coefficient is based on runs of homozygosity (ROH). The F_{ROH} statistic, introduced by McQuillan et al. (McQuillan et al., 2008), was calculated as follows: $F_{ROH} = \mathrm{L_{ROH}}/\mathrm{L_{AUT}}$, where $\mathrm{L_{ROH}}$ is

the total length of all ROH in the genome of an individual, and L_{AUT} is the specific length of the autosomal genome covered by SNPs.

Calculation of nucleotide diversity The nucleotide diversity (π) for each population was calculated using VCFtools v0.1.14 (Danecek et al., 2011), based on whole genome SNPs.

Linkage disequilibrium decay Genome-wide LD was evaluated between *in situ* and *ex situ* groups. The average LD of a pair of SNPs in a 300 kb sliding window was estimated using Haploview (Barrett et al., 2004), and the LD decay curves were generated using R v.3.5.3 and Adobe Illustrator CC 2018.

Estimation of population differentiation using F_{ST}The fixation index (F_{ST}), a measure of population differentiation and population structure (Weir and Cockerham, 1984), was estimated using VCFtools v0.1.14 (Danecek et al., 2011) with a 100 kb window and 10 kb step size.

Effective population size Different methods can be used to compute the effective population size (Ne). When based on the number of parents, Ne was calculated with the formula Ne= $\frac{4NmNf}{Nm+Nf}$, where Nf and Nm represent number of dams and sires, respectively (Groeneveld et al., 2009). Ne can also be based on the rate of inbreeding using the formula Ne= $\frac{1}{2*F}$(Groeneveld et al., 2009). Here, we used NeEstimator v.2.01(Do et al., 2014a) to implement the linkage disequilibrium (LD) approach of Waples and Do (Waples and Do, 2008) to estimate effective population size. Ne estimates for each subpopulation were calculated as the average of the estimates for macrochromosomes (gga1-gga5) (Axelsson et al., 2005).

Runs of homozygosity (ROH) To investigate recent inbreeding and the distribution of homozygosity, we identified ROH based on the autosomal SNPs using PLINK v1.9 (Purcell et al., 2007). The analysis was conducted using the default parameter *-homozygosity* , and the following criteria were also selected: (1) a sliding window of 50 SNPs across the genome; (2) one heterozygous and five missing calls were allowed per window to account for genotyping error; (3) the minimum number of consecutive SNPs included in a run of homozygosity was set to 50 and the minimum length for a run was set to 100 kb; (4) the required minimum SNP density to define a run was 1 SNP per 50 kb; (5) the maximum distance between two consecutive SNPs in a run was 1000 kb (Zhang et al., 2018).

Differences in genome-wide homozygosity between *in situ* and*ex situ* populations were tested for statistical significance with three measures: numbers of runs of homozygosity (NSEG); total length of runs (KB); average length of runs (KBAVG).

Adaptation analysis To analyze the underlying genetic mechanisms of adaptation between the *in situ* and *ex situ*conserved populations, we employed multiple statistical tests to identify genomic regions harboring footprints of positive selection between the groups. We used F_{ST} (Akey et al., 2002; Holsinger and Weir, 2009; WRIGHT, 1949), Pi (nucleotide diversity) (Nei and Li, 1979; Wang et al., 2016), and XP-EHH (cross-population extended haplotype homozygosity) (Sabeti et al., 2007). A sliding window approach (100 kb windows sliding in 10 kb steps) was applied to quantify the polymorphism levels, using pairwise nucleotide variation as a measure of variability ($\vartheta\pi$) and genetic differentiation (F_{ST}) between populations. Genome signatures with significantly high F_{ST} values corresponding to the top 5% of values), and $\vartheta\pi$ ratios in the top 5% of values ($\vartheta\pi$,*in situ* /$\vartheta\pi$, *ex situ*) were classified as extensively diversified. XP-EHH scores were calculated with Selscan (Szpiech and Hernandez, 2014) with default parameters to compare whole genome SNPs in all three chicken breeds between *in situ* and *ex situ*conserved populations.The scores for each SNP were then frequency-normalized over all chromosomes using the script norm, provided with Selscan.

Breed characteristics Variants and genes that underlie phenotypic changes in the domestic chicken likely evolved rapidly after domestication. Based on the genomic variation data obtained in this study, we identified regions that may have evolved rapidly in domestic chickens, and thus might have contributed to trait differences amongst breeds. Comparing these regions with data from the Chicken QTL database, we obtained candidate genomic regions, identified as "genomic conservation units", and used them as markers for breed-specific characteristics.

Typically, regions or loci that have evolved rapidly and have experienced selection exhibit specific signatures

of variation, including high population differentiation, significantly reduced nucleotide diversity levels, and long-range haplotype homozygosity (Sabeti et al., 2006). To detect such regions and identify the "genomic conservation unit" for a given chicken breed, we first calculated F_{ST} values to measure population differentiation using a non-overlapping window approach with VCFtools (Danecek et al., 2011). Then, we calculated the statistic $di = \sum_{j \neq i}^{\infty} \left(\frac{\mathrm{FST^{ij}} - E[\mathrm{FST^{ij}}]}{sd[\mathrm{FST^{ij}}]} \right)$ (Akey et al., 2010)

for each SNP, where $\mathrm{FST^{ij}}$ and $sd[\mathrm{FST^{ij}}]$ represent the expected value and standard deviation of F_{ST} between breeds i and j calculated from 28 autosomes. Finally, Di was averaged over SNVs in non-overlapping 100 kb windows, and we empirically selected the significantly high F_{ST} values (the most extreme 5% occupying the right tail) as candidate signals. As a further measure of selection, Pi was calculated for each population with non-overlapping 100 kb windows using VCFTools. Then, the Pi ratio for the three breeds was calculated by the formula: Pi ratio = log2(Pi1/ (Pi2 + Pi3)/2).

Genome Annotation and Functional enrichment analysis We used the Ensembl *Gallus gallus* BioMart webtools to retrieve genes associated with selected genomic regions identified using the methods described above. Retrieved regions were compared to the Animal QTL Database (Hu et al., 2013) (*http://www.animalgenome.org/QTLdb*) to identify candidate regions or genes associated with interesting phenotypic or economic traits. Functional enrichment analyses for Gene Ontology (GO) terms, metabolic pathways, and InterPro domains were performed using the R "clusterprofiler" package (Yu et al., 2012). All chicken genes that were annotated in Ensembl were used as a background set. P values (i.e., EASEscore), indicating significance of the overlap between various gene sets, were calculated using Benjamini-corrected modified Fisher's exact test. Only terms with a P value less than 0.05 were considered as significant.

Results

Genome sequencing and identification of variants. To detect genome-wide variation in three Chinese chicken breeds that have been conserved *in situ* , we genotyped 91 individuals using high-throughput sequencing (Figure 1). Alignment of 79.95 Gb of sequence data against the *Gallus gallus* 5.0 reference genome yielded > 8× average read depth (Table S1). In combination with genomic data obtained from the same three breeds conserved in *ex situ* programs (Zhang et al., 2018), a total of 5,070,414 variants were identified, including 4,709,112 SNPs and 361,302 short insertions and deletions (indels). Variants were evenly distributed across the genome (Figure S1a and Figure S2). 31.58% of the total SNPs were novel and had not been included in the dbSNP database at NCBI (Table S2 and Figure S1b). After removal of variations that did not meet quality criteria (MAF 0.01, HWE 10e-6), 1,518,758 SNPs remained for further analysis.

Population structure analysis. To investigate phylogenetic relationships and population structure amongst the 361 chickens, we constructed a neighbor-joining tree using a pairwise genetic distance matrix (Figure 2a) and performed principal component analysis (PCA) based on the variance-standardized genotype relationship matrix (Figure 2b). The neighbor-joining tree suggests that the samples form six major clusters that correspond to the three Chinese domestic chicken breeds, with further subdivision of each breed into *in-situ* and *ex-situ* populations. This pattern was further confirmed by PCA. The first principal component (PC1, variance explained = 11.65%) successfully separated the Langshan chicken breed from the other groups. The second principal component (variance explained = 10.7%) separated all populations in the three chicken breeds (Figure S3). Notably, the PCA separated the *in situ* and *ex situ* populations, especially for Beijing You chicken (Figure S3). To better understand p population ancestry, we used ADMIXTURE to estimate the number of ancestral populations (Alexander et al., 2009) and allowed population number (K) to vary from 2 to 9. The minimum estimated cross-validation error occurred at K=6 (Figure S4). These results suggest that the three Chinese domestic chicken breeds have distinct backgrounds and also differ between *in-situ* and *ex-situ* populations, consistent with the results from the NJ tree and principal components analyses. The likelihood model based on K=6 resolves the three Chinese domestic chicken populations into six genetic clusters (Figure 2c). One individual from the *in-situ* conserved population of Beijing You chickens (YBYC) had a genetic background that was distinct from other individuals in the YBYC population, based on the NJ tree, PCA, and ADMIXTURE results. We therefore removed this individual from subsequent analyses.

Genomic diversity assessment. Analyses of genomic genetic variability parameters for the six sub-populations are presented in Figure 3 and Table 2. The parameters include observed (Ho) and expected (He) heterozygosity, allelic richness (A_R), proportion of polymorphic SNPs (P_N), and inbreeding coefficient (F). The genomic diversity in populations conserved *in situ* was higher than in those conserved *ex situ* . Observed heterozygosity (Ho) and expected heterozygosity (He) were similar for all three breeds in both *in situ* and *ex situ* conserved populations. For example, changes in genetic diversity between *in situ* conserved populations of the Beijing You chicken (YBYC, Ho = 0.2646, He = 0.2714) vs. *ex situ* (BYC15, Ho = 0.2729, He = 0.2658) were less than 5%. In contrast, allelic richness (A_R) and proportion of polymorphic markers (P_N) for *in situ* conserved populations (A_R = 1.209, P_N = 0.7891) were higher than for *ex situ* (A_R = 1.198, P_N = 0.7258).

Estimation of inbreeding coefficients. To estimate the degree of inbreeding in *in situ* and *ex situ* conserved populations, we calculated F_{ES} andF_{ROH} across subpopulations. As expected,F_{ES} values increased while conservation procedures were maintained. This trend is also evident in the comparison of F_{ES} in *in situ* vs. *ex situ*conserved chicken populations. Conservation practices have been applied for a longer period (conservation time; CT) for the *in situ*population than the *ex situ* population, and theF_{ES} values for the *in situ* population are correspondingly higher.

Since F_{ROH} is better at detecting both rare and common variants, we focused on this measurement in subsequent analyses. The inbreeding coefficient based on runs of homozygosity (F_{ROH}) was relatively low, ranging from 0.0463 to 0.0958. Except for Langshan chickens, F_{ROH} in*in situ* conserved populations was lower thanF_{ROH} in*ex situ* populations. The difference may be caused by the small size of the Langshan chicken *in situ* conserved population and the long conservation time (CT = 60 years). Inbreeding coefficients are compared for the current generation of all three chicken breeds in Figure 3 and Table 2.

Calculation of nucleotide diversity The results of a nucleotide diversity (Pi) survey are shown for the three breeds in Figure 4a. The YLSC (Pi = 0.000112582) had the highest average nucleotide diversity amongst the 12 subpopulations, followed in descending order by YBEC, LSC15, LSC12, YBYC, BEC10, BEC07, LSC10, BYC07, BYC10, BEC15, and BYC15. Pi was markedly higher in *in situ*conserved populations than in *ex situ* in all three chicken breeds, and highly significant differences (P¡0.001) were observed between populations within breeds.

Linkage disequilibrium decay Differences in LD decay between *in situ* and *ex situ* conserved populations are shown in Figure 4b. The highest maximum average LD (r^2 = 0.2235) was observed in Beijing You chickens (BYC15), and the lowest (r^2 = 0.1806) occurred in Baier Yellow chickens (YBEC). Compared to the current generation of *ex situ* conserved populations (BYC15, BEC15, and LSC15), maximum average LD values were lower in the *in situ*conserved Beijing You chicken and Baier Yellow chicken populations, while higher values were observed in Langshan chickens. This may indicate that YBYC and YLSC have greater genetic diversity than BYC15 and LSC15. As expected, LD declined as the physical distance increased between pairwise SNPs. As shown in Figure 4b, LD decay in*in situ* conserved populations declined markedly compared with*ex situ* populations for Beijing You chicken and Baier Yellow chicken. In contrast, LD decay. is similar in *in situ* and*ex situ* conserved populations for Langshan chickens. Using Beijing You chickens as an example, r^2decreased by half (from 0.1982 to 0.0991) over a span of 11.84 kb in the*in situ* group, while LD decayed by half over a span of 14.68 kb in the *ex situ* conserved population (BYC15).

Estimation of population differentiation using Fst To estimate population differentiation, we calculated pairwiseF_{ST} values across the sub-populations (Table S3). Values ranged from 0.004826 to 0.1508. F_{ST} values for all pair-wise comparisons are shown in Figure 5. For all three breeds, F_{ST} values amongst three successive generations were lower than 0.05. This result indicates that no or little genetic differentiation has occurred in the conserved populations from one generation to the next. Significant or moderate genetic differentiation is observed between breeds, and the maximumF_{ST} value was calculated between LSC15 and BYC15 (F_{ST} = 0.1508). Notably,F_{ST} values between *in situ* and *ex situ* conserved populations for all three breeds were greater than 0.05. In the case of the Beijing You chicken, F_{ST}values have increased with time of conservation, and the maximumF_{ST} value was 0.1379 between BYC15 and YBYC. Overall, moderate genetic

differentiation has occurred in *in situ*and *ex situ* conserved populations among the three chicken breeds.

Effective population size (Ne) Ne is an important measure in conservation genetics, and conservation efforts strive to increase Ne. In order to estimate current Ne for conserved Chinese domestic chicken breeds, we used NeEstimator v2 (Do et al., 2014b), which applies a method based on linkage disequilibrium (LD) to calculate Ne using whole-genome SNPs markers. Effective population size (Ne) was estimated for autosomal chromosomes gga1 through gga28 (Table S4). Ne ranged from 2.7 to 167.4, with a mean of 43.81. Amongst macro-chromosomes (gga1-gga5), BEC15 exhibited the smallest estimated Ne (50.96), suggesting that BEC15 is a limited pool, whereas YBEC had the largest value (130.28), suggesting much higher genetic diversity. Importantly, Ne in *in situ* conserved populations was higher than in current generations of *ex situ*conserved populations (Figure 6).

Runs of homozygosity The abundance and genomic distribution of ROH provide information about the demographic history of a livestock species. ROH were identified in the genomes of all *in situ* and *ex situ* conserved populations (Table S5). A genome-wide survey for autozygosity was conducted to identify regions with signatures that reflect ancient or recent inbreeding effects. Using estimates of F_{ROH} , maximum values were found in Beijing You chickens subjected to in *ex situ* conservation. In contrast, the minimum values occurred in Baier Yellow chicken breeds enrolled in *in situ* conservation programs (Table 3). BYC15, the current generation in an *ex situ* conserved population, had the highest level of inbreeding (0.1018). As expected, YBYC in the *in situ* conservation population had a lower level of inbreeding (0.0777) than BYC15. YBEC had the lowest level of inbreeding (0.0463) amongst all populations. However, within the Langshan chicken breed, YLSC ($F_{ROH} = 0.0745$) had a higher level of inbreeding than LSC15 ($F_{ROH} = 0.0604$).

All ROH were then assessed to determine whether any populations exhibit evidence of recent inbreeding. For BYC and BEC, the *ex situ*conserved populations had longer ROH and lower genomic diversity than the *in situ* conserved populations in these breeds (Figure 7a). In contrast, the *in situ* conserved LSC population had a higher level of inbreeding than the *ex situ* conserved population. We also mapped ROH to the genome, and found that the homozygosity segments in populations subjected to *in situ* vs. *ex situ* conservation were distributed differently (Figure 7b).

Adaptation analysis

Results from the population structure and fixation index (F_{ST}) analyses show that all three chicken breeds exhibit genetic differentiation between the *in situ* and *ex situ* conserved populations. Since climate and living conditions differ between the populations maintained in *in situ* and *ex situ*programs, we hypothesized that genetic adaptation has occurred in response to these changes. Livestock populations that have adapted to different environmental niches (known as ecotypes) cannot always be distinguished easily by phenotype. In order to detect the signals of genetic differentiation, we determinedF_{ST} , Pi, and XP-EHH in 100 kb windows across the genome in the three chicken breeds (Figure 8). Candidates were defined as regions with signals that ranked in the uppermost 5% of values. In order to decrease the number of false positives, only regions identified by all three methods were retained in the final list of positive selection candidates. 186, 212, and 161 candidate regions were obtained for the three chicken breeds, Beijing You chicken, Langshan chicken, and Baier Yellow chicken, respectively (Table S7). Genes that may have experienced selection and adaptation were identified in the regions by comparison with annotated sequences. Protein-coding genes with highF_{ST} values (3719 genes), XP-EHH values (4435 genes), and $\vartheta\pi$ ratios (2504 genes) were identified in Beijing You chickens. Of these genes, 857 were identified by all three methods (Figure S5(a)). Figures S5b and S5c shows the corresponding results for Baier Yellow chickens and Langshan chickens, respectively. Clusterprofiler (Yu et al., 2012) was used to conduct Gene Ontology (GO) and KEGG pathway analyses to investigate potential functions associated with the candidate genes. Significantly enriched GO terms and KEGG pathways are shown in Figure S6.

Breed characteristics

The chicken conservation programs use random mating and random selection in *in situ* conservation, and random mating with within-family selection in *ex situ* conservation. Genome-wide SNP markers are often

used to assess genetic diversity in programs that conserve animal genetic resources for specific populations. However, this approach can result in the loss of genetic diversity for some breed-specific traits, thereby reducing performance (Sun et al., 2018).

The three chicken breeds in this study have specific performance traits and phenotypes of value, as discussed in our previous study (Zhang et al., 2018). Because these characteristics are important, programs that maintain the overall genetic diversity of each breed must also maintain the distinct phenotypic variations. Since the size of a conservation population is usually small (30 males and 300 females), genetic drift can easily occur and alleles that contribute to special traits can be lost. Therefore, it is crucial to monitor and maintain the genetic diversity associated with these special traits. Here, we utilized Di and Pi to sweep the selective signatures to identify genomic regions related with breed characteristics. We then compared these regions to the Chicken QTL database (Hu et al., 2013) (http://www.animalgenome.org/cgi-bin/QTLdb). Regions or SNPs that overlapped with QTLs were classified tentatively as the "genomic conservation unit" for a specific breed.

Based on the population-scale genetic differences between Beijing You chicken and the other breeds, we hypothesized that specific genome signals have appeared in the Beijing You chicken population during domestication. To localize these selective sweeps in the Beijing You chicken genome, we calculated two genome-wide statistics, Di and the Pi ratio, for the Beijing You chicken vs. the Baier Yellow and Langshan populations. Focusing on the regions in the top 5% of the Di and Pi ratio empirical distribution (Figure 9), we identified 59 significant regions (Di $>$ 0.5036 and $\log_2$ Pi ratio $>$ 0.9433) harboring 255 candidate genes (Table S5b). Comparing the genomic regions with the Chicken QTL database (Hu et al., 2013), we found that these selected regions were related to important economic traits such as growth, body weight, and feed conversion ratio. In the Baier Yellow chicken, we identified 50 candidate genomic regions (Di $>$0.4762 and $\log_2$ Pi ratio $>$0.6137) harboring 202 genes (Figure 10, Table S5a). These genomic regions overlapped with QTLs for body weight (day of first egg), egg number, egg production rate, and earlobe color. Finally, 36 genomic regions were identified in Langshan chickens (Di$>$0.5064 and log2 Pi ratio$>$0.6572) (Figure 11 and Table S5c).

A phylogenetic tree reflects the genetic distance among animals and enables the selection of individuals for breeding (Sun et al., 2018). Based on the pairwise distance matrix, a phylogenetic tree for the three *ex situ* populations was constructed using the neighbor-joining method (Saitou and Nei, 1987) (Figure S6a). SNPs located in the selected genome regions were defined as breed-specific markers, and the selected genome regions were defined as the "genomic conservation unit". Using these criteria, 9029, 7260, and 3907 SNPs located within genomic conservation units were identified in the Beijing You chicken, Baier Yellow chicken, and Langshan chicken populations, respectively. The phylogenetic trees were then reconstructed using these SNPs (Figure S6b).

Using the phylogenetic trees together, it is possible to identify individuals that embody whole genomic diversity as well as breed-specific characteristics. For example, consider the 6 pedigrees in Figure S7a and 7 pedigrees in Figure S6b for the Beijing You chicken. If the conservation program was constructed using only the genome-wide SNPs, individuals 2, 4, 7, 23, 39, and 52 might not be chosen for breeding, resulting in the loss of the genetic diversity associated with important economic traits. It is therefore critical to select individuals that represent both pedigrees in Figure S6a and Figure S6b. Similar precautions would need to be observed for the Baier Yellow and Langshan populations. These additional steps improve the ability of the conservation program to retain critical breed characteristics.

Discussion

Because China has a long history of animal husbandry and diverse geographical conditions, it has developed extensive genetic resources for the chicken, with 107 different indigenous breeds. Chickens are one of the most widely distributed livestock animals in China. Worldwide, they also have a significant role as a source of income and high-quality protein. Indigenous chickens possess enormous genetic diversity, especially in adaptative traits, including the ability to survive harsh conditions, shifting climate, urbanization, disease

epidemics, selection errors, and many other potential stresses (Anderson, 2003; Anderson and Centonze, 2007). Most Chinese domestic chicken breeds also have unique meat and /or egg qualities, as well as other useful breed characteristics. However, the majority of these chickens are currently maintained as small populations. Generally, the smaller a livestock population, the greater is its vulnerability to extinction (Biscarini et al., 2015; Henson et al., 1992; Ramstad et al., 2004). Many favorable alleles can be lost as a result of selection or genetic drift. The successful preservation and utilization of these local breeds depend on the accurate assessment of conservation efficiency, which is the essential measure of an effective conservation program. Both *in situ* and *ex situ* programs have been established for the management of poultry genetic resources in China. In this study, we used SNPs obtained by high-throughput genome sequencing to assess genomic diversity for chickens managed in *in situ* and*ex situ* conservation programs, and proposed a strategy for conserving the specific breed characteristics of three Chinese domestic chicken breeds (Beijing You chicken, Langshan chicken, and Baier Yellow chicken).

The results show that all three chicken breeds, conserved both in*in situ* and *ex situ* , have maintained rich genetic diversity as measured by heterozygosity (Ho, He), proportion of polymorphic markers (P_N), and allelic richness (A_R). Breeds conserved *in situ* exhibited higher genetic diversity than those conserved *ex situ* . Although conservation time was longer for the *in situ* than the *ex situ* populations, the *in situ* populations were larger. We estimated the effective population size (Ne) based on whole genome SNPs for the conserved populations, and evaluated Ne based on the macrochromosome class (gga1–5). For all three breeds, Ne was larger for chickens enrolled in *in situ* conservation programs (Figure 6 and Table S4). We also estimated inbreeding coefficients based on whole genome SNPs markers. As expected, the inbreeding coefficient for Baier Yellow chickens and Beijing You chickens conserved *in situ* were lower than for those conserved *ex situ* . However, the opposite was true for Langshan chickens. The discrepancy may reflect the fact that the *in situ* conservation time for the Langshan chicken was the longest among the three breeds, and is currently 60 years. The largest inbreeding coefficient was 0.0958 for the Beijing You chicken (*ex situ*), which meets our program goal to maintain 90% of whole genomic diversity from the initial population, and limit the inbreeding coefficient to less than 0.1 for 100 years .

Based on assessments using NJtree, PCA, STRUCTURE, F_{ST}, and the distribution of ROH, genetic differentiation has occurred between the two conserved populations for all three chicken breeds. The adaptation may have been driven by environmental differences, or selection may differ between the *in situ* and the *ex situ*programs. Few studies have examined and compared the structures of*in situ* and *ex situ* conserved populations, and it is not known how the genetics of domestic chickens may change in response to a shift from the *in situ* to the *ex situ* conditions over several decades. To explore the genetic mechanisms underlying the differentiation between the *in situ* and *ex situ* conserved chickens, we used F_{ST}, Pi and XP-EHH to detect regions that were different between the two populations. Based on the gene functions revealed by KEGG and GO term analysis, we hypothesize that the genetic differences may be related to adaptation to local environmental conditions. For example, the selection and conservation of the Beijing You chicken began at the BAAFS institute of Animal Husbandry and Veterinary Medicine in 1972. This *in situ* conservation program reached its 47^{th} year in 2018. In 1976, Beijing You chickens were obtained from the Beijing program and transferred to Yangzhou, Jiangsu (National Chickens Genetic Resources) to establish an*ex situ* program, which reached its 40^{th} year in 2015. Climate conditions in the two locations are markedly different. In contrast, the conservation programs for Baier Yellow chickens and Langshan chickens were conducted under nearly identical climate conditions at Zhejiang and Jiangsu. The population sizes for these chickens were very small at the onset of the conservation program, so it is possible that the genetic differentiation has been caused by genetic drift over several decades. Alternatively, the very small founder populations used in these programs may have sampled different variants from the original populations simply by chance.

Breed conservation has recently entered the genomic era. For both*in situ* and *ex situ* conservation, breeders now use genomic markers to estimate genomic diversity and reconstruct pedigrees, enabling them to reveal relationships among animals in a population and select individuals for mating. However, this cannot be pursued blindly. In the case of domestic chickens, some breeds have special characteristics and it is important to preserve these traits. More specifically, utilization of markers for the sole purpose of optimizing genomic

diversity will not necessarily preserve breed-specific characteristics. This shortcoming can be addressed by genome sequencing, which makes it possible to identify and conserve variations responsible for breed-specific characteristics. We first used Di and Pi to identify the "genomic conservation unit" that is associated with specific breed characteristics. SNP markers located in these regions can be used to safeguard breed-specific characteristics in a conservation program. Specifically, trait-specific markers are not only useful for rebuilding phylogenetic trees for identifying valuable individuals, but are also useful as reference markers for custom SNP chips that can track breed-specific characteristics at the molecular level to monitor conservation efficiency.

Conclusion

Maintaining the genomic diversity of Chinese domestic chicken breeds is important for economic and cultural reasons. We found that the small*ex situ* conserved populations that are maintained in controlled environments retain less genetic diversity than populations conserved*in situ*. In addition, the transfer of conservation populations from their place of origin to another site results in genetic differentiation. This may be caused by genetic drift or adaptation. Finally, we identified the "genomic conservation unit" for all three chicken breeds and proposed that phylogenetic trees should be used to select valuable indi18viduals for breeding within a conservation program, based on breed-specific characteristics markers as well as genome-wide markers. The results of this study provide a basis for further optimization of conservation programs for domestic chicken breeds in China.

Acknowledgements

This study was supported by the Program for Changjiang Scholars and Innovation Research Teams in the University (IRT_15R62) and the Beijing Innovation Consortium of Swine Research System (BAIC02-2019).

Reference

Beckmann, J. & Soller, M. (1986). Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35, 111–124.

Blears, M., De Grandis, S., Lee, H. & Trevors, J. (1998). Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. Journal of Industrial Microbiology and Biotechnology 21, 99–114.

FAO (2013). Status and Trends of Animal Genetic Resources 2012. Food and Agriculture Organization, Commission on Animal Genetic Resources, Rome.

Henson, E. L. (1992). In Situ Conservation of Livestock and Poultry. Food and Agriculture

Organization of the United Nations, Rome.

Hiemstra, S. J., Drucker, A. G., Tvedt, M. W., Louwaars, N., Oldenbroek, J., Awgichew, K., Kebede, S. A., Bhat, P. & Da Silva Mariante, A.

(2006 a). Exchange, use and Conservation of Animal Genetic Resources. Food and Agriculture Organization of the United Nations, Rome.

http://hgdownload.cse.ucsc.edu/goldenPath/galGal5/bigZips/galGal5.fa.gz.

https://ftp.ncbi.nlm.nih.gov/snp/organisms/archive/chicken_9031/VCF/.

R-Project*https://www.r-project.org/.*

The second report on the state of the world's animal geentic resources for food and agricultural.

Tixier-Boichard, M. (2014). Status and gaps in characterization of animal genetic

resources. In 10th World Congress on Genetics Applied to Livestock Production. August 17–22, 2014, p. 29. American Society of Animal Science, Vancouver.

Wu CX. Poultry genetic resources in China: conservation and utilization. Beijing: XXV World's Poultry Congress; 2016. p. 88–91.

Akey, J.M., Ruhe, A.L., Akey, D.T., Wong, A.K., Connelly, C.F., Madeoy, J., Nicholas, T.J., and Neff, M.W. (2010). Tracking footprints of artificial selection in the dog genome. Proc Natl Acad Sci U S A *107* , 1160-1165.

Akey, J.M., Zhang, G., Zhang, K., Jin, L., and Shriver, M.D. (2002). Interrogating a high-density SNP map for signatures of natural selection. Genome Res *12* , 1805-1814.

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

Ali, B.A., Huang, T.-H., Qin, D.-N., and Wang, X.-M. (2004). A review of random amplified polymorphic DNA (RAPD) markers in fish research. Reviews in Fish Biology and Fisheries *14* , 443-453.

Anderson, S. (2003). Animal genetic resources and sustainable livelihoods. Ecological Economics *45* , 331-339.

Anderson, S., and Centonze, R. (2007). Property Rights and the Management of Animal Genetic Resources. World Development *35* , 1529-1541.

Andersson, L., and Georges, M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. Nature Reviews Genetics*5* , 202-212.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., and Saunders, N.C. (1987). INTRASPECIFIC PHYLOGEOGRAPHY: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics *18* , 489-522.

Avise, J.C., Ellis, D., Clarke Bryan, C., Robertson, A., and Jeffreys Alec, J. (1986). Mitochondrial DNA and the evolutionary genetics of higher animals. Philosophical Transactions of the Royal Society of London B, Biological Sciences *312* , 325-342.

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

Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2004). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics*21* , 263-265.

Biscarini, F., Nicolazzi, E.L., Stella, A., Boettcher, P.J., and Gandini, G. (2015). Challenges and opportunities in genetic improvement of local livestock breeds. Front Genet *6* , 33.

Bruford, M.W., Bradley, D.G., and Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. Nature Reviews Genetics*4* , 900-910.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T.*, et al.* (2011). The variant call format and VCFtools. Bioinformatics*27* , 2156-2158.

de Cara, M.A.R., Fern'?ndez, J., Toro, M.A., and Villanueva, B. (2011). Using genome-wide information to minimize the loss of diversity in conservation programmes. Journal of Animal Breeding and Genetics*128* , 456-464.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. (2014a). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour *14* , 209-214.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. (2014b). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources *14* , 209-214.

Dodgson, J.B., Cheng, H.H., and Okimoto, R. (1997). DNA marker technology: a revolution in animal genetics. Poultry Science *76* , 1108-1114.

Groeneveld, E., Westhuizen, B., Maiwashe, A., Voordewind, F., and Ferraz, J.B. (2009). POPREP: a generic report for population management. Genet Mol Res *8* , 1158-1178.

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E.*, et al.* (2010). Genetic diversity in farm animals q€“ a review. Animal Genetics *41* , 6-31.

Harrison, R.G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends in Ecology & Evolution*4* , 6-11.

Henson, E.L., Food, and Nations, A.O.o.t.U. (1992). In Situ Conservation of Livestock and Poultry (Food and Agriculture Organization of the United Nations).

Holsinger, K.E., and Weir, B.S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. Nature Reviews Genetics *10* , 639.

Hu, Z.L., Park, C.A., Wu, X.L., and Reecy, J.M. (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Res *41* , D871-879.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X., and Wilson, A.C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci U S A *86* , 6196-6200.

Koh, M.C., Lim, C.H., Chua, S.B., Chew, S.T., and Phang, S.T.W. (1998). Random amplified polymorphic DNA (RAPD) fingerprints for identification of red meat animal species. Meat Science *48* , 275-285.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol *33* , 1870-1874.

Levin, I., Crittenden, L.B., and Dodgson, J.B. (1993). Genetic Map of the Chicken Z Chromosome Using Random Amplified Polymorphic DNA (RAPD) Markers. Genomics *16* , 224-230.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25* , 2078-2079.

Liu, Z.J., and Cordes, J.F. (2004a). DNA marker technologies and their applications in aquaculture genetics. Aquaculture *238* , 1-37.

Liu, Z.J., and Cordes, J.F. (2004b). Erratum to “DNA marker technologies and their applications in aquaculture genetics” [Aquaculture 238 (2004) 1–37]. Aquaculture *242* , 735-736.

McMahon, B.J., Teeling, E.C., and H’?glund, J. (2014). How and why should we implement genomics into conservation? Evolutionary Applications *7* , 999-1007.

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A.*, et al.* (2008). Runs of homozygosity in European populations. Am J Hum Genet *83* , 359-372.

Morin, P.A., Luikart, G., Wayne, R.K., and the, S.N.P.w.g. (2004). SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution*19* , 208-216.

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

Parsons, Y.M., and Shaw, K.L. (2001). Species boundaries and genetic diversity among Hawaiian crickets of the genus Laupala identified using amplified fragment length polymorphism. Molecular Ecology *10* , 1765-1772.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet *38* , 904-909.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J.*, et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81* , 559-575.

Rambaut, A. (2018). FigTree v1. 4.4, a graphical viewer of phylogenetic trees. Available from*http://tree.bio.ed.ac.uk/software/figtree*.

Ramstad, K.M., Woody, C.A., Sage, G.K., and Allendorf, F.W. (2004). Founding events influence genetic population structure of sockeye salmon (Oncorhynchus nerka) in Lake Clark, Alaska. Molecular Ecology *13* , 277-290.

Rescources, C.N.C.O.A.G. (2011). Animal Genetic Rescources In China Poultry.

Sabeti, P.C., Schaffner, S.F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T.S., Altshuler, D., and Lander, E.S. (2006). Positive Natural Selection in the Human Lineage. Science*312* , 1614.

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

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution *4* , 406-425.

Savelkoul, P.H.M., Aarts, H.J.M., de Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J.L.W., Schouls, L., and Lenstra, J.A. (1999). Amplified-Fragment Length Polymorphism Analysis: the State of an Art. Journal of Clinical Microbiology *37* , 3083.

Sun, H., Olasege, B.S., Xu, Z., Zhao, Q., Ma, P., Wang, Q., Lu, S., and Pan, Y. (2018). Genome-Wide and Trait-Specific Markers: A Perspective in Designing Conservation Programs. Front Genet *9* , 389.

Szpiech, Z.A., and Hernandez, R.D. (2014). selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol *31* , 2824-2827.

Szpiech, Z.A., Jakobsson, M., and Rosenberg, N.A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics *24* , 2498-2504.

Thurston, L.M., Siggins, K., Mileham, A.J., Watson, P.F., and Holt, W.V. (2002). Identification of Amplified Restriction Fragment Length Polymorphism Markers Linked to Genes Controlling Boar Sperm Viability Following Cryopreservation1. Biology of Reproduction *66* , 545-554.

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. Genetics, selection, evolution : GSE *34* , 275-305.

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

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

Waples, R.S., and Do, C. (2008). ldne: a program for estimating effective population size from data on linkage disequilibrium. Mol Ecol Resour *8* , 753-756.

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

WRIGHT, S. (1949). THE GENETICAL STRUCTURE OF POPULATIONS. Annals of Eugenics *15* , 323-354.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology *16* , 284-287.

Zane, L., Bargelloni, L., and Patarnello, T. (2002). Strategies for microsatellite isolation: a review. Molecular Ecology *11* , 1-16.

Zeder, M.A., Emshwiller, E., Smith, B.D., and Bradley, D.G. (2006). Documenting domestication: the intersection of genetics and archaeology. Trends in Genetics *22* , 139-155.

Zhang, D.-X., and Hewitt, G.M. (1996). Nuclear integrations: challenges for mitochondrial DNA markers. Trends in Ecology & Evolution *11* , 247-251.

Zhang, M., Han, W., Tang, H., Li, G., Zhang, M., Xu, R., Liu, Y., Yang, T., Li, W., Zou, J., *et al.* (2018). Genomic diversity dynamics in conserved chicken populations are revealed by genome-wide SNPs. BMC Genomics *19* , 598.

Author contributions

K.L.W. conceived and designed the experiments. M.M.Z performed bioinformatics analyses and experiments, wrote the manuscript, and prepared the figures. All authors reviewed and approved the manuscript.

Data accessibility

All relevant raw data descripted in our manuscript are available publicly prior to publication. However, we are uncertain which repository will be most appropriate for our data, and would sincerely appreciate your guidance.

Figure Legends

Figure 1. Map of China (including the South China Sea Islands, lower right corner) showing chicken populations included in this study. BEC, Baier Yellow chicken (*ex-situ*); YBEC, Baier Yellow chicken (*in-situ*); BYC, Beijing You chicken (*ex-situ*); YBYC, Beijing You chicken (*in-situ*); LSC, Langshan chicken (*ex-situ*); YLSC, Langshan chicken (*in-situ*). Male and female specimens are shown for the three breeds. Each subpopulation in the study consisted of 10 males and 20 females (green and brown areas in pie charts, respectively). Airplane glyphs indicate that individuals from each breed were moved from their original locations (*in situ*in Beijing, Hangzhou, and Rudong) to Yangzhou for *ex situ*conservation under the auspices of NCGR (National Chicken Genetic Resources) in Jiangsu.

Figure 2. Population genetic structure. (a). Neighbor-joining tree constructed using SNP data from *ex-situ* and *in-situ*conserved populations of the three chicken breeds. (b). PCA analysis of subpopulations. The first three principal components are shown, and the subpopulations are color-coded according to the key to the right. (c). Inferred population genetic structure using the maximum-likelihood method under a model with ancestral components varying from K=2 to K=6.

Figure 3. Analysis of genomic diversity between *in situ*and *ex situ* populations within breeds. Ho, observed heterozygosity; He, expected heterozygosity; P_N, proportion of polymorphic markers; A_R, allelic richness;

F_{ROH}, inbreeding coefficients based on ROH; F_{ES}, inbreeding coefficient based on pedigree.

Figure 4. (a). Estimates of population nucleotide diversity (Pi-LSC10 = 0.000102248, LSC12 = 0.000105831, LSC15 = 0.000106998, YBYC = 0.000105085, BYC07 = 0.000100251, BYC10 = 9.84366e-05, BYC15 = 9.46419e-05, YBEC = 0.000107512, BEC07 = 0.000104584, BEC10 = 0.000104865, BEC15 = 9.71941e-05). (b). LD decay determined by squared correlations of allele frequencies (r^2) against the distance between polymorphic sites.

Figure 5. Matrix showing pairwise differentiation estimates (F_{ST}) between *in situ* and *ex situ*populations.

Figure 6. Boxplots showing effective population sizes of the*in situ* and *ex situ* conserved populations for each breed.

Figure 7. (a). Number of runs of homozygosity (ROH) as a function of run size in kb. (b). Circos plot showing genomic location of runs of homozygosity for each of the three chicken breeds in *in situ* and *ex situ* conserved populations.

Figure 8. Circos Manhattan plots showing results forF_{ST} , Pi, and XP-EHH analyses in (a). Baier Yellow chicken, (b). Beijing You chicken, and (c). Langshan chicken.

Figure 9. Selection sweep of Beijing You chicken population. (a) Classification of sweep regions for Beijing You chicken. The X-axis represents the Di value and Y-axis represents the Pi value. The green circles satisfy criteria for sweep regions. (b) Manhattan plot for genome-wide distribution of Di and Pi values. Baier Yellow chicken and Langshan chicken served as reference groups.

Figure 10. Selection sweep of Baier Yellow chicken population. (a) Classification of sweep regions for Baier Yellow chicken. The X-axis represents the Di value and Y-axis represents the Pi value. The red circles satisfy criteria for sweep regions. (b) Manhattan plot for genome-wide distribution of Di and Pi values. Beijing You chicken and Langshan chicken served as reference groups.

Figure 11. Selection sweep of Langshan chicken chicken population. (a) Classification of sweep regions for Langshan chicken chicken. The X-axis represents the Di value and Y-axis represents the Pi value. The red circles satisfy criteria for sweep regions. (b) Manhattan plot for genome-wide distribution of Di and Pi values. Beijing You chicken and Baier Yellow chicken served as reference groups.

Supporting information

Figure S1.(a) SNP density and distribution across the genome (b). Number of novel SNPs vs. those found within the dbSNP database.

Figure S2. Indel density and distribution across the genome.

Figure S3. Biplots showing PC1 vs. PC2, PC1 vs PC3, and PC2 vs PC3.

Figure S4. The CV error associated with each K value.

Figure S5. Venn diagrams showing numbers of genes identified using F_{ST} , Pi, and XP-EHH analyses for (a). Baier Yellow Chicken. (b). Beijing You Chicken. (c). Langshan Chicken.

Figure S6. Go term and KEGG analysis for (a). Beijing You Chicken. (b). Baier Yellow Chicken. (c). Langshan Chicken.

Figure S7. (a). Phylogenetic tree constructed using the genome-wide SNPs; (b). Phylogenetic tree constructed using SNPs associated with economic traits. (a). Beijing You chicken. (b). Baier Yellow Chicken. (c). Langshan chicken.

Table S1. Summary statistics for genome sequencing.

Table S2. Summary of genome sequencing and annotation of variants for the three Chinese domestic chicken breeds.

Table S3. Estimation of the pairwise genetic differentiation statistic among breeds (F_{ST}).

Table S4. Effective population size (Ne) estimated for the three breeds in *in situ* and *ex situ* conservation programs.

Table S5. Statistical summary of analysis for runs of homozygosity in *in situ* and *ex situ* conserved chicken populations.

Table S6. Candidate regions in (a) Baier Yellow chicken. (b). Beijing You chicken. (c). Langshan chicken.

Table 1. Pedigree information of the In-situ and ex-situ conserved chicken populations

Breeds	In-situ	In-situ	In-situ	In-situ	In-situ	In-situ	In-situ	In-situ	Ex-situ	Ex-situ	Ex-situ	Ex-situ	Ex-situ
	Conservation first generation	Conservation scale	Conservation scale	Samples	Samples	Sampling collection location	Sampling collection time	Code	Conservation first generation	Conservation scale	Conservation scale	Samples	Samp
		Sire	Dam	Sire	Dam					Sire	Dam	Sire	Dam
Baier Yellow Chicken	1982	[?]30	[?]300	10	20	Zhejiang	2018	YBEC	1998	30	300	10	20
												10	20
												10	20
Beijing You Chicken	1972	[?]30	[?]300	10	20	Beijing	2018	YBYC	1976	30	300	10	20
												10	20
												10	20
Langshan Chicken	1959	[?]30	[?]300	10	20	Jiangsu	2018	YLSC	1998	30	300	10	20
												10	20
												10	20

NCGR: National Chickens Genetic Resources (Jiangsu).

Table 2. Parameters of genomic diversity of three domestic chicken populations

Populations	Ho	He	P_N (%)	A_R	$F_{ES}{}^a$	$F_{ROH}{}^b$
BEC07	0.2690	0.2635	0.8098	1.218	0.0135	0.0494
BEC10	0.2764	0.2681	0.7983	1.221	0.0175	0.0500
BEC15	**0.2793**	**0.2649**	**0.7481**	**1.209**	**0.0241**	**0.0719**
YBEC	**0.2711**	**0.2635**	**0.8327**	**1.226**	**0.1602**	**0.0463**
BYC07	0.2690	0.2612	0.7833	1.211	0.0424	0.0818
BYC10	0.2732	0.2634	0.7627	1.208	0.0463	0.0679
BYC15	**0.2729**	**0.2658**	**0.7258**	**1.198**	**0.0528**	**0.0958**
YBYC	**0.2646**	**0.2714**	**0.7891**	**1.209**	**0.1942**	**0.0777**
LSC10	0.2818	0.2686	0.7753	1.218	0.0175	0.0481
LSC12	0.2796	0.2699	0.7995	1.223	0.0201	0.0502
LSC15	**0.2815**	**0.2721**	**0.8013**	**1.226**	**0.0241**	**0.0604**

Populations	Ho	He	P_N (%)	A_R	F_{ES} [a]	F_{ROH} [b]
YLSC	**0.2802**	**0.2784**	**0.8251**	**1.231**	**0.2410**	**0.0745**

He Expected heterozygosity, Ho Observed heterozygosity, P_N Proportion of polymorphic SNPs, A_R Allelic richness, F_{ES} [a], inbreeding coefficient based on pedigree; F_{ROH} [b], inbreeding coefficient based on the runs of homozygosity

Figure 1

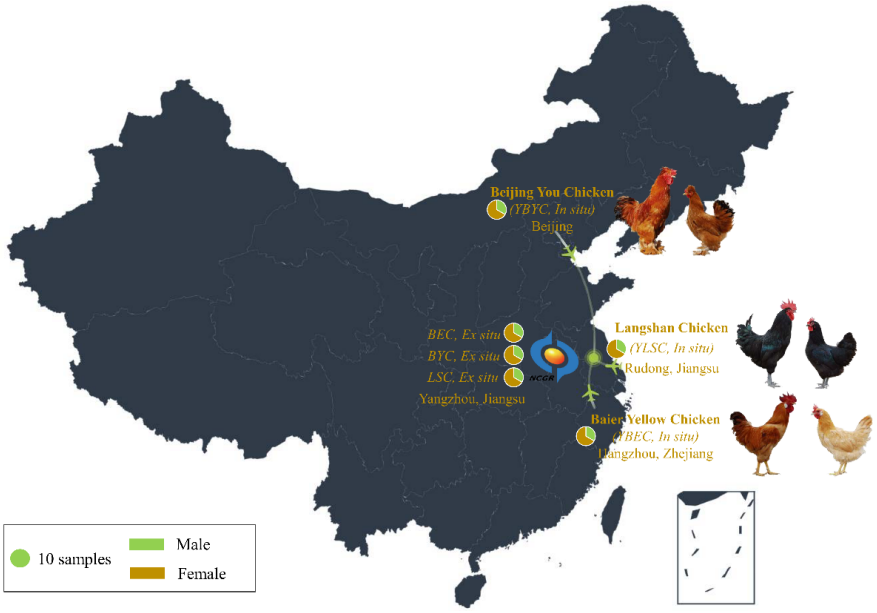

Figure 2

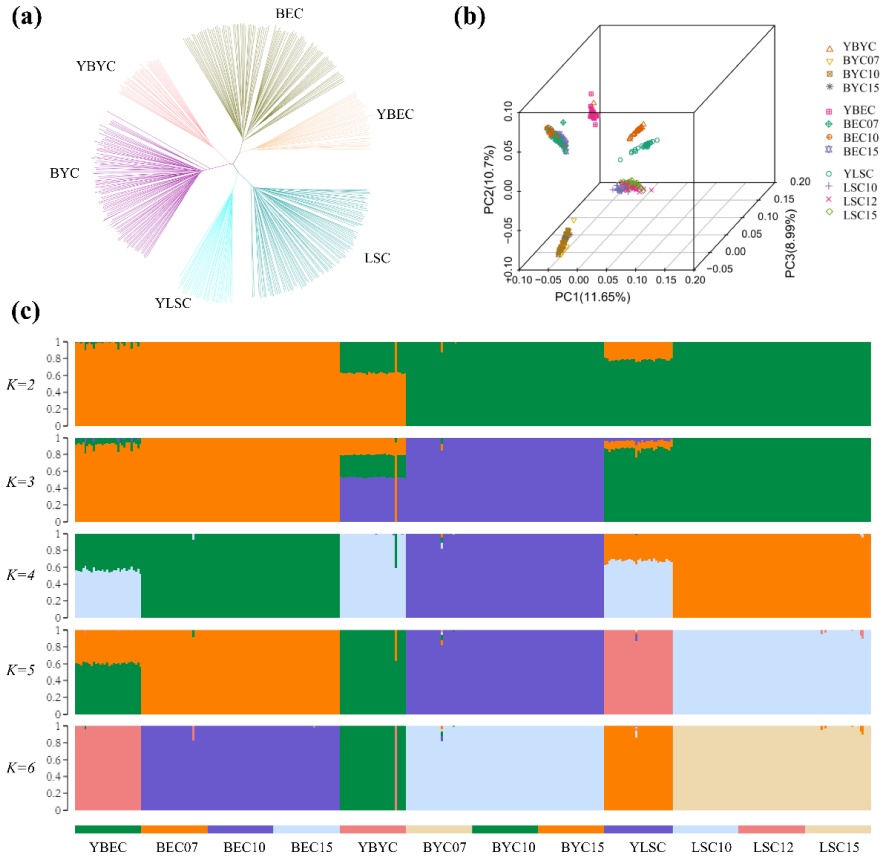

Figure 3

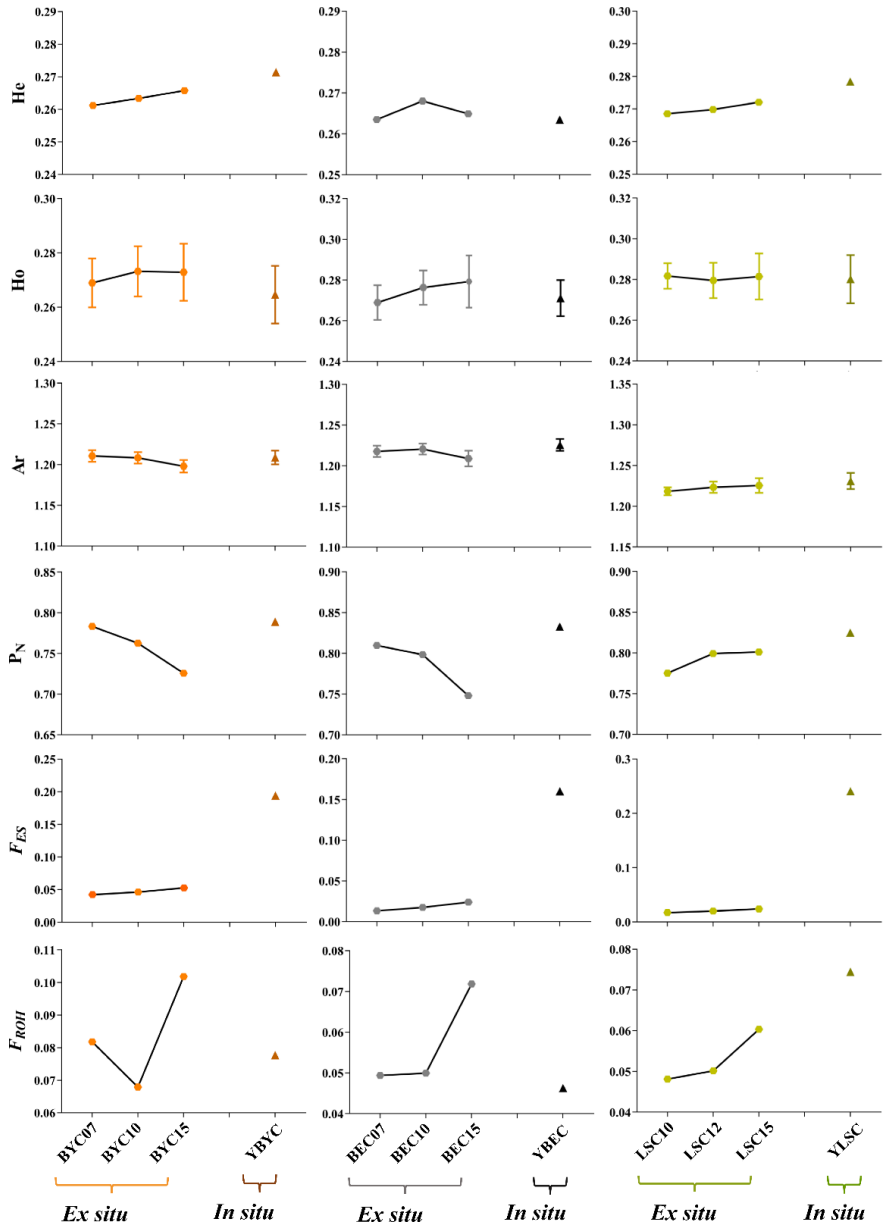

Figure 4

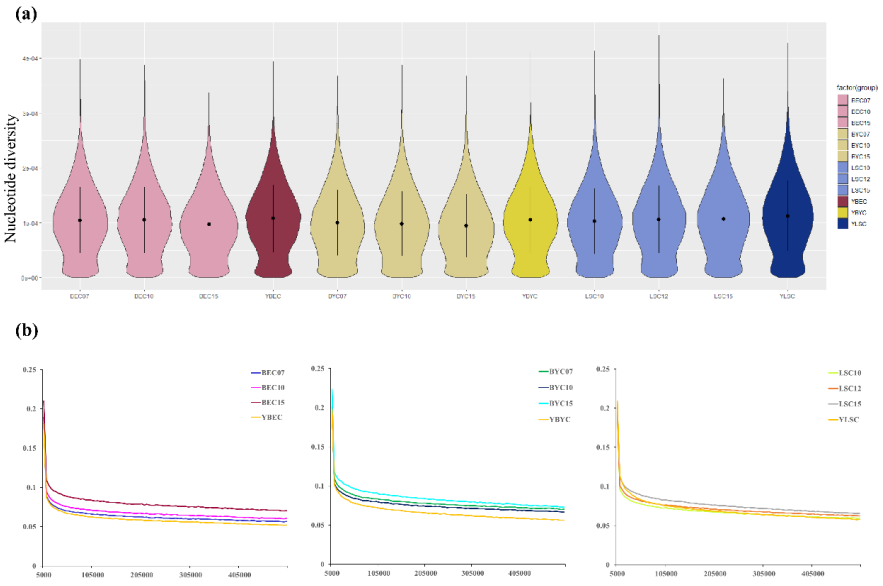

Figure 5

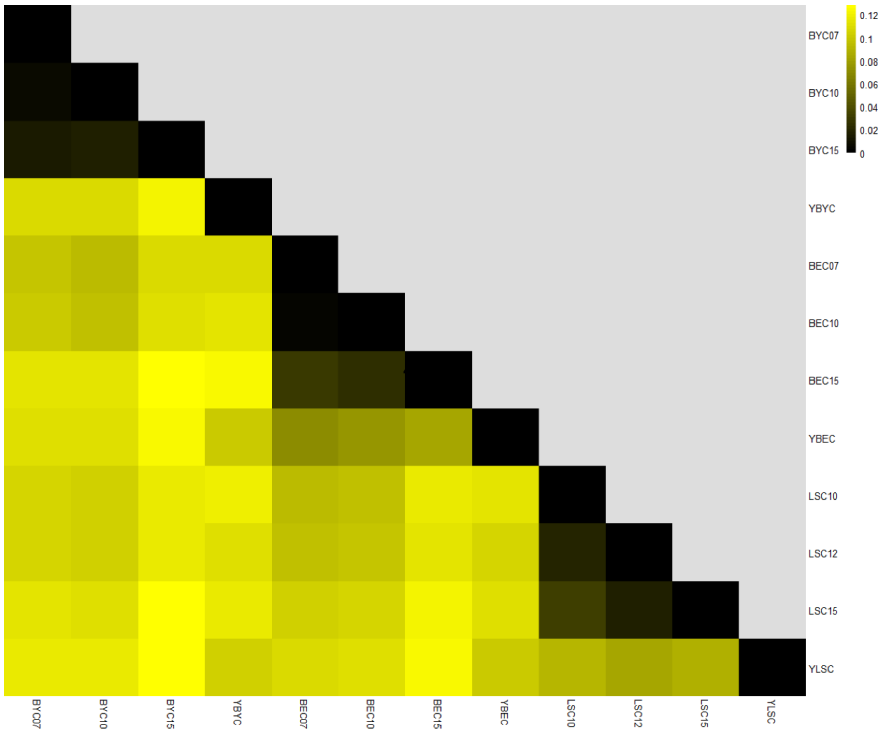

Figure 6

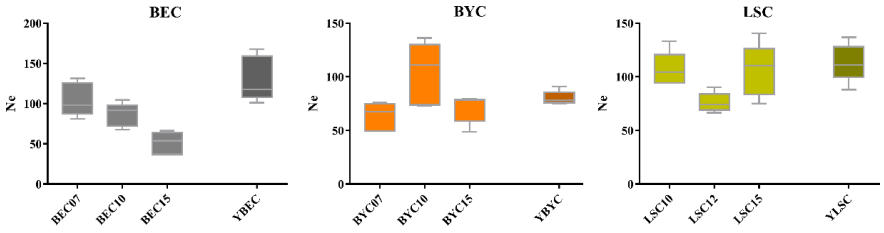

Figure 7

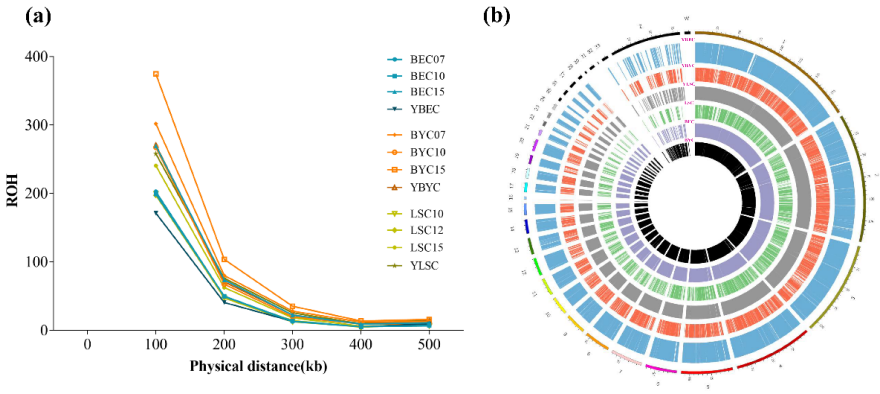

Figure 8

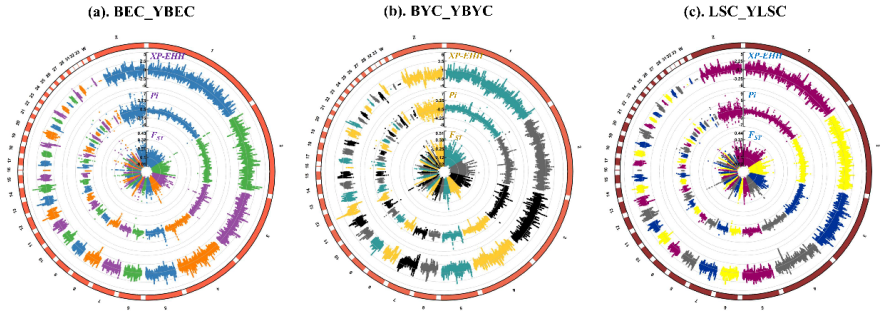

Figure 9

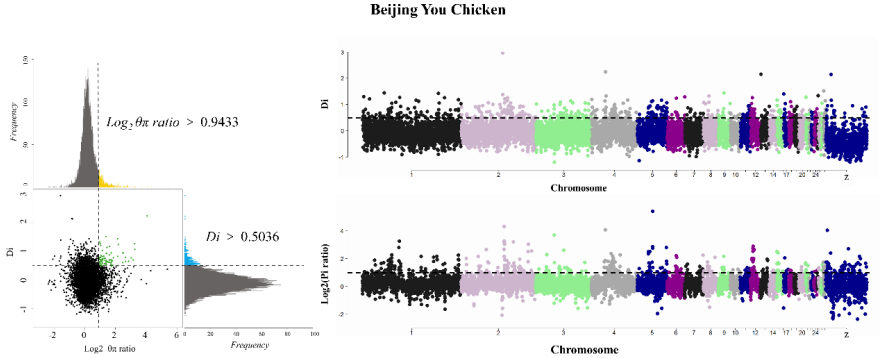

Figure 10

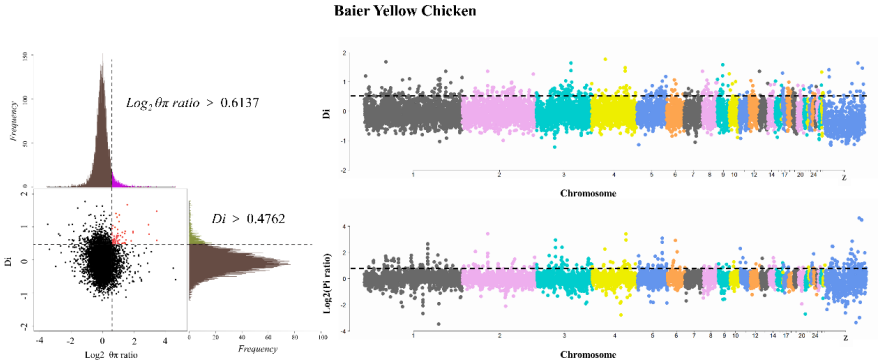

Figure 11

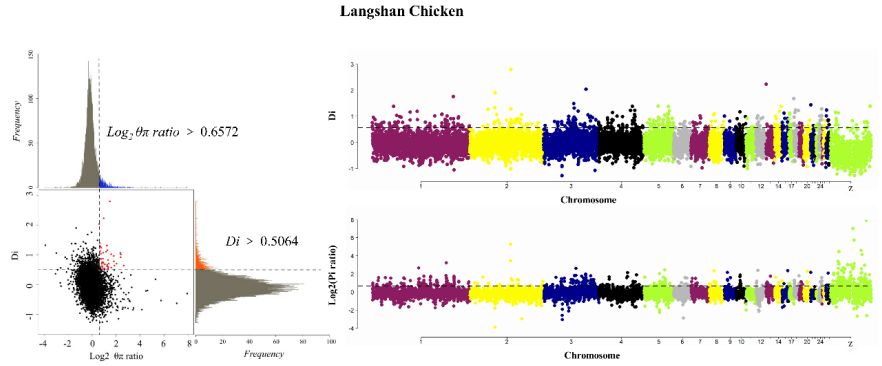

Figure S1

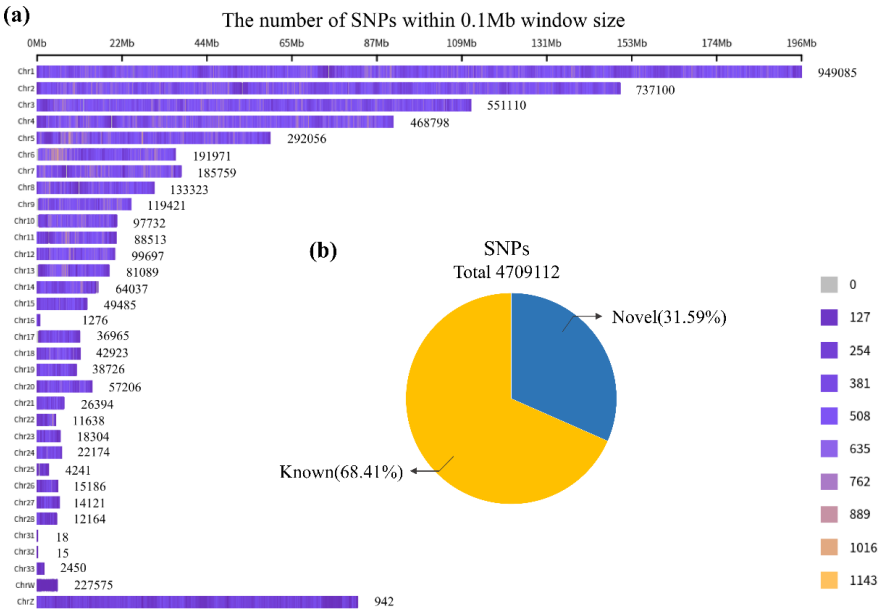

Figure S2

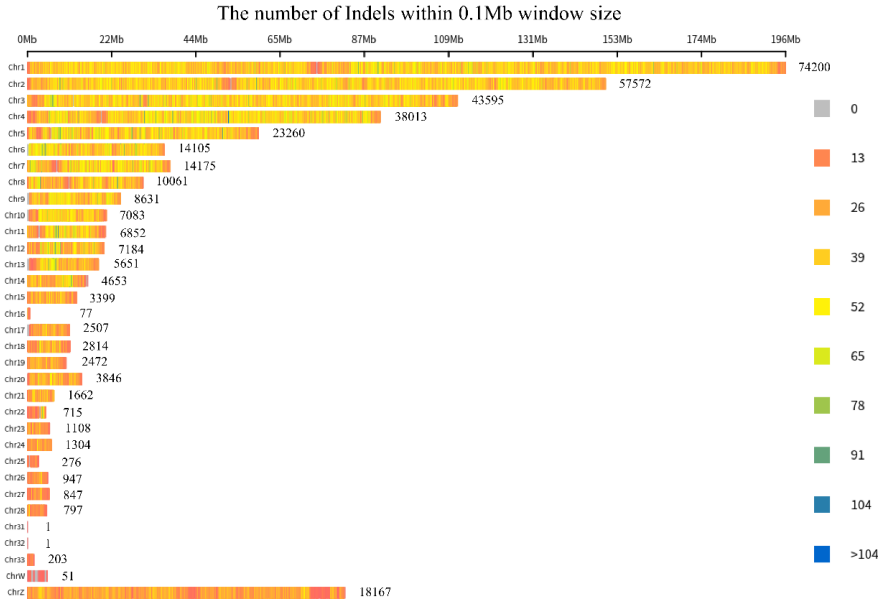

Figure S3

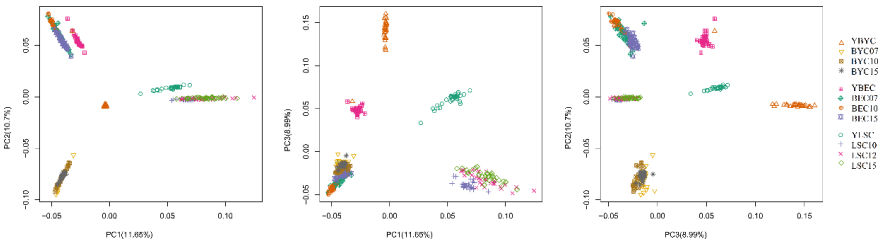

Figure S4

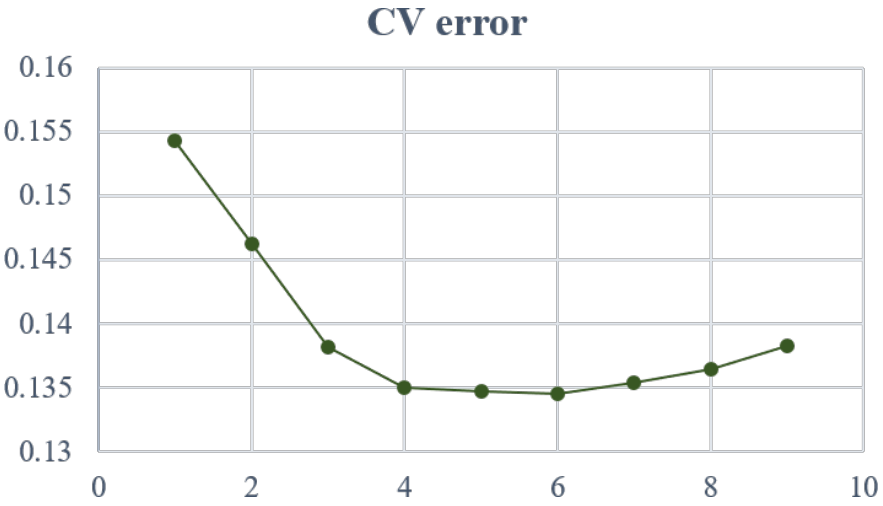

Figure S5

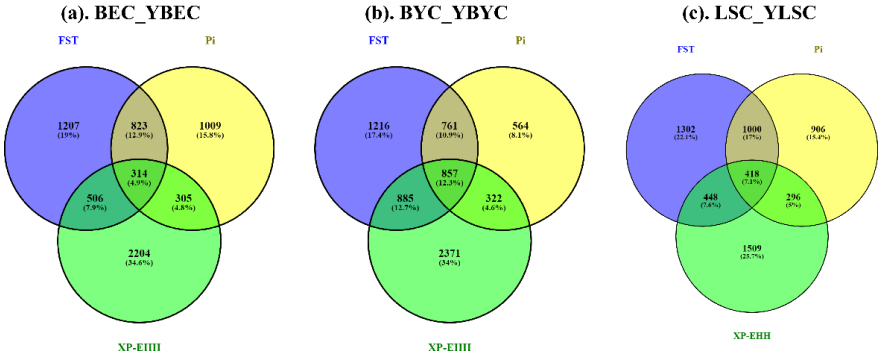

Figure S6

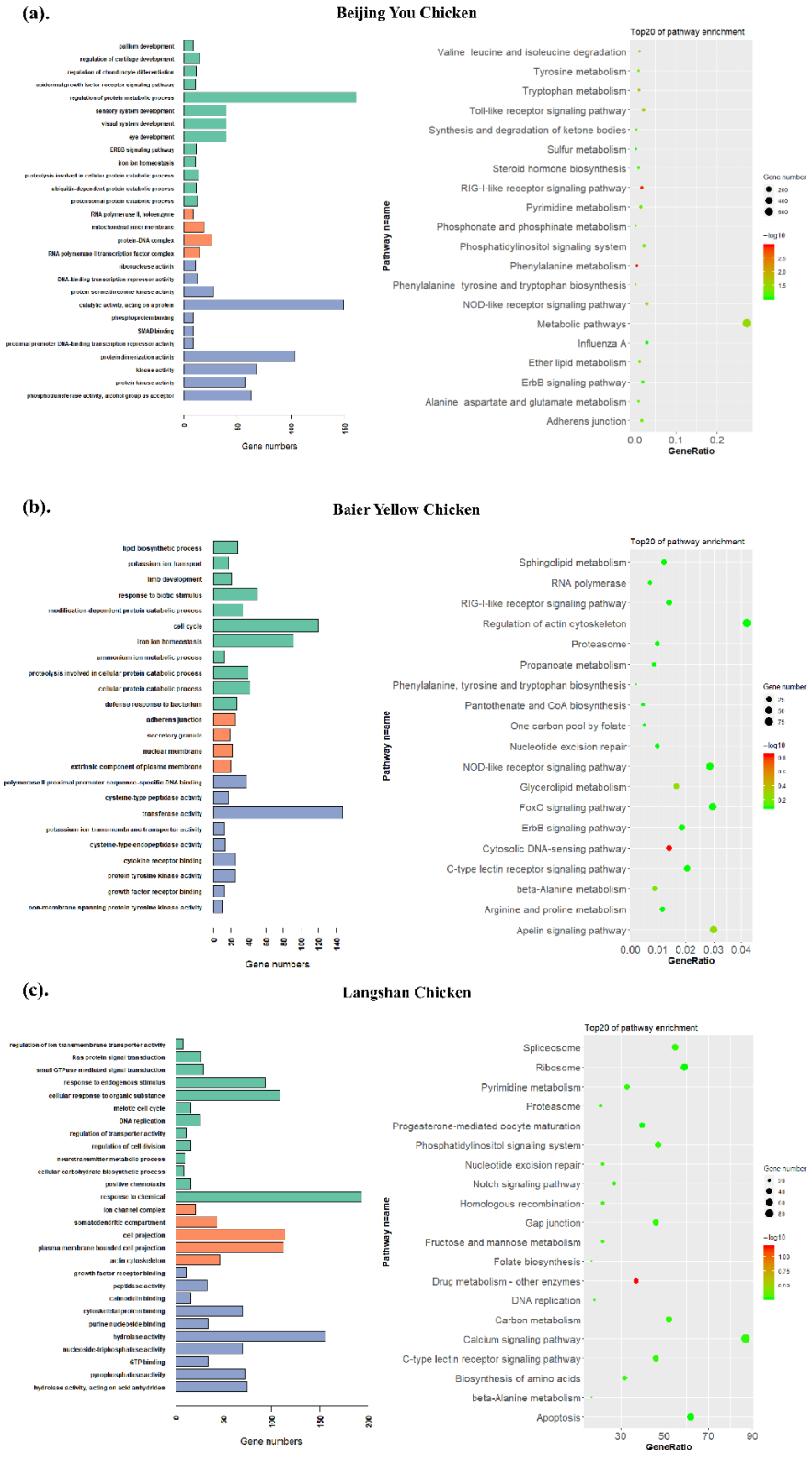

Figure S7