



## OPEN ACCESS

## EDITED BY

Oladeji Bamidele,  
International Livestock Research  
Institute, Nigeria

## REVIEWED BY

Krishnamoorthy Srikanth,  
Cornell University, United States  
Xiquan Zhang,  
South China Agricultural University, China  
Jang-Il Sohn,  
Genome4me (Inc.), Republic of Korea

## \*CORRESPONDENCE

Nagarjun Vijay,  
✉ [nagarjun@iiserb.ac.in](mailto:nagarjun@iiserb.ac.in)

RECEIVED 06 March 2023

ACCEPTED 05 June 2023

PUBLISHED 22 June 2023

## CITATION

Shinde SS, Sharma A and Vijay N (2023),  
Decoding the fibromelanosis locus  
complex chromosomal rearrangement of  
black-bone chicken: genetic  
differentiation, selective sweeps and  
protein-coding changes in  
Kadakhnath chicken.  
*Front. Genet.* 14:1180658.  
doi: 10.3389/fgene.2023.1180658

## COPYRIGHT

© 2023 Shinde, Sharma and Vijay. This is  
an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Decoding the fibromelanosis locus complex chromosomal rearrangement of black-bone chicken: genetic differentiation, selective sweeps and protein-coding changes in Kadakhnath chicken

Sagar Sharad Shinde, Ashutosh Sharma and Nagarjun Vijay\*

Computational Evolutionary Genomics Lab, Department of Biological Sciences, IISER Bhopal, Bhauri, Madhya Pradesh, India

Black-bone chicken (BBC) meat is popular for its distinctive taste and texture. A complex chromosomal rearrangement at the fibromelanosis (*Fm*) locus on the 20th chromosome results in increased endothelin-3 (*EDN3*) gene expression and is responsible for melanin hyperpigmentation in BBC. We use public long-read sequencing data of the Silkie breed to resolve high-confidence haplotypes at the *Fm* locus spanning both Dup1 and Dup2 regions and establish that the *Fm\_2* scenario is correct of the three possible scenarios of the complex chromosomal rearrangement. The relationship between Chinese and Korean BBC breeds with Kadakhnath native to India is underexplored. Our data from whole-genome re-sequencing establish that all BBC breeds, including Kadakhnath, share the complex chromosomal rearrangement junctions at the fibromelanosis (*Fm*) locus. We also identify two *Fm* locus proximal regions (~70 Kb and ~300 Kb) with signatures of selection unique to Kadakhnath. These regions harbor several genes with protein-coding changes, with the bactericidal/permeability-increasing-protein-like gene having two Kadakhnath-specific changes within protein domains. Our results indicate that protein-coding changes in the bactericidal/permeability-increasing-protein-like gene hitchhiked with the *Fm* locus in Kadakhnath due to close physical linkage. Identifying this *Fm* locus proximal selective sweep sheds light on the genetic distinctiveness of Kadakhnath compared to other BBC.

## KEYWORDS

Kadakhnath, black-bone chicken, genetic linkage, fibromelanosis, *Fm* locus

## Introduction

Domestication of chicken during the Neolithic period involved a complex pattern of interbreeding with various jungle fowl species (Eriksson et al., 2008; Tixier-Boichard et al., 2011; Wang et al., 2020a; Lawal and Hanotte, 2021). After domestication, chickens have spread worldwide and occur as commercial, exotic, and indigenous village breeds. Humans use chickens as a research model due to their physiology and behavior (Cogburn et al., 2007), as game fowl and for religious reasons, or more commonly for egg or meat production

(Rubin et al., 2010). Chicken is the most preferred source of meat for humans due to its easy availability and affordability (Kralik et al., 2016). Hence, understanding the genetics of meat features is commercially relevant (Mir et al., 2017). Although understudied, indigenous village chicken breeds with unique properties provide an opportunity to understand the genetics of meat. For instance, the black-bone chicken (BBC) is a delicacy due to its texture, color, firmness, flavor, and use in traditional medicine (Jaturasitha et al., 2008; Dou et al., 2022). The black color of the BBC results from melanin deposition throughout the body, i.e., melanin hyperpigmentation or fibromelanosis caused by the *Fm* allele (Muroya et al., 2000). Bateson and Punnett were the pioneers in identifying the autosomal dominant *Fm* allele (Bateson et al., 1911). Modern studies found a chromosomal rearrangement on chromosome 20 is involved in the *Fm* locus (Dorshorst et al., 2011; Dharmayanthi et al., 2017; Sohn et al., 2018). The overexpression of the endothelin-3 (*EDN3*) gene located within the *Fm* locus is responsible for hyperpigmentation seen in the BBC (Dorshorst et al., 2011; Shinomiya et al., 2012; Zhang et al., 2022).

BBC breeds have anti-fatigue and anti-hypoxic abilities, with their meat having antioxidant properties (Tu et al., 2009; Dou et al., 2022), high carnosine (Tian et al., 2007; Sharma et al., 2022b), and lower fat and cholesterol content (Jaturasitha et al., 2008; Tian et al., 2011). Some BBC breeds also have local adaptations. For instance, the Korean Ogye has improved fetal viability and innate immunity against microbial and viral infections (Cho et al., 2022). BBC breeds occur globally and have distinct names, such as Ayam Cemani (Indonesia), Black H'Mong (Vietnam), Tuzo (Argentina), Svarthöna (Sweden) (Johansson and Nelson, 2015; Dharmayanthi et al., 2017), Yeonsan Ogye (Korea) (Sohn et al., 2018), and Thai BBC (Thailand) (Buranawit et al., 2016). China has a high diversity of BBC breeds, including Silkie, Jiangshan, Lueyang, Sichuan, Xingwen, Yugan, Dehua, Jinhu, Muchuan, Wumeng, Yanjin, Xichuan, Tuanfu, Wuliangshan, Emei, and Miyi fowl (Zhu et al., 2014; Huang et al., 2020b; Li et al., 2020; Kriangwanich et al., 2021; Dou et al., 2022). India has a single breed of BBC, commonly known as Kadaknath (Sharma et al., 2022b).

Kadaknath is considered native to the Jhabua, Alirajpur, and Dhar districts of Madhya Pradesh, but its farming has recently spread across India (Jadhao et al., 2022; Tripathi et al., 2022). The Kadaknath breed has long been documented as a distinctive indigenous Indian breed and is also called Kali Masi or Karaknath (Slater, 1945). At least three distinctive phenotypes (i.e., jetblack, pencil, and golden) occur within the Kadaknath breed (Haunshi and Prince, 2021). In jetblack, all the body parts like plumage, comb, internal organs, eyes, skin, beak, shank, and claw are entirely black, whereas pencil and golden have white and golden color patches on the plumage, respectively (Haunshi and Prince, 2021). Several unique characteristics, such as earlier egg-laying maturity, high-protein content, better disease resistance, and adaptation to the local environment, are attributed to the Kadaknath chicken (Rout et al., 1992; Haunshi et al., 2011; 2022; Thakur et al., 2015; Jena et al., 2018; Shanmathy et al., 2018; Sahu et al., 2019; Haunshi and Prince, 2021; Sehrawat et al., 2021; Sharma et al., 2022b; Jadhao et al., 2022). Understanding the genetics behind these traits will help

establish the uniqueness of Kadaknath and guide de-extinction efforts and breeding programs.

Despite its immense popularity and commercial value, the genomics of the Kadaknath chicken breed has received limited attention. Therefore, further research and genomic analyses are required to understand adaptations in this breed and its genetic history. The aims of this study are as follows: (1) Performing whole-genome re-sequencing of Kadaknath and evaluating its relationship with other black-bone and non-black-bone chicken breeds. (2) Assessing whether Kadaknath and other BBC breeds share a common origin for the *Fm* locus by comparing the chromosomal rearrangement junction and resolving the correct arrangement of duplicated regions at the *Fm* locus. (3) Evaluating how the BBC breeds dispersed to various parts of Asia. (4) Using population genetic statistics to identify signatures of selection in Kadaknath compared to other BBC breeds.

## Materials and methods

### Population sampling

The study was approved by the Institutional Ethics Committee (IEC) of the Indian Institute of Science Education and Research, Bhopal, vide reference number IISERB/IEC/Certificate/2018-11/03 dated 8th June 2018. We purchased the meat of six individuals (two individuals each from jetblack, pencil, and golden morphs (Supplementary Figure S1)) of Kadaknath from an Food Safety and Standards Authority of India (FSSAI)-licensed shop in Bhopal, Madhya Pradesh, India. We procured two other individuals with a black-bone phenotype from the same FSSAI-licensed shop to examine hybrids. One had a golden-pencil external appearance, and another was completely white on the exterior. We also sampled two non-BBCs from the same shop to determine the genetic relationship of Kadaknath with the native-village chicken and broiler reared in the same poultry (Supplementary Figure S1). To reduce the chances of inter-breeding and have relatively pure Kadaknath samples, we obtained three (one individual each from jetblack, pencil, and golden morphs) additional Kadaknath breed chickens from Jhabua, Madhya Pradesh. Whole-genomic DNA with high purity and quality was extracted from the liver tissue samples using DNeasy Blood & Tissue Kits (QIAGEN). We generated >25x coverage whole-genome short-read paired-end data (using Illumina Novaseq) for all 13 individuals sampled (Supplementary Table S1).

To evaluate the relationship of Kadaknath chicken globally, we compared our dataset with publicly available chicken re-sequencing data. High-quality re-sequencing datasets of 88 chicken individuals from other BBCs, commercial chicken lines, and other distinctive chicken breeds with a coverage >20x were selected and obtained from European Nucleotide Archive (ENA) and Korean National Agricultural Biotechnology Information Center (KNABIC) (<https://nabic.rda.go.kr/>). Out of 88 individuals, 23 are BBCs, which we considered for further analysis. Hence, we analyze a dataset of 101 individuals (88 public +13 sequenced as part of this study) from different breeds (Supplementary Table S2 for more details) (Fan et al., 2013; Roux et al., 2014; Wang et al., 2015; Ulfah et al., 2016; Sohn et al., 2018; Qanbari et al., 2019; Li et al., 2020;

Mariadassou et al., 2021; Cho et al., 2022). Transcriptomic datasets from four individuals each for five native Indian breeds (Kadaknath, Ankleshwar, Aseel, Punjab brown, and Nicobari) and broiler chicken from India generated from the breast muscle are publicly available (Supplementary Table S2). We only include these 24 transcriptome samples with the 101 WGS samples to assess the population structure.

## Read mapping, variant calling, and phylogeny

We mapped the paired-end raw reads of 101 individuals to the chicken genome assembly (genome assembly version is *Gallus\_gallus.GRCg6a*) using the BWA (Li and Durbin, 2009) (Burrows–Wheeler aligner) v0.7.17-r1188 mem read mapper with default parameters. We added the read group information using Picard tools and removed duplicate reads in all 101 individual BAM files (<https://github.com/broadinstitute/picard>). We performed the variant calling using FreeBayes (Garrison and Marth, 2012) with different quality control flags such as `-min-alternate-count -C 10`, `-min-mapping-quality -m 20`, `-min-base-quality -q 20`, and `-min-coverage 10`. Similarly, we used the bcftools (Li and Barrett, 2011; Danecek et al., 2021) with mapping quality flags such as `-mapping quality -C 50`, `-min base quality -Q 20`, and `-min mapping quality -q 20` for robust variant calling. We removed the indels from variant calls using VCFtools (Danecek et al., 2011) with the `-remove-indels` flag and extracted the common SNPs using both SNP callers for retaining reliable SNP calls. The single-nucleotide polymorphisms (SNPs) identified by both variant callers (bcftools v1.9 and FreeBayes v1.0.0) were used for subsequent analysis. To identify the effect of genetic variants, we used the snpEff v SnpEff 4.3t and *GRCg6a.96* databases for annotation (Cingolani et al., 2012). For the prediction of the possible impact of non-synonymous fixed variants in BPIL on the protein's structure and function, we used the Polymorphism Phenotyping v2 (PolyPhen-2) tool (Adzhubei et al., 2010). Common SNPs from both vcf files were used to construct phylogeny for 34 and 101 individuals using SNPhylo (Lee et al., 2014) v20180901, based on a maximum likelihood tree with 1,000 bootstrap values. We excluded the scaffolds, Z, W, and MT chromosomes in phylogeny analysis. The local phylogeny for Dup1 and Dup2 regions was generated using the vk phylo command (using both NJ and UPGMA methods) implemented in the VCF-kit (Cook and Andersen, 2017). The Mt genome haplotype median-joining networks were constructed using PopART (Leigh and Bryant, 2015) and SplitsTree (Huson, 1998).

We mapped the paired-end transcriptomic raw read data of 24 individuals to chicken genome assembly (genome assembly version is *Gallus\_gallus.GRCg6a*) using the STAR (Dobin et al., 2013) (v2.7.0d) read mapper with default parameters. We perform the variant calling using bcftools with flags `-mapping quality -C 50`, `-min base quality -Q 20`, and `-min mapping quality -q 20`. Using VCFtools, we extracted 67,617 SNPs only considering autosomes where no data were missing in any individuals using the flag `-max-`

missing 1. For assessing the population structure of all 125 individuals (101 with whole genomic data +24 transcriptomic data), we performed the variant calling only on 67,617 SNPs (extracted from transcriptomic data of 24 individuals) using bcftools.

## *Fm* locus junction identification

The *Fm* locus consists of a complex chromosomal rearrangement composed of two different non-paralogous regions (Dup1 (~127 Kb) and Dup2 (~170 Kb)) separated by an intermediate (Int) region. Dup1 and Dup2 regions are both duplicated and are involved in a complex rearrangement consisting of two junctions: (A) Dup1 + (inverted Dup2) and (B) (inverted Dup1) + Dup2. To identify the base-pair level positions of Dup1, Dup2, and Int regions, we compared the short-read coverage of black and non-black chicken in 1 Kb windows along chromosome 20. We used the makewindows command of bedtools (Quinlan and Hall, 2010) (v2.26.0) to create 1 Kb non-overlapping windows along chromosome 20. The number of reads in each 1 Kb window was calculated using the bedtools coverage command. We shortlisted adjacent windows with drastically different read coverage in black individuals but not in non-black individuals. The base-pair level coordinates of Dup1 and Dup2 in the *Gallus\_gallus.GRCg6a* genome were narrowed down further using coverage estimates in 1bp windows. Dup1 starts at 20:10766772 and ends at 20:10894151. Dup2 occurs further along the chromosome and starts at 20:11306686 and ends at 20:11477501. In addition to Dup1, Dup2, and Int regions, we defined ~500 Kb flanking regions as Flank1 (20:10263555-10766771) and Flank2 (20:11477502-11980000).

## Black-bone-specific *Fm* locus junction

The junction between the rearranged regions in the BBC does not occur in the *Gallus\_gallus.GRCg6a* genome (Supplementary Figure S2, 3). In searching for a completely assembled *Fm* locus, we check the previously published assembly of BBC breeds. We find that the two genome assemblies of Silkie (Silkie2 (GCA\_024679325.10 and Silkie3 (GCA\_024653025.1) generated by Li et al. (2022b) and one genome assembly of Yeosan Ogye generated by Sohn et al. (2018) have been published. We generated the dot plot of chromosome 20 of chicken genome assembly (*Gallus\_gallus.GRCg6a*) with chromosome 20 of Yeosan Ogye (CM008847.1) and Silkie3 (CM045235.1) using the Gepard tool (Krumbsiek et al., 2007). The rearranged *Fm* locus is partially assembled at Chr 20 in Yeosan Ogye (Supplementary Figure S2, 3). In the Silkie3 genome, parts of the *Fm* locus occur on chr20 and unplaced scaffold (JAJMOI010001544.1) (Supplementary Figure S4, 5). However, in Silkie2 genome assembly, chr20 is not assembled, so Dup1 and Dup2 are independently present on four different unplaced scaffolds (JAJMOM010018106.1, JAJMOM010006651.1, JAJMOM01000015 6.1, and JAJMOM010000154.1). None of the existing assemblies provide the full-length assembly of the *Fm* locus in BBC breeds. While the junction sequences have been reported earlier (Dorshorst et al., 2010) for the Silkie breed, it is unclear whether all BBC breeds

share the same junction sequences. Hence, we searched for the *Fm* locus junction in the Korean BBC PacBio data from ENA (SRR6189090). Based on our search of PacBio reads that mapped to the *Gallus\_gallus.GRCg6a* genome, we shortlisted reads that simultaneously aligned to two of the five (Flank1, Dup1, Int, Dup2, and Flank2) genomic regions we have defined. We found several reads spanning Flank1-Dup1, Dup1-Int, Int-Dup2, and Dup2-Flank2, which is unsurprising because these regions are adjacent.

We found five reads (SRR6189090.111279, SRR6189090.56386, SRR6189090.880702, and SRR6189090.387043, and SRR6189090.54824) that span across both Dup1 and Dup2. Dup1 and Dup2 are far apart in their genomic location in *Gallus\_gallus.GRCg6a* assembly. Hence, these reads support a rearrangement that leads to junctions between these two regions. Three of these five reads (SRR6189090.111279, SRR6189090.880702, and SRR6189090.387043) support the junction Dup1 + (Inverted Dup2) (i.e., START-DUP1-END-END-DUP2-START), and the other two reads (SRR6189090.54824 and SRR6189090.56386) support the junction (inverted Dup1) + Dup2 (i.e., END-DUP1-START-START-DUP2-END). The read coverage at these junctions consistently differed between all black-bone and non-black-bone chickens. Based on the read coverage across the *Fm* locus junctions, we identified 34 out of 101 samples analyzed as BBC.

## Long-read data mapping and haplotype-specific read identification

Recently, a new high-coverage multi-platform genomic public dataset for the Silkie BBC became available on the European Nucleotide Archive (ENA) as part of the Bioproject# PRJNA805080 (we thank China Agricultural University for publishing these data). With a ~65X (Nanopore) and >660X coverage (PacBio) of the same Silkie individual, this dataset is well suited to resolve the haplotypes at the *Fm* locus and identify the exact order of rearrangement. For mapping the PacBio and Nanopore long-read data of this Silkie individual, we used the *Gallus\_gallus.GRCg6a* (*galgal6a* genome assembly) genome with the BWA (Burrows-Wheeler aligner) v0.7.17-r1188 bwasmapper with the flags -t 24, -a2, -b3, -q2, -r2, and -z1 to obtain high-quality read mapping.

We used the bam-readcount (Khanna et al., 2022) tool to obtain the number of reads supporting each of the four nucleotide bases at each position along the Dup1 (~127 Kb) and Dup2 (~170 Kb) regions. Potentially heterozygous sites with reliable read support were shortlisted as sites with at least 10 reads supporting each of the two alleles. Sites with at least 10 reads supporting three or more bases were excluded as potentially tri-allelic sites. Using these criteria, we found around 1% of the sites that are potentially heterozygous in Dup1 (1,214/127000) and Dup2 (1774/170000) regions. Subsequent steps were performed using the phasing.sh script (provided on our GitHub page). For each of these 2,988 sites, we labeled the reads based on the nucleotide base (i.e., A, T, G, and C) present in that read at each position using the biostar214299 program from jVarkit (Lindenbaum, 2015). We prepared the list of reads supporting each of the four bases for each site using the reads labeled in the

previous step. Starting from the sites (i.e., 10766895 for Dup1 and 11476819 for Dup2) identified by visual inspection of the long-read alignments in IGV, we extended the haplotypes using the phasing.sh script. The rationale of the script is to find the next site which can distinguish the haplotypes with at least five reads supporting each haplotype. The number of reads supporting each of the four nucleotide bases at the two sites is calculated for every pair of sites the script considers. Hence, the number of reads supporting each of the 16 possible combinations of the four bases at the two sites is counted. We used this script sequentially to extend the haplotypes by identifying haplotype-defining sites by manually inspecting these counts. Of the 2,988 sites considered, 49 haplotype-defining sites were sufficient to span entire Dup1 and Dup2 regions (Supplementary Table S3).

## Principal component analysis and admixture

Principal components analysis (PCA) was performed using PCAngsd (Meisner and Albrechtsen, 2018) based on genotype likelihood estimates from Analysis of Next-Generation Sequencing Data (ANGSD) v0.935 (Korneliusen et al., 2014). PCA was performed both genome-wide and *Fm* locus region-wise (Flank1, Dup1, Int, Dup2, and Flank2) for the 34 black-bone and all 101 chicken individuals. We used several flags in ANGSD for population structure analysis as follows: -GL 2, -doMaf 1, -minMapQ 30, -minQ 20, -doGlf 2, and -SNP\_pval < 1e-6. Genotype likelihood values from ANGSD were used to identify principal components using PCAngsd and genome-wide admixture proportions using NGSadmix (Skotte et al., 2013). NGSadmix was run for different values of K from K1 to K10 using each K with 15 iterations with flags -minMaf 0.05 and -minInd as specified. Admixture analysis suggests Kadaknath has a sub-structure but remains distinct from the other black-bone breeds at best K = 7 using median values of Ln (Pr Data) or log probability of the data (referred to as Ln Pr (X|K), i.e., k for which Pr(K = k) is highest, while according to ΔK (Evanno's best K method), the best K is 8. We used the Evanno (Evanno et al., 2005) method implemented in the Clumpak (Kopelman et al., 2015) web server to find the best K.

Using the variant calls of 125 individuals for 67,617 SNPs, we performed the admixture analysis through AdmixPipe v3 (Mussmann et al., 2020) for K1–K15, and we ran 10 iterations for each K. During the admixture analysis, two individuals of the Brown layer (BROL1 and BROL2, Supplementary Table S2) got excluded due to the poor quality of variant calls. In the admixture analysis of 123 individuals, we found that according to Evanno's best K method, the best K is 2. The observation of K = 2 in admixture analysis using ΔK (Evanno's best K method) is reported in more than 50% of recent studies (Janes et al., 2017). Hence, we followed the recommendations of Gilbert et al., (2012) and Janes et al. (2017) to assess the population structure. We re-ran the admixture pipeline for 109, 100, and 88 individuals after removing the populations (Rhode Island Red, White Leghorn, and broiler), respectively, which separated at K = 2. As per the suggestion by Janes et al. (2017) to infer the optimal K value, we are providing the Ln Pr (X|K) and ΔK plots and bar plots for the multiple K values in supplementary. PCA was also performed using PCAngsd for *Fm* locus region-wise



(Flank1, Dup1, Int, Dup2, and Flank2) using 123 individuals (BROL1 and BROL2 are excluded).

## Population genetic analysis

Chinese black-bone breeds have diverged to differing extents from Kadaknath and have limited sample sizes. Hence, we combined the individuals from closely related XBBC (Xichuan black-bone chicken), LCEM (Emei black fowl), and LCMY (Miyi black fowl) breeds into a single population representative of Chinese black-bone (CHIN,  $n = 9$ ) chicken and JETB, PENC, and GOLD into another population representative of Kadaknath (KADK,  $n = 9$ ) (Supplementary Table S4). We calculated population genetic statistics genome-wide to identify signatures of selection. However, to avoid false positives, we excluded genomic regions (50 Kb windows with <80 percent callable sites) with poor callability. We used the CallableLoci walker of GATK (McKenna et al., 2010) on the BAM files to quantify callability. For identifying callable regions from mapped BAM files for CHIN and KADK individuals, we used GATK with different flags such as `-minMappingQuality 20`, `-minBaseQuality 20`, `-minDepth 10`, `-minDepthForLowMAPQ 20`, and `-maxFractionOfReadsWithLowMAPQ 20`. Furthermore, we filtered 50 Kb windows with >0.1 repeat element fraction to rule out the possibility of false positives (Supplementary Figure S6–8). We used stringent coverage and quality criteria (`-GL 2`, `-dosaf 1`, `-baq 1`, `-C 50`, `-setMinDepthInd 6`, `-minInd 3, 4 or 9`, `-minMapQ 30`, `-minQ 20`, and `-doCounts 1`) in ANGSD to calculate all population genetic statistics. The list of individuals in each population and the population pairs compared is in Supplementary Table S4. We found that at least 20,000 of the 21,659 50 Kb windows had sufficient high-quality data in all populations.

We used the folded site frequency spectrum (SFS) approach implemented in ANGSD to calculate genome-wide population-specific estimates of  $\pi$  (the average pairwise differences), Watterson's  $\theta$  (the average number of segregating sites),  $\tau$  (Tajima's D), and Fu and Li's D. We estimated inter-population genomic differentiation ( $F_{ST}$ ) and divergence ( $D_{xy}$ ) using ANGSD and `popgenWindows` python script ([https://github.com/simonhmartin/genomics\\_general](https://github.com/simonhmartin/genomics_general)) for each population pair. The  $F_{ST}$  estimates from the two methods were highly correlated (Pearson's  $\rho = 0.98$ ,  $p$ -value <2.2  $e^{-16}$ ) (Supplementary Figures S9,10). We defined  $F_{ST}$  outlier regions as 50 Kb windows in the top 1 percent of the genome-wide estimates and merged adjoining windows using bedtools. Similarly, the 50 Kb windows in the top 10 percent of the genome-wide estimates of  $D_{xy}$  were deemed to have high levels of divergence. We identified fixed sites as those sites with  $F_{ST} > 0.9$ . We estimate haplotype-based statistics  $iHS$  and  $XP-EHH$  in the `rehh` (Gautier and Vitalis, 2012) R package with option `polarized = FALSE` using genotype data phased with SHAPEIT (Delaneau et al., 2011).

## Results

### Whole-genome re-sequencing

We generated >25X coverage whole-genome re-sequencing Illumina data for 13 chicken individuals from India. Our

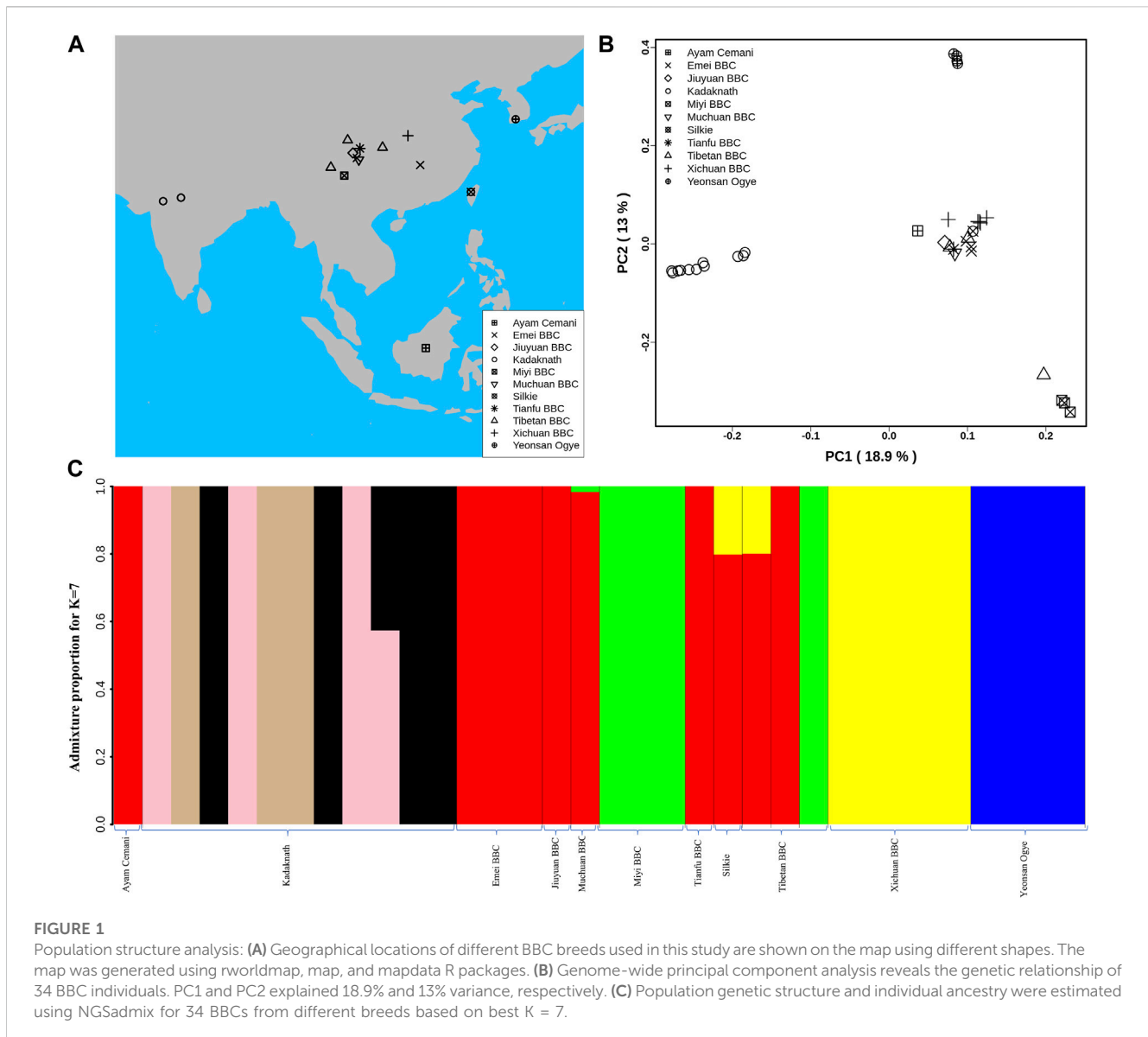
collection includes nine Kadaknath samples as representatives of all three extant morphs jetblack (JETB  $n = 3$ ), golden (GOLD  $n = 3$ ), and pencil (PENC  $n = 3$ ). We also sampled one individual with a golden–pencil-like phenotype (GOPE  $n = 1$ ), one with complete white plumage with black bone (CROS  $n = 1$ ), one non-black-bone individual where plumage was black (NONB), and one individual of broiler breed (Supplementary Figure S1; Supplementary Table S1). We compared these Kadaknath chicken genomes with public re-sequencing datasets of chicken breeds. The sampling location (Figure 1A for black-bone chicken) of individuals analyzed in this study is spread across Asia (for more detail, see Supplementary Table S2, which includes the complete list of all chicken breeds examined). A detailed map of China depicts the locations of all the Chinese black-bone breeds (Supplementary Figure S11).

### Kadaknath is a distinct breed of black-bone fowl

The major axes of genetic variation in the BBC (PC1:18.9% and PC2:13%) separated the Kadaknath, Yeosan Ogye (YOSK  $n = 4$ ), and Chinese BBC ( $n = 17$ ) (see Figure 1B; Supplementary Figure S12). We observed that all Kadaknath individuals form a single cluster distinct from other black-bone breeds (from Indonesia, China, and South Korea) in the PCA (Figure 1B; Supplementary Figure S12). Similarly, in the genome-wide PCA analysis of 101 individuals consisting of both BBCs and non-BBCs, Kadaknath forms a separate cluster from other commercial, native, and BBC breeds in different PC comparisons (Supplementary Figure S13A–C).

Admixture analysis suggests that Kadaknath has sub-structure but remains distinct from the other black-bone breeds at best  $K = 7$  (see Figure 1C and Supplementary Figure S14A–C; Supplementary Figure S15). In Chinese BBC, the admixture analysis identified three groups supporting the PCA clustering pattern, while YOSK forms a separate cluster (see Figure 1C; Supplementary Figure S12). We observed lower genetic diversity ( $\pi = 0.002$ ,  $\theta = 0.002$ ) in YOSK ( $n = 4$ ), while Kadaknath ( $n = 9$ ) and CHIN ( $n = 9$ ) BBC have comparable genetic diversity ( $\pi = 0.004$  and  $0.003$ ,  $\theta = 0.003$  in both) (Supplementary Figure S16A, B). Genetic diversity can be affected by crossbreeding with native chicken breeds. Genome-wide admixture analysis of 101 individuals consisting of both BBCs and non-BBCs supports gene flow between Chinese BBC and native Chinese non-BBC breeds (Supplementary Figure S17A–C). Crossbreeding between Ogye and broiler is also supported by the genome-wide admixture analysis. Including native Indian non-BBC breeds in the admixture analysis suggests crossbreeding between Kadaknath and Ankleshwar breeds (Supplementary Figure S18A–C). However, Kadaknath is distinct from other native Indian breeds, and the Kadaknath samples from the transcriptome dataset cluster with the genome sequencing data generated in this study (Supplementary Figure S19–21A–C).

In contrast to the nuclear genome, the mitochondrial genome haplotype network did not separate Kadaknath from other BBC breeds (Supplementary Figure S22, 23). The *Fm* locus region on chromosome 20, which codes for the black-bone phenotype, is the defining feature of all BBCs. Even after excluding chromosome 20,



the PCA of the remaining chromosomes finds that Kadaknath is genetically distinct from other BBC breeds (Supplementary Figure S24A–C). Hence, the genetic distinctiveness of Kadaknath is spread across the entire genome.

### *\*Fm<sub>2</sub>* is the correct arrangement of duplicated regions at the *Fm* locus

In the non-black chicken, Dup1 (~127 Kb), Int (~412 Kb), and Dup2 (~170 Kb) regions occur in a single copy, are arranged sequentially, and are flanked by Flank1 (~500 Kb) and Flank2 (~500 Kb) regions (Figure 2A). While the Dup1 region contains five protein-coding genes (*EDN3*, *ZNF831*, *SLMO2*, *ATP5E*, and *TUBB1*), the Dup2 region probably consists of only long-non-coding RNA genes. All non-BBCs have a single copy of this region, referred to as *\*N* locus (Figure 2A). The corresponding locus in the BBC is known as the *Fm* locus. The *Fm* locus consists of

three different non-paralogous regions (Dup1, Int, and Dup2) that form a complex chromosomal rearrangement in which both Dup1 and Dup2 regions are duplicated, giving rise to two junctions: (A) Dup1 + (inverted Dup2) and (B) (inverted Dup1) + Dup2. Although the exact ordering of these regions in the rearrangement is not conclusively established, both Dup1 and Dup2 regions are known to be duplicated due to these regions having a sequencing coverage that is twice the genomic average (Dorshorst et al., 2011). The presence of Dup1 + (inverted Dup2) and (inverted Dup1) + Dup2 junctions has been verified in several BBC breeds (Dorshorst et al., 2011; Dharmayanthi et al., 2017; Sohn et al., 2018). Based on this information, three possible scenarios have been proposed by earlier studies (Figure 2B). The *\*Fm<sub>2</sub>* scenario is supported based on crosses between black and non-black-bone chicken (Dorshorst et al., 2011). However, both *\*Fm<sub>2</sub>* and *\*Fm<sub>3</sub>* scenarios require two rearrangement events compared to a single rearrangement event needed for the *\*Fm<sub>1</sub>* scenario (Sohn et al., 2018).

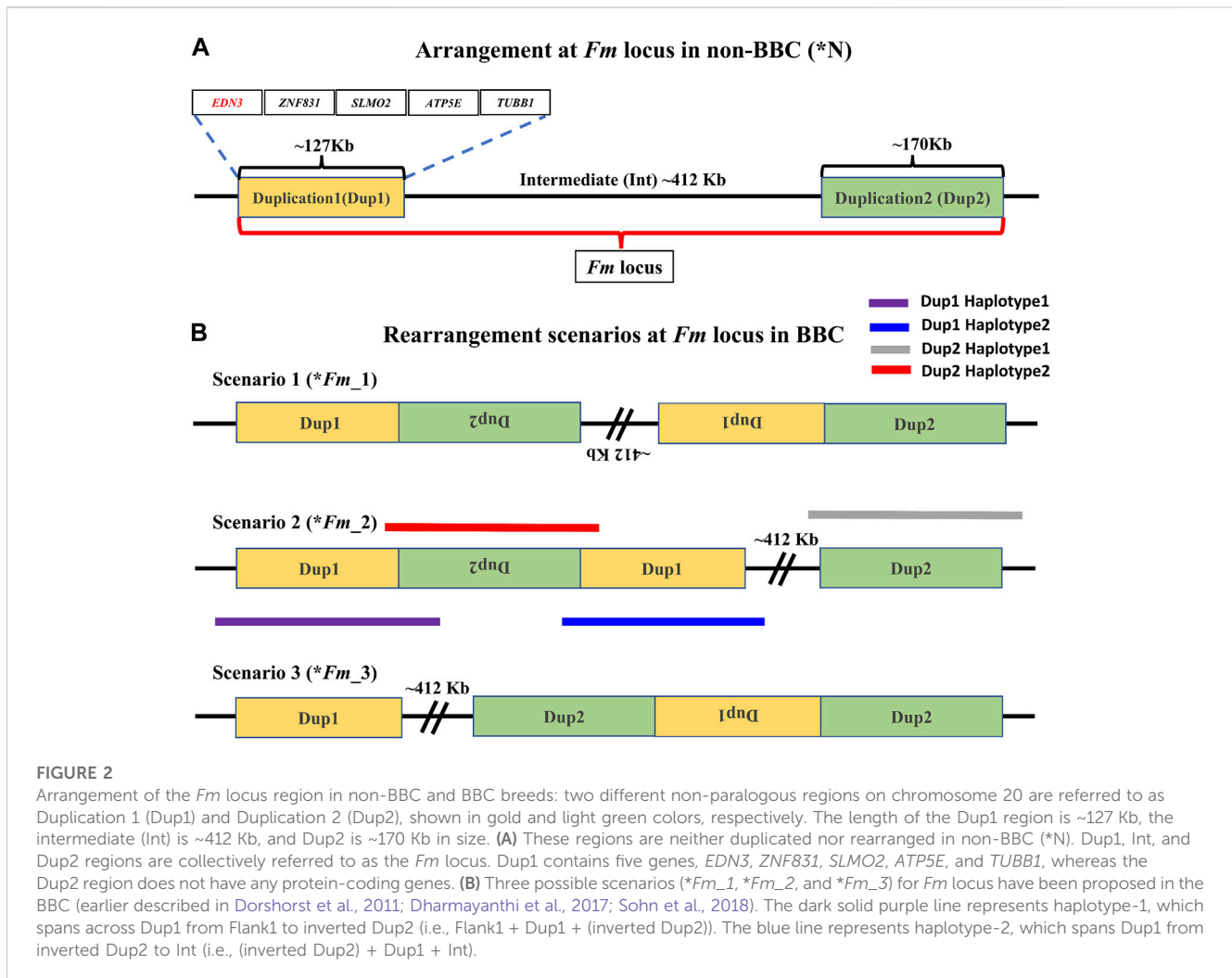
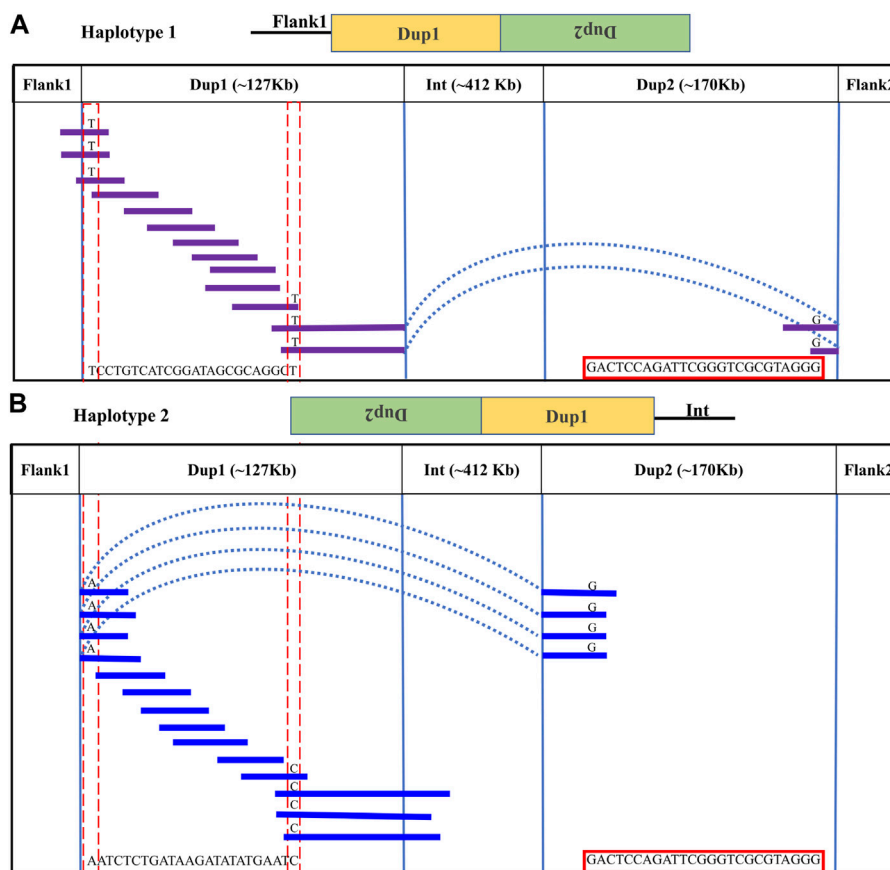


FIGURE 2

Arrangement of the *Fm* locus region in non-BBC and BBC breeds: two different non-paralogous regions on chromosome 20 are referred to as Duplication 1 (Dup1) and Duplication 2 (Dup2), shown in gold and light green colors, respectively. The length of the Dup1 region is ~127 Kb, the intermediate (Int) is ~412 Kb, and Dup2 is ~170 Kb in size. (A) These regions are neither duplicated nor rearranged in non-BBC (\*N). Dup1, Int, and Dup2 regions are collectively referred to as the *Fm* locus. Dup1 contains five genes, *EDN3*, *ZNF831*, *SLMO2*, *ATP5E*, and *TUBB1*, whereas the Dup2 region does not have any protein-coding genes. (B) Three possible scenarios (\**Fm*<sub>1</sub>, \**Fm*<sub>2</sub>, and \**Fm*<sub>3</sub>) for *Fm* locus have been proposed in the BBC (earlier described in Dorshorst et al., 2011; Dharmayanthi et al., 2017; Sohn et al., 2018). The dark solid purple line represents haplotype-1, which spans across Dup1 from Flank1 to inverted Dup2 (i.e., Flank1 + Dup1 + (inverted Dup2)). The blue line represents haplotype-2, which spans Dup1 from Flank1 to Dup2 (i.e., (inverted Dup2) + Dup1 + Int).

Distinguishing between these three (\**Fm*<sub>1</sub>, \**Fm*<sub>2</sub>, and \**Fm*<sub>3</sub>) scenarios requires long-range connectivity information such as long-read sequencing data (PacBio, Nanopore, Synthetic long reads, etc.), Hi-C (high-resolution chromosome conformation capture) (Kronenberg et al., 2021; Wang et al., 2022), or optical mapping (Weissensteiner et al., 2020). Both \**Fm*<sub>1</sub> and \**Fm*<sub>2</sub> contain the same adjacent regions for Dup1 (defined as haplotype-1 and haplotype-2 in Figure 2B) inconsistent with \**Fm*<sub>3</sub>. Hence, long-range information that can span the entire ~127 Kb Dup1 region in a single read will be able to distinguish between \**Fm*<sub>1</sub> and \**Fm*<sub>2</sub> vs. \**Fm*<sub>3</sub>. However, to differentiate between \**Fm*<sub>1</sub> and \**Fm*<sub>2</sub> scenarios, we must span the entire ~170 Kb Dup2 region in a single read. Spanning Dup1 or Dup2 regions with PacBio (average read lengths of 10–25 Kb) or Nanopore (average read lengths of 10–30 Kb) technologies with a single read is challenging but not impossible as reads longer than 1 Mb are possible (Amarasinghe et al., 2020; Hon et al., 2020). Without such individual reads that can span the entire region, it is possible to perform read-based phasing to infer haplotypes that extend to distinct junctions (Patterson et al., 2015). Using a public long-read dataset, we have inferred distinct high-confidence Dup1 and Dup2 haplotypes (supporting the \**Fm*<sub>2</sub> scenario).

Upon visual inspection of the long-read alignments in IGV at the Flank1–Dup1 junction region, three heterozygous sites (10766895, 10766948, and 10767078 at chromosome 20) were recognized at the Dup1 start region potentially separating the two haplotypes of Dup1 (Supplementary Figure S25). We extended these haplotypes toward the end of Dup1 by relying upon overlapping long-reads at haplotype-defining sites (Methods). To ensure the robustness of our read-backed phasing approach, we required that more than 10 sequencing reads support haplotype-specific alleles at each pair of adjacent sites (Supplementary Table S3). Dup1 haplotypes are anchored by (Flank1 end)/(Dup2 start) at the Dup1 start and (Int start)/(Dup2 end) at the Dup1 end (Figure 3). The Flank1 + Dup1 junction containing haplotype is referred to as Dup1 haplotype-1 (D1H1) and contains the allele T at position 10766895. The (inverted Dup2) + Dup1 junction containing haplotype is referred to as Dup1 haplotype-2 (D2H2) and contains the allele A at position 10766895. Each haplotype can be distinguished at 24 sites (Supplementary Table S3) along the Dup1 region. Each haplotype-defining site shares more than 10 long reads with haplotype-specific alleles at adjacent positions. Dup1 haplotype-1 (Figure 3A; Supplementary Figure S26A, B) spans Dup1 from Flank1 to Dup2 haplotype-2 (D2H2) end, and haplotype-2



**FIGURE 3**

Long-read-based haplotypes resolve the sequence spanning the Dup1 region at *Fm* locus: the two haplotypes spanning the Dup1 region are distinguished by distinct alleles at 24 positions using long sequencing reads. Red dotted vertical boxes highlight the alleles that differ between haplotype-1 and haplotype-2 at the same positions. The alleles at 24 sites that separate these two haplotypes at various positions along Dup1 are presented sequentially between the red dotted boxes. **(A)** Haplotype-1 of Dup1 spans from Flank1 to inverted Dup2 (i.e., Flank1 + Dup1 + (inverted Dup2)). The purple lines represent overlapping reads of Nanopore containing Dup1 haplotype-1 alleles. The light blue dotted line represents the span of the same read from the end of haplotype-1 of Dup1 to the end of haplotype-2 of Dup2. **(B)** Haplotype-2 spans Dup1 from haplotype-2 of Dup2 to the Int region (i.e., (inverted Dup2) + Dup1 + Int). The blue lines represent the overlapping reads of Nanopore containing haplotype-2 alleles. The light blue dotted line represents the span of the same read from the start of haplotype-2 of Dup1 to the start of haplotype-2 of Dup2. Both haplotypes of Dup1 are connected to the single haplotype of Dup2 (i.e., haplotype-2 of Dup2) but at different ends.

(Figure 3B; Supplementary Figure S26A, B) spans Dup1 from Dup2 haplotype-2 (D2H2) start to Int start. Our inference of Dup1 haplotype-1 (i.e., Flank1 + Dup1 + (inverted Dup2)) and haplotype-2 (i.e. (inverted Dup2) + Dup1 + Int) sequences rule out the \**Fm*\_3 scenario, suggesting that either \**Fm*\_1 and \**Fm*\_2 scenario is possible.

In the Dup2 region, one haplotype was anchored to the Dup2+ Flank2 junction and another haplotype to the Dup2 + (inverted Dup1) junction. The Dup2 + Flank2 junction containing the haplotype is Dup2 haplotype-1 (D2H1) and contains the allele A at position 11476819. The Dup2 + (inverted Dup1) junction containing the haplotype is Dup2 haplotype-2 (D2H2) and contains the allele G at position 11476819 (Supplementary Figure S27). To verify the correct arrangement, we used the same read phasing approach on the Dup2 region, and we could distinguish the two haplotypes at the Dup2 region based on 25 sites (Supplementary Table S3). Haplotype-1 of the Dup2 region (Figure 4A; Supplementary Figure S28, 29) spans Dup2 from the end of Int

to the start of Flank2, and haplotype-2 (Figure 4B; Supplementary Figure S28, 29) spans Dup2 from Dup1 haplotype-2 (D1H2) start to Dup1 haplotype-1 (D1H1) end. Our inference of haplotype-1 of Dup2 (i.e., Int+ Dup2 + Flank2) and haplotype-2 (i.e., start of Dup1 + (inverted Dup2) + Dup1 end) sequences rule out the \**Fm*\_1 and \**Fm*\_3 scenarios, suggesting that the \**Fm*\_2 scenario is correct (Figure 5).

### All black-bone chicken breeds share the same rearrangement junctions at the *Fm* locus

Conclusive inference regarding which scenario is present in the BBC has been extremely challenging due to the large (~1 Mb) size and complexity of the rearrangement. However, in the case of Kadaknath, it has not even been established whether Dup1 + (inverted Dup2) and (inverted Dup1) + Dup2 junctions identified



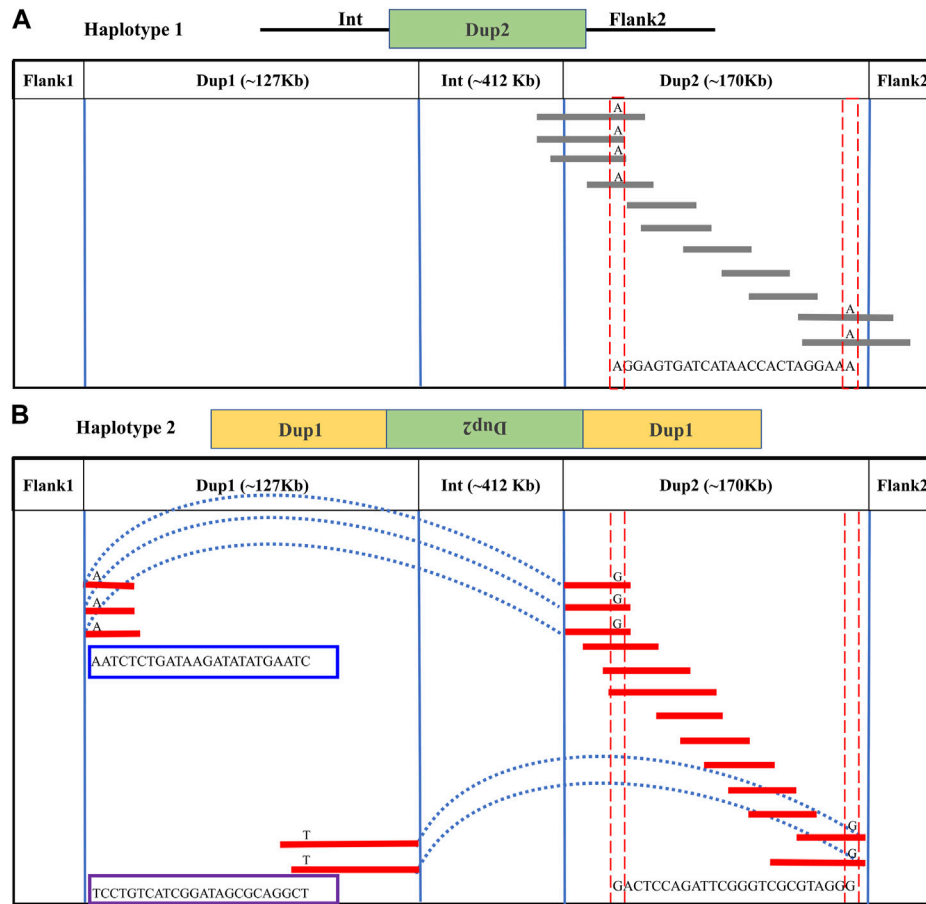


FIGURE 4

Long-read-based haplotypes resolve the sequence spanning the Dup2 region at *Fm* locus: The two haplotypes spanning the Dup2 region are distinguished by distinct alleles at 25 positions using long sequencing reads. Red dotted vertical boxes highlight the alleles that differ between haplotype-1 and haplotype-2 at the same positions. The alleles at 25 sites that separate these two haplotypes at various positions along Dup2 are presented sequentially between the red dotted boxes. (A) Haplotype-1 of Dup2 spans from Int to Flank2 (i.e., Int + Dup2 + Flank2). The gray lines represent overlapping reads of Nanopore containing Dup2 haplotype-1 alleles. (B) Haplotype-2 of Dup2 spans from the start of haplotype-2 of Dup1 to the end of haplotype-1 of the Dup1 region (i.e., Dup1 +(inverted Dup2)+ Dup1). The red lines represent the overlapping reads of Nanopore containing Dup2 haplotype-2 alleles. The light blue dotted line represents the span of the same reads from the start of haplotype-2 of Dup1 to the start of haplotype-2 of Dup2. Similarly, the same reads span the end of haplotype-1 of Dup1 to the end of haplotype-2 of Dup2.

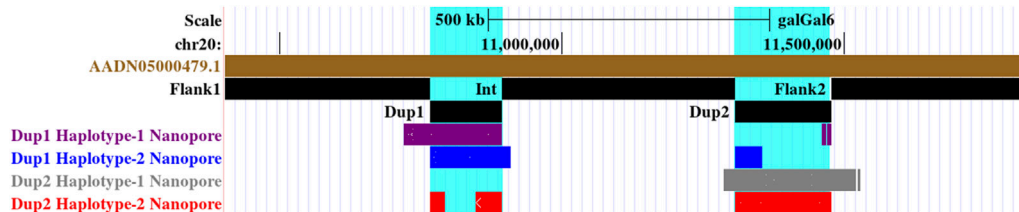
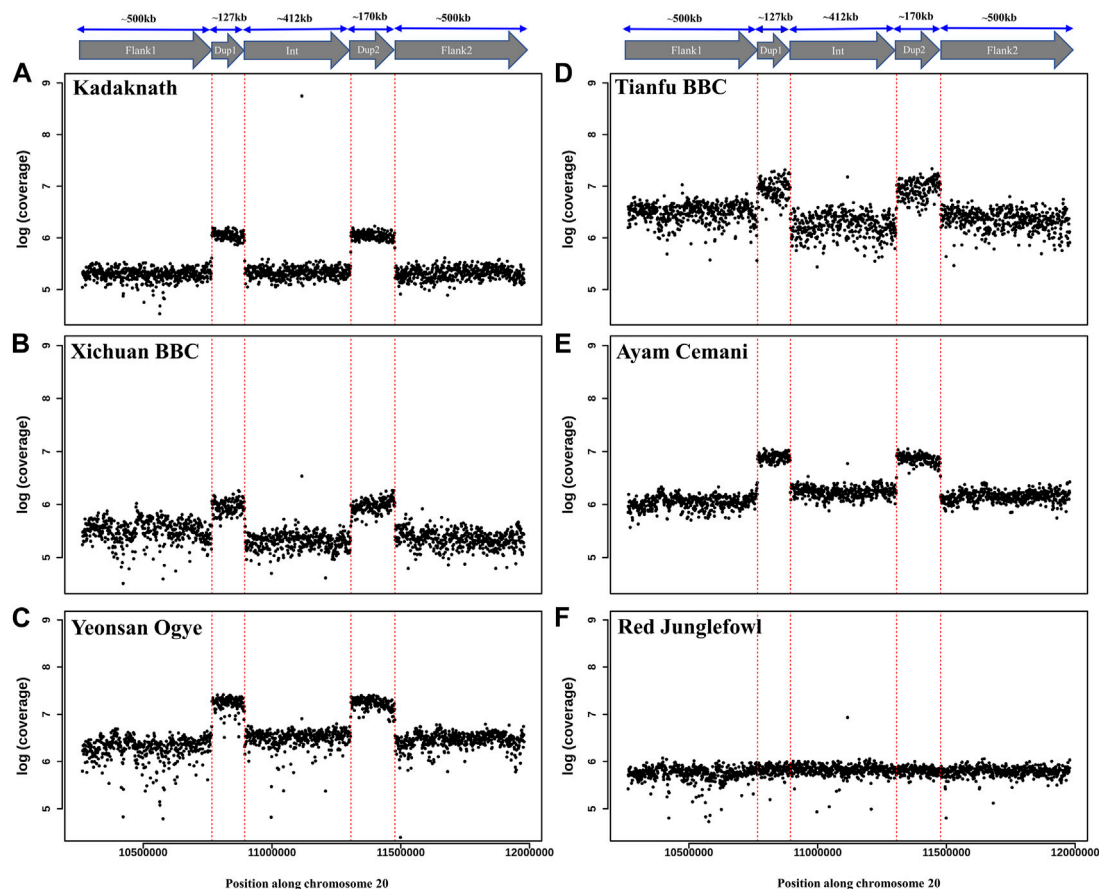


FIGURE 5

Screenshot of the UCSC genome browser of two haplotypes at Dup1 and Dup2 regions highlighted in cyan color identified using Nanopore long reads. Haplotype-1 of Dup1, shown in purple, spans from Flank1 to the end of inverted Dup2 (i.e., Flank1 + Dup1 + (inverted Dup2)). Haplotype-2 of Dup1, shown in blue, spans from the start of inverted Dup2 to the Int region (i.e., (inverted Dup2) + Dup1 + Int). Haplotype-1 of Dup2, shown in gray, spans from Int to Flank2 (i.e., Int + Dup2 + Flank2). Haplotype-2 of Dup2, shown in red, spans from the start of haplotype-2 of Dup1 to the end of haplotype-1 of the Dup1 region (i.e., Dup1 +(inverted Dup2)+ Dup1).

in other BBCs are present. We have compared the normalized short-read coverage at these junctions to evaluate whether all BBC breeds share the same rearrangement junctions at the *Fm*

locus. The normalized short-read coverage in BBCs (Figures 6A–E) abruptly increases at Dup1 and Dup2 regions, while no such increase occurs in non-BBCs (Figure 6F). The drastic



**FIGURE 6**

Comparison of *Fm* locus read coverage in black-bone vs. non-black-bone chicken breeds: read coverage along chromosome 20 at the *Fm* locus is shown in 1 Kb sliding windows. Two duplicated genomic loci, Dup1 and Dup2, are denoted by vertical dotted red lines and have a higher coverage in BBC breeds: (A) Kadaknath, (B) Xichuan, (C) Yeosan Ogye, (D) Tianfu, (E) Ayam Cemani than non-BBC, and (F) red jungle fowl (RJF).

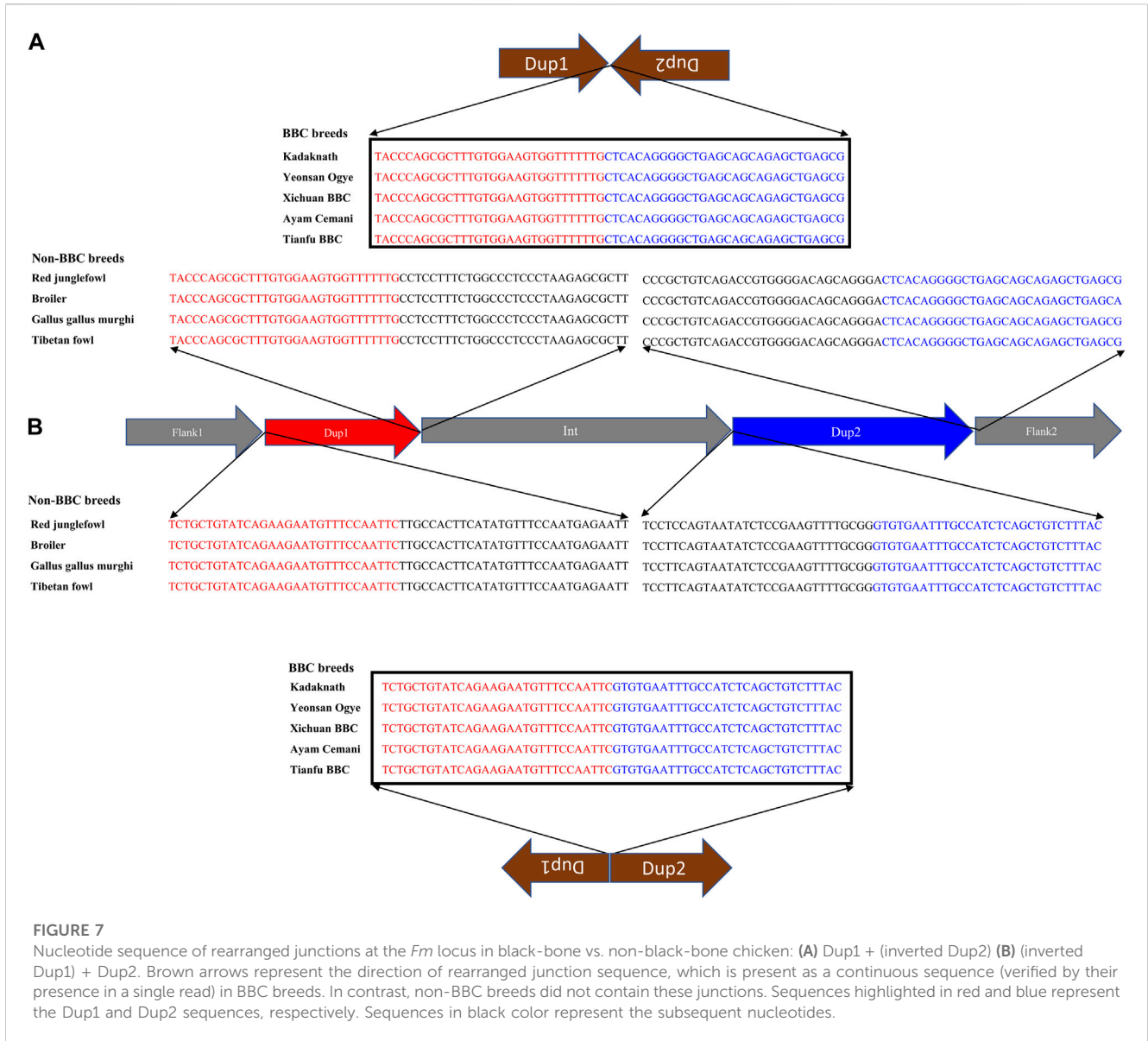
change in coverage at Dup1 and Dup2 boundaries occurs at the same base in all BBCs (Supplementary Figure S30–34). In the case of non-BBCs, no change in coverage occurs at this position (Supplementary Figure S35). We also identified the black-bone-specific *Fm* locus junctions Dup1 + (inverted Dup2) and (inverted Dup1) + Dup2 using published PacBio data (Figures 7A,B and Methods). Short-read coverage spanning these black-bone-specific *Fm* locus junctions is present in the BBC and missing in the non-BBC (Supplementary Figure S36–38). We can infer that BBC defining *Fm* locus originated through a single event based on coverage and rearranged junction sequence.

Crossbreeding with native and commercial breeds affects the genome-wide patterns of genetic variation in black-bone breeds (Supplementary Figure S17–21A–C; Supplementary Figure S39, 40). Hence, we focused on the *Fm* locus to investigate the history of the BBC breeds. The  $F_{ST}$  between the BBC and non-BBC along chromosome 20 is elevated at Dup1 and Dup2 regions (see Supplementary Figure S41). However, other population genetic parameters, such as genetic diversity ( $\pi$  and  $\theta$ ) and divergence ( $D_{xy}$ ), lack any prominent signatures (Supplementary Figure S42–44). As evident from the elevated  $F_{ST}$ , the major axis of genetic variation (assessed using local PCA) in Dup1 and Dup2 regions separates the BBC from non-BBCs (Supplementary

Figure S45, 46A–E). A phylogenetic tree of the SNPs from Dup1 and Dup2 regions also largely separates the BBC breeds from non-BBC breeds (Supplementary Figure S47A, B). Our evaluation of the genetic differentiation ( $F_{ST}$ ) landscape between BBC breeds found reduced  $F_{ST}$  at Dup1 and Dup2 regions compared to the genomic background (Supplementary Figure S48). However, comparing individual BBC breeds with non-black chicken breeds showed the opposite pattern with elevated  $F_{ST}$  at Dup1 and Dup2 regions (Supplementary Figure S49, 50). Hence, the patterns of genetic differentiation ( $F_{ST}$ ), local PCA, and phylogenetic tree also support a common origin of the *Fm* locus in all BBC breeds.

## Isolation by distance pattern suggests dispersal between India and China

Our analysis discovered that the rearrangement junctions in all BBCs are identical to the one in Kadaknath and strongly support a common origin for the *Fm* locus. An independent origin for the *Fm* locus would mean separate rearrangement events have created identical junction sequences. Given the lack of repeat sequences at the junctions, such independent origins

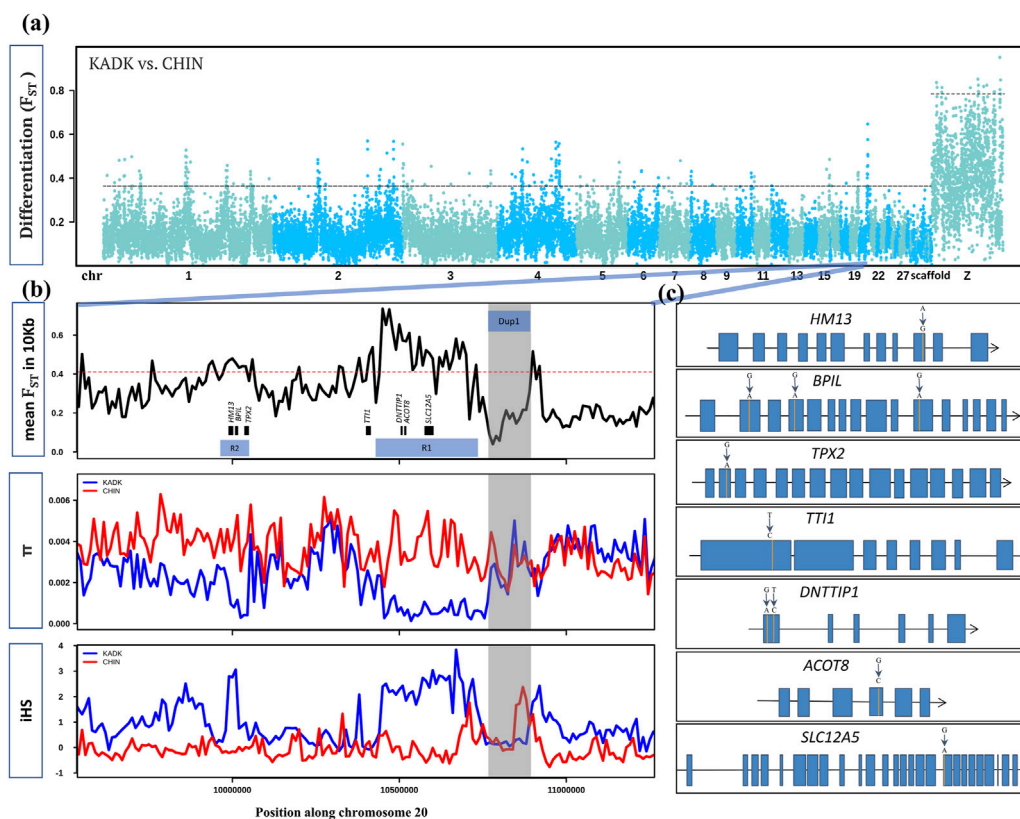


**FIGURE 7**  
 Nucleotide sequence of rearranged junctions at the *Fm* locus in black-bone vs. non-black-bone chicken: **(A)** Dup1 + (inverted Dup2) **(B)** (inverted Dup1) + Dup2. Brown arrows represent the direction of rearranged junction sequence, which is present as a continuous sequence (verified by their presence in a single read) in BBC breeds. In contrast, non-BBC breeds did not contain these junctions. Sequences highlighted in red and blue represent the Dup1 and Dup2 sequences, respectively. Sequences in black color represent the subsequent nucleotides.

seem highly unlikely and are not consistent with the genetic relatedness of BBC and non-BBC breeds at the *Fm* locus. Hence, the current distribution of BBC breeds across Asia needs an explanation. We hypothesized (see a schematic of the proposed scenario in [Supplementary Figure S51](#)) that the dispersal of BBC occurred after a common origin for the *Fm* locus, followed by recent crossbreeding with native and commercial breeds.

To test our hypothesis and delineate the dispersal route of different black-bone breeds, we first evaluated if a pattern of isolation by distance (IBD) is prevalent. The IBD plot shows a consistent increase in genetic distance with an increase in the geographic distance ([Supplementary Figure S52](#), mantel's  $r = 0.64$ ,  $p$ -value = 0.0002). The IBD pattern persisted even after we repeated the analysis using ANGSD/variant-call-based estimates of nucleotide differentiation ( $F_{ST}$ ) after excluding various populations to avoid errors due to auto-correlation and biased estimates from isolated populations ([Supplementary Figure](#)

[S53–59](#)). Among the black-bone breeds, Tibetan black-bone (TBTC) and Sichuan black-bone (LCTMJ) chickens are genetically closest to Kadakhnath (mean  $F_{ST} \sim 0.17$ ). The other black-bone breeds that are geographically more distant from Jhabua (LCEM, mean  $F_{ST} = 0.21$ ; XBBC, mean  $F_{ST} = 0.22$ ; LCMY, mean  $F_{ST} = 0.3$ ; and YOSK, mean  $F_{ST} = 0.36$ ) occur at increasing genetic distances from Kadakhnath ([Supplementary Table S5](#)). The pattern of IBD spanning India and China suggests potentially human-mediated dispersal. We lack conclusive evidence to identify the direction of dispersal. However, the analysis of our dataset, even after excluding alleles found in dbSNP (Sherry et al., 2001) ([Supplementary Figure S60, 61](#)), found that Kadakhnath has almost twice the number of private alleles than Chinese and Korean BBC breeds and suggests an India-to-China dispersal. More extensive fine-scale sampling may provide a definitive answer regarding the direction and timing of the dispersal.



**FIGURE 8**

Within and between population comparison of KADK with CHIN: **(A)** genome-wide landscape of pairwise genetic differentiation ( $F_{ST}$ ) between KADK and CHIN in 50 Kb non-overlapping windows. The dark slate gray and deep sky blue colors represent the alternative chromosomes. The dotted horizontal black line marks the 99th percentile outlier of estimated  $F_{ST}$  for autosome and the Z chromosome, respectively. **(B)** Highest  $F_{ST}$  region with a prominent pattern occurs on chromosome 20. The panels from top to bottom show estimates of  $F_{ST}$  in 10 Kb non-overlapping windows within population pairwise nucleotide diversity ( $\pi$ ) and integrated haplotype score (iHS) with the major allele as ancestral and the minor allele as derived. The genes are represented as black boxes with names. The horizontal red dotted line marks the 99th percentile 10 Kb  $F_{ST}$  outlier. The transparent gray color highlights the Dup1 region in all panes in the B panel, while R1 and R2 represent the selective sweep regions. **(C)** Non-synonymous changes (shown above the exon) within genes in the regions R1 and R2 are denoted by vertical orange lines. Blue boxes are the exons connected by black lines representing introns.

## Genome-wide signatures of selection in Kadaknath chicken

We screened the genome for selection signatures to identify all Kadaknath-specific regions by comparing Kadaknath with Chinese BBC. We also compared Kadaknath and a Chinese BBC population (XBBC,  $n = 4$ ) with the Korean Yeonsan Ogye to identify population-specific signatures. Despite the sizeable geographic separation, the genome-wide mean  $F_{ST}$  between KADK and CHIN is 0.16, making it amenable to identifying selection signatures (Figure 8A; Supplementary Table S6). A pairwise comparison between KADK and CHIN revealed 137 genic regions in the top 1%  $F_{ST}$  windows (Supplementary Table S7; Supplementary Figure S62–94). We further shortlisted candidates using the additional criteria that genetic diversity ( $\pi$  and  $\theta$ ) and integrated haplotype score (iHS) should be strikingly different between KADK and CHIN. Two prominent signatures from these shortlisted regions indicative of selective sweeps in the KADK breed are evident on the 20th chromosome in the vicinity of the Dup1 region. The first region (R1) of ~300 Kb is ~160 Kb

before Dup1, and the second region (R2) of ~70 Kb is ~0.7 Mb before Dup1. In both regions, the genetic diversity ( $\pi$  and  $\theta$ ) and Tajima's D ( $\tau$ ) are strongly reduced in KADK compared to CHIN (Figure 8B; Supplementary Figure S81). Compared to CHIN, the elevated iHS in KADK and the positive value of extended haplotype homozygosity (XP-EHH) support a selective sweep in KADK. This region also occurs in the top 10% genome-wide genetic divergence ( $D_{xy}$ ) windows (Supplementary Figure S95).

In the first region (R1), deoxynucleotidyltransferase terminal interacting protein 1 (*DNTTIP1*), acyl-CoA thioesterase 8 (*ACOT8*), and solute carrier family 12 member 5 (*SLC12A5*) genes contain non-synonymous changes. The TELO2-interacting protein 1 homolog (*TTI1*), which occurs near the first sweep region (R1), also has a non-synonymous change (Figure 8C; Supplementary Table S8, 9). The second sweep region (R2) contains three genes with non-synonymous changes: histocompatibility minor 13 (*HM13*), bactericidal/permeability-increasing protein-like (*BPIIL*), and targeting protein for Xklp2 (*TPX2*). The most striking differences were in the *BPIIL* gene, which has three fixed sites with non-synonymous changes. Of these changes, two alleles are



unique to the KADK population (Supplementary Figure S96). The first non-synonymous change (CGG->CAG, R->Q) occurs at position 69 in exon 2. The second non-synonymous change (GAG->AAG, E->K) is within the first BPI superfamily protein domain at position 159 in exon 4. The third and last non-synonymous change (CGC->CAC, R->H) is in the second domain at position 405 in exon 11. The impact of non-synonymous changes (R69Q, E159K, and R405H) in *BPIL* specific to Kadaknath is annotated as a moderate effect in the snpEff database. The PolyPhen-2 (Polymorphism Phenotyping v2) tool found that the non-synonymous changes R69Q and E159K are possibly damaging, while R405H is probably damaging (Supplementary Figure S97–99). We also found that *HM13*, *SLC12A5*, and *DNTTIP1* have novel Kadaknath-specific changes. The *HM13* gene is highly expressed in the muscle tissue, and we could assess the Kadaknath-specific change in all the native Indian breeds. Only the Kadaknath transcriptome contained the change identified in the genome sequencing data.

The Korean Yeonsan Ogye breed is closer to the Chinese XBBC (mean  $F_{ST}$  = 0.26) than to Kadaknath (mean  $F_{ST}$  = 0.3). However, the genome-wide mean  $F_{ST}$  is still low enough to distinguish selection signatures. A comparison of YOSK with KADK confirmed that R1 and R2 regions are specific to Kadaknath (Supplementary Figure S100; Supplementary Table S6). Surprisingly, our comparisons with YOSK identified a third region (R3) of ~570 kb, which is ~2.16 Mb before Dup1 with a sweep specific to YOSK (Supplementary Figure S101; Supplementary Table S9). Local PCA of the *Fm* locus, R1, and R2 regions demonstrates that Flank1, R1, and R2 regions have changed only in Kadaknath and not in other BBC (Supplementary Figure S102A–E; Supplementary Figure S103, 104). A careful examination of entire chromosome 20 failed to identify any other sweep regions (Supplementary Figure S105). Similar to the sweep in KADK, we found two genomic regions (on Chr 4 and Chr 9) with signatures of a sweep in the CHIN population (Supplementary Table S9). The sweep region on Chr 4 contains the *PCDH7* gene, and the region on Chr 9 contains the *COL4A3* and *MFF* genes (Supplementary Figure S65, 70). However, none of these genes have any non-synonymous fixed differences.

## Discussion

### All BBC breeds share a common origin of the *Fm* locus

As part of this study, we have generated the first whole-genome dataset of the Kadaknath breed spanning all three morphs. Hence, we can evaluate the relationship between all BBC breeds by comparing our data with public datasets. Our comparison of the *Fm* locus junction region in genome sequencing data from BBC breeds conclusively establishes a common origin for the complex chromosomal rearrangement (Dharmayanthi et al., 2017). The following four lines of evidence support this conclusion: (1) the short-read coverage along the *Fm* locus, (2) the high-sequence identity of the rearranged junctions across the BBC breeds, (3) the local and phylogenetic relationship between chicken breeds at the *Fm* locus, and (4) the patterns of pairwise genetic differentiation

between BBC and non-BBC breeds. Independent structural variants can produce the same phenotypes, as seen in the case of blue eggshell (Wang et al., 2013). However, in the case of BBC breeds, the hyperpigmentation phenotype results from a single shared complex chromosomal rearrangement.

### Spread of BBC across Asia

The earliest records of BBC have been found in the writings of Marco Polo, the Venetian explorer–writer who traveled through Asia (Dorshorst et al., 2011). The Compendium of Materia Medica or Bencao Gangmu, compiled and edited by Li Shizhen and published in the late 16th century, attributes various medicinal properties to BBCs (Wang et al., 2021). In Korea, the BBC is thought to have had a royal connection, and “Dongui Bogam,” a traditional Korean medical encyclopedia compiled and edited by Heo Jun in 1,613, records the medicinal use of BBC (Sohn et al., 2018). While an earlier study notes that “In 1635 AD, the finding of chickens with black meat (typical of fibromelanosis, FM mutation) in Mozambique suggested direct introductions from India” (Tixier-Boichard et al., 2011), we could not find literature on how Kadaknath ended up in Jhabua, India. However, Indian chickens are thought to have entered Africa through Egypt (contributing to breeds such as Fayoumi) before dispersing into Europe (Eltanany and Hemed, 2016).

Jhabua is close to the ancient port cities of Bharuch (also known as Bharukaccha, Barygaza, and Broach) (260 Km) and Lothal (300 Km) on the west coast of India. Given the proximity to port cities and the prevalence of BBC only in Western India (and the spread of BBC to Africa), we suspected that Kadaknath may have traveled through a marine route. The ancient sea trade between India, Korea, and other parts of Asia are well documented (Acri, 2018). Moreover, the movement of domesticated breeds during the period of colonialism was also facilitated by the common rule of large parts of Asia by various European powers (Sudrajad et al., 2020). We find evidence of crossbreeding between Kadaknath and Ankleshwar chicken breeds. The Ankleshwar breed of chicken is named after the city of Ankleshwar, located near modern-day Bharuch. Similar to Kadaknath, the Ankleshwar breed is reared by tribal communities of southern Gujarat. Evidence of gene flow between these native Indian breeds suggests that the rearing of Kadaknath and Ankleshwar breeds overlapped. None of the other native Indian breeds share ancestry with Kadaknath and support the movement of Kadaknath through the port city of Bharuch via a marine route.

Historical records are patchy, prone to error/obfuscation, and fail to provide conclusive information about the spread of BBC across Asia (Peters et al., 2022). Hence, we evaluated the relationship between BBC breeds and patterns of isolation by distance to infer the ancient dispersal route. Although each BBC breed has considerable genetic distinctiveness, population genetic analyses support the common heritage of all BBC breeds and reveal a trend of isolation by distance. While IBD patterns are well established in Chinese native chicken breeds (Nie et al., 2019), our dataset spans BBC breeds from India, China, Indonesia, and Korea. The sample size of some BBC breeds in our analysis is limited, and the sampling does not cover the entire geographic distribution of some breeds. However, our analysis achieves genome-wide pan-BBC breed sampling by including all the major breeds from Asia.

The BBC chicken from Tibet is genetically closest to the Kadaknath breed and suggests that the old Tibet–Nepal salt trade route or the maritime silk route may have facilitated the spread of BBC. Interestingly, the Tibetan BBC is genetically more similar to Kadaknath than some Chinese BBC breeds. Unfortunately, we lack BBC samples from Nepal. However, we believe BBC poultry in Nepal is Kadaknath chicken, recently imported from India. Hence, the ancestral stock of the BBC spread across Asia may not be currently available. Changes in trade routes and the introduction of commercial poultry breeds limit our ability to trace the historical prevalence of BBC breeds.

## Did black-bone chicken originate in India?

The lack of data from India and the considerable interest in BBC breeds in Europe, China, and Korea have meant that studies have focused mainly on non-Indian BBCs. Hence, Southern China and Tibet are considered the source of all BBC breeds (Zhu et al., 2014; Zhang et al., 2018; Xue et al., 2023). However, despite a traditionally restricted geographic distribution, the Kadaknath breed has nucleotide diversity comparable to Chinese BBC breeds and much higher than the Korean BBC. The number of private alleles identified in Kadaknath is also greater than that in Chinese and Korean BBC breeds. Moreover, the mean genome-wide  $F_{ST}$  of 0.11 between Jhabua and Bhopal is comparable to the differentiation between some Chinese BBC breeds. The presence of three distinct morphs within Kadaknath chicken suggests phenotypic diversity derived either from commercial breeds, native Indian breeds such as Ankleshwar, or existing variations within Kadaknath. The extent of phenotypic diversity within BBC breeds from China and Korea is not documented to allow a fair comparison. The high genetic diversity in Kadaknath supports the potential origin of all BBC in Jhabua, India. The export of black chicken from India to Africa in ~1600 AD also supports that BBC was present in ancient India (Tixier-Boichard et al., 2011).

While we cannot conclude whether BBC had an Indian, Chinese, or more Southeast Asian origin, our data suggest that any of these sources are plausible. More widespread geographic sampling and analysis of allele-sharing patterns may provide a definitive answer regarding the origin of BBC. For instance, the human-aided dispersal of the ginkgo tree out of China could be traced back to samples from eastern China using extensive sampling (Zhao et al., 2019). Future studies using such large-scale sampling could provide a more definitive answer regarding the origin and dispersal of BBC. Unlike the initial domestication of the chicken, which may have occurred independently (Tixier-Boichard et al., 2011; Mariadassou et al., 2021) in several locations, the BBC has a single origin linked to the rearrangement at the *Fm* locus. Irrespective of the source of all BBCs, we identify several Kadaknath-specific genetic changes. Hence, the genetic distinctiveness of the Kadaknath breed has long diverged from other BBC breeds and is a result of its unique heritage sustained by the Bhil and Bhilala tribal communities of Madhya Pradesh. The beliefs and practices of tribal communities may have contributed to the domestication and conservation of Kadaknath (Dar et al., 2019).

## Conclusive resolution of the complex chromosomal rearrangement

Earlier studies of the *Fm* locus have proposed three possible scenarios for the complex chromosomal rearrangement. Although the *\*Fm\_2* scenario was favored based on crosses between BBC and non-BBC breeds, the correct scenario was not established by previous studies as the genome assembly of this region is challenging. We use a haplotype phasing approach relying on published long-read datasets to conclusively resolve the correct arrangement to be the *\*Fm\_2* scenario at the *Fm* locus. In the *\*Fm\_2* scenario, the distal region (Dup1+Int + Dup2+Flank2) resembles the *\*N* arrangement found in non-BBC breeds. However, the proximal region (Flank1+Dup1+(inverted Dup2)) has a very different arrangement than the non-BBC *\*N* arrangement. In contrast to the *\*Fm\_2* scenario, the proximal region would be similar to the *\*N* arrangement in *\*Fm\_3* scenario. In the *\*Fm\_1* scenario, only the first occurrence of Dup1 and the last occurrence of Dup2 are similar to the *\*N* arrangement. Hence, in the *\*Fm\_2* and *\*Fm\_3* scenarios, recombination with *\*N* arrangement may be easier (i.e., recombination is suppressed only in the inverted ~127 Kb (Dup1) or 170 Kb (Dup2) region) compared to the *\*Fm\_1* scenario (i.e., recombination is suppressed in ~709 Kb (inverted (Dup2+Int + Dup1)) region). The ease of recombination with *\*N* arrangement could explain the prevalence of the *\*Fm\_2* scenario, which requires two rearrangement events. Ongoing improvements to genome assembly methods such as haplotype-aware *de novo* assembly (Chin et al., 2016; Korch et al., 2017; Garg et al., 2018; Sohn and Nam, 2018) and/or Strand-seq (Falconer et al., 2012; Sanders et al., 2017; Ghareghani et al., 2018) should allow easier resolution of such complex chromosomal rearrangements.

## Selective sweep near the *Fm* locus may be a consequence of linkage

Our analyses identified two genomic regions (R1 and R2) in the vicinity of the *Fm* locus with prominent signatures of a selective sweep in the KADK population. The most promising candidate gene in this region that may have been the focus of selection is the *BPIL* gene which has accumulated three non-synonymous changes predicted to be functionally important. The *BPIL* gene is part of the innate immune defense system, which binds and neutralizes lipopolysaccharides (LPS) from the outer membrane of Gram-negative bacteria (Holweg et al., 2011; Sun and Sun, 2016; Monson et al., 2019). *BPIL* is associated with upregulated expression after exposure to heat stress and LPS treatment in a comparative transcriptomic study between Fayoumi and broiler chicken breeds (Monson et al., 2019). The Fayoumi chicken breed is relatively resistant to Newcastle disease virus (NDV) compared to Leghorns (Deist et al., 2017) and potentially shares ancestry with Indian chicken breeds (Eltanany and Hemed, 2016). Earlier studies have shown that the Kadaknath has a high level of disease tolerance against Newcastle disease, also known as Ranikhet disease, compared to other chicken breeds (Kokate et al., 2017; Gumasta et al., 2021; Malarmathi et al., 2023). Along with resistance against Newcastle disease, it has been reported that the Kadaknath breed is less sensitive to coccidial infection (Thakur et al., 2015). The non-synonymous variants in the *BPIL* gene might enhance the disease resistance in Kadaknath against various bacterial and viral infections and be associated with a better immune response. Further work to

functionally evaluate the effect of these changes will provide a clear answer.

Five *BPI*-like genes occur near *BPI* (Chiang et al., 2011), which help arrest bacterial growth, are prominent in neutrophil phagocytosis, and work as a bactericidal protein (Elsbach and Weiss, 1998; Weiss, 2003). In chickens, defense against Gram-negative bacteria may be especially important as they lack the complement *C9* gene required to effectively eliminate pathogens through the membrane attack complex (MAC) formation (Sharma et al., 2022a). Notably, another BBC breed from Korea, YOSK, has signatures of selection at the *TLR4* gene involved in detecting Gram-negative bacteria (Cho et al., 2022). Hence, various immune genes may have been selected in different chicken breeds to protect against Gram-negative bacteria.

Several other genes with non-synonymous changes occur in R1 and R2 genomic regions. Although we cannot identify the functional consequences of these changes, the strong signature of selection and KADK-specific protein-coding alterations suggest phenotype-altering ability. Predictions from snpEff and PolyPhen-2 also suggest a major effect of these protein-coding changes located within the domain regions. The genomic region of high differentiation is close to the *Fm* locus, while weak genome-wide differentiation (except for chr 4 and chr9 regions) between Kadaknath and Chinese BBCs is strongly suggestive of hitchhiking. Moreover, the sweep signature is in the Chinese BBC in chr 4 and chr 9 cases. Hence, the proximity of R1 and R2 regions to the *Fm* locus indicates that selective sweep may be due to close physical linkage with the *Fm* locus. Such co-selection of traits due to linkage and pleiotropy has occurred during animal domestication, including in chicken (Rubin et al., 2010; Feder et al., 2014; Webster et al., 2015). Hitchhiking of alleles can increase the frequency of both beneficial and mildly deleterious alleles (Kirkpatrick and Barton, 2006; Bierne, 2010; Huff et al., 2012; Flaxman et al., 2013; Franssen et al., 2015; Martin et al., 2020; Hale et al., 2021; Weist et al., 2022). Whether the changes in KADK are beneficial or deleterious need to be investigated in future studies.

The genomic co-occurrence of economically important traits in domesticated plants and animals is also known (Zhang et al., 2010; Kijas et al., 2012; Beissinger et al., 2014; Fang et al., 2017; Kong et al., 2018; Huang et al., 2020a; Wang et al., 2020b; Jayakodi et al., 2020; Guan et al., 2021; Li et al., 2022a). The alleles causing phenotypic changes can occur in closely linked genes and will undergo selection for any of the phenotypes being favored by the breeding process. Notably, R1 and R2 selective sweeps are found only in KADK, R3 in YOSK, and none in Chinese BBC breeds. Hence, the genetic variants and the selective sweep may represent recent events after the BBC breeds separated from each other. Comparing different BBC breeds provides snapshots of the selection process and the build-up of genetically linked co-selected allelic changes and may help understand the hitchhiking process during domestication.

Genome-wide data generated as part of this study and comparative analysis with other BBC breeds establish the genetic uniqueness of Kadaknath that extends beyond the *Fm* locus. We also identify specific genes with selection signatures that are likely responsible for the Kadaknath-specific phenotypes. The co-selection of genes that are in linkage, as shown in Kadaknath, is widespread in domestic species (Rubin et al., 2010; Zhang et al., 2010; Webster et al., 2015; Fang et al., 2017; Kong et al., 2018; Li et al., 2022a). Such clustering of linked alleles under selection is favored in low recombination regions and near chromosomal rearrangements. Our work exemplifies the interaction

of artificial selection and chromosomal rearrangement-linked traits in domesticated species.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ebi.ac.uk/ena, PRJEB51457>. Scripts and data are available at: [https://github.com/ceglabsagarshinde/Kadaknath\\_Project](https://github.com/ceglabsagarshinde/Kadaknath_Project) and <https://doi.org/10.17632/8f9dn6h76h.1>.

## Ethics statement

The study was approved by the Institutional Ethics Committee (IEC) of the Indian Institute of Science Education and Research, Bhopal, vide reference number IISERB/IEC/Certificate/2018-11/03 dated 8th June 2018.

## Author contributions

Conceptualization: SSS, AS, and NV. Methodology: SSS and AS. Investigation: SSS, AS, and NV. Visualization: SSS and AS. Supervision: NV. Writing—original draft: SSS and AS. Writing—review and editing: SSS, AS, and NV. All authors contributed to the article and approved the submitted version.

## Funding

This study was supported by the Council of Scientific and Industrial Research Fellowship (SSS), University Grants Commission Ph.D. scholarship (AS), Department of Biotechnology, Ministry of Science and Technology, India (Grant No. BT/11/TYBA/2018/03) (NV), and Science and Engineering Research Board (Grant No. ECR/2017/001430) (NV). No funds are available for open-access publication fees.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1180658/full#supplementary-material>



## References

- Acri, A. (2018). Maritime buddhism. *Oxf. Res. Encycl. Relig.* doi:10.1093/ACREFORE/9780199340378.013.638
- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. doi:10.1038/NMETH0410-248
- Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., and Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21, 1–16. doi:10.1186/S13059-020-1935-5
- Bateson, B. W., -a, F., Punnett, R. C., and Contents, M. (1911). The inheritance of the peculiar pigmentation of the silky fowl. *J. Genet.* 13 (1), 185–203. doi:10.1007/BF02981551
- Beissinger, T. M., Hirsch, C. N., Vaillancourt, B., Deshpande, S., Barry, K., Robin Buell, C., et al. (2014). A genome-wide scan for evidence of selection in a maize population under long-term artificial selection for ear number. *Genetics* 196, 829–840. doi:10.1534/GENETICS.113.160655
- Bierne, N. (2010). The distinctive footprints of local hitchhiking in a varied environment and global hitchhiking in a subdivided population. *Evolution* 64, 3254–3272. doi:10.1111/j.1558-5646.2010.01050.x
- Buranawit, K., Chailungka, C., Wongsunsi, C., and Laenoi, W. (2016). Phenotypic characterization of Thai native black-bone chickens indigenous to northern Thailand. *Thai J. Vet. Med.* 46, 547–554.
- Chiang, S. C., Veldhuizen, E. J. A., Barnes, F. A., Craven, C. J., Haagsman, H. P., and Bingle, C. D. (2011). Identification and characterisation of the BPI/LBP/PLUNC-like gene repertoire in chickens reveals the absence of a LBP gene. *Dev. Comp. Immunol.* 35, 285–295. doi:10.1016/j.dci.2010.09.013
- Chin, C. S., Peluso, P., Sedlazeck, F. J., Nattestad, M., Concepcion, G. T., Clum, A., et al. (2016). Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* 13, 1050–1054. doi:10.1038/NMETH.4035
- Cho, Y., Kim, J. Y., and Kim, N. (2022). Comparative genomics and selection analysis of Yeosan Ogye black chicken with whole-genome sequencing. *Genomics* 114. doi:10.1016/j.ygeno.2022.110298
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). 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. (Austin)* 6, 80–92. doi:10.4161/FLY.19695/SUPPL\_FILE/KFLY\_A\_10919695\_SM0001.ZIP
- Cogburn, L. A., Porter, T. E., Duclos, M. J., Simon, J., Burgess, S. C., Zhu, J. J., et al. (2007). Functional genomics of the chicken-a model organism. *Poult. Sci.* 86, 2059–2094. doi:10.1093/PS/86.10.2059
- Cook, D. E., and Andersen, E. C. (2017). VCF-Kit: Assorted utilities for the variant call format. *Bioinformatics* 33, 1581–1582. doi:10.1093/BIOINFORMATICS/BTX011
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/BIOINFORMATICS/BTR330
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, 1–4. doi:10.1093/GIGASCIENCE/GIAB008
- Dar, J. A., Subashree, K., Raha, D., Kumar, A., Khare, P. K., and Khan, M. L. (2019). Tree diversity, biomass and carbon storage in sacred groves of Central India. *Environ. Sci. Pollut. Res.* 26, 37212–37227. doi:10.1007/S11356-019-06854-9
- Deist, M. S., Gallardo, R. A., Bunn, D. A., Kelly, T. R., Dekkers, J. C. M., Zhou, H., et al. (2017). Novel mechanisms revealed in the trachea transcriptome of resistant and susceptible chicken lines following infection with Newcastle disease virus. *Clin. Vaccine Immunol.* 24. doi:10.1128/CVI.00027-17
- Delaneau, O., Marchini, J., and Zagury, J. F. (2011). A linear complexity phasing method for thousands of genomes. *Nat. Methods* 9 (9), 179–181. doi:10.1038/nmeth.1785
- Dharmayanthi, A. B., Terai, Y., Sulandari, S., Zein, M. S. A., Akiyama, T., and Satta, Y. (2017). The origin and evolution of fibromelanosis in domesticated chickens: Genomic comparison of Indonesian Cemani and Chinese Silkie breeds. *PLoS One* 12. doi:10.1371/JOURNAL.PONE.0173147
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). Star: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635
- Dorshorst, B., Okimoto, R., and Ashwell, C. (2010). Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the silky chicken. *J. Hered.* 101, 339. doi:10.1093/hered/esp120
- Dorshorst, B., Molin, A. M., Rubin, C. J., Johansson, A. M., Strömstedt, L., Pham, M. H., et al. (2011). A complex genomic rearrangement involving the Endothelin 3 locus causes dermal hyperpigmentation in the chicken. *PLoS Genet.* 7. doi:10.1371/journal.pgen.1002412
- Dou, T., Yan, S., Liu, L., Wang, K., Jian, X., Xu, Z., et al. (2022). Integrative analysis of transcriptomics and metabolomics to reveal the melanogenesis pathway of muscle and related meat characters in Wuliangshan black-boned chickens. *BMC Genomics* 23, 1–15. doi:10.1186/s12864-022-08388-w
- Elsbach, P., and Weiss, J. (1998). Role of the bactericidal/permeability-increasing protein in host defence. *Curr. Opin. Immunol.* 10, 45–49. doi:10.1016/S0952-7915(98)80030-7
- Eltanany, M. A., and Hemed, S. A. (2016). Deeper insight into maternal genetic assessments and demographic history for Egyptian indigenous chicken populations using mtDNA analysis. *J. Adv. Res.* 7, 615–623. doi:10.1016/J.JARE.2016.06.005
- Eriksson, J., Larson, G., Gunnarsson, U., Bed'hom, B., Tixier-Boichard, M., Strömstedt, L., et al. (2008). Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.* 4. doi:10.1371/JOURNAL.PGEN.1000010
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14, 2611–2620. doi:10.1111/j.1365-294X.2005.02553.X
- Falconer, E., Hills, M., Naumann, U., Poon, S. S. S., Chavez, E. A., Sanders, A. D., et al. (2012). DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* 9, 1107–1112. doi:10.1038/nmeth.2206
- Fan, W. L., Ng, C. S., Chen, C. F., Lu, M. Y. J., Chen, Y. H., Liu, C. J., et al. (2013). Genome-wide patterns of genetic variation in two domestic chickens. *Genome Biol. Evol.* 5, 1376–1392. doi:10.1093/GBE/EVT097
- Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., et al. (2017). Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* 18. doi:10.1186/S13059-017-1289-9
- Feder, J. L., Nosil, P., and Flaxman, S. M. (2014). Assessing when chromosomal rearrangements affect the dynamics of speciation: Implications from computer simulations. *Front. Genet.* 5, 295. doi:10.3389/fgene.2014.00295
- Flaxman, S. M., Feder, J. L., and Nosil, P. (2013). Genetic hitchhiking and the dynamic buildup of genomic divergence during speciation with gene flow. *Evolution* 67, 2577–2591. doi:10.1111/evo.12055
- Franssen, S. U., Nolte, V., Tobler, R., and Schlotterer, C. (2015). Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila melanogaster* populations. *Mol. Biol. Evol.* 32, 495–509. doi:10.1093/MOLBEV/MSU320
- Garg, S., Rautiainen, M., Novak, A. M., Garrison, E., Durbin, R., and Marschall, T. (2018). A graph-based approach to diploid genome assembly. *Bioinformatics* 34, i105–i114. doi:10.1093/BIOINFORMATICS/BTY279
- Garrison, E., and Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing*.
- Gautier, M., and Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* 28, 1176–1177. doi:10.1093/BIOINFORMATICS/BTS115
- Ghareghani, M., Porubská, D., Sanders, A. D., Meiers, S., Eichler, E. E., Korbel, J. O., et al. (2018). Strand-seq enables reliable separation of long reads by chromosome via expectation maximization. *Bioinformatics* 34, i115–i123. doi:10.1093/BIOINFORMATICS/BTY290
- Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J. S., et al. (2012). Recommendations for utilizing and reporting population genetic analyses: The reproducibility of genetic clustering using the program STRUCTURE. *Mol. Ecol.* 21, 4925–4930. doi:10.1111/j.1365-294X.2012.05754.X
- Guan, J., Xu, Y., Yu, Y., Fu, J., Ren, F., Guo, J., et al. (2021). Genome structure variation analyses of peach reveal population dynamics and a 1.67 Mb causal inversion for fruit shape. *Genome Biol.* 22. doi:10.1186/S13059-020-02239-1
- Gumasta, P., Argade, S., Satpaty, S., Sahu, H., Porte, C., Ghosh, R. C., et al. (2021). Occurrence of various disease conditions and their pathology in Kadaknath chickens. *Pharma Innov. J.* 10, 218–222. doi:10.22271/TPI.2021.V10.I8SD.7213
- Hale, M. C., Campbell, M. A., and McKinney, G. J. (2021). A candidate chromosome inversion in Arctic charr (*Salvelinus alpinus*) identified by population genetic analysis techniques. *G3 (Bethesda)* 11. doi:10.1093/G3/JOURNAL/JKAB267
- Haunshi, S., and Prince, L. L. L. (2021). Kadaknath: A popular native chicken breed of India with unique black colour characteristics. *World's Poult. Sci. J.* 77, 427–440. doi:10.1080/00439339.2021.1897918
- Haunshi, S., Niranjana, M., Shanmugam, M., Padhi, M. K., Reddy, M. R., Sunitha, R., et al. (2011). Characterization of two Indian native chicken breeds for production, egg and semen quality, and welfare traits. *Poult. Sci.* 90, 314–320. doi:10.3382/PS.2010-01013
- Haunshi, S., Devatkal, S., Prince, L. L. L., Ullengala, R., Ramasamy, K., and Chatterjee, R. (2022). Carcass characteristics, meat quality and nutritional composition of Kadaknath, a native chicken breed of India. *Foods (Basel, Switz.)* 11, 3603. doi:10.3390/FOODS11223603
- Holweg, A., Schnare, M., and Gessner, A. (2011). The bactericidal/permeability-increasing protein (BPI) in the innate defence of the lower airways. *Biochem. Soc. Trans.* 39, 1045–1050. doi:10.1042/BST0391045
- Hon, T., Mars, K., Young, G., Tsai, Y. C., Karalija, J. W., Landolin, J. M., et al. (2020). Highly accurate long-read HiFi sequencing data for five complex genomes. *Sci. Data* 7, 1–11. doi:10.1038/s41597-020-00743-4



- Huang, K., Andrew, R. L., Owens, G. L., Ostevik, K. L., and Rieseberg, L. H. (2020a). Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype. *Mol. Ecol.* 29, 2535–2549. doi:10.1111/MEC.15428
- Huang, X., Otecko, N. O., Peng, M., Weng, Z., Li, W., Chen, J., et al. (2020b). Genome-wide genetic structure and selection signatures for color in 10 traditional Chinese yellow-feathered chicken breeds. *BMC Genomics* 21. doi:10.1186/S12864-020-6736-4
- Huff, C. D., Witherspoon, D. J., Zhang, Y., Gatenbee, C., Denson, L. A., Kugathasan, S., et al. (2012). Crohn's disease and genetic hitchhiking at IBD5. *Mol. Biol. Evol.* 29, 101. doi:10.1093/MOLBEV/MSR151
- Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics* 14, 68–73. doi:10.1093/BIOINFORMATICS/14.1.68
- Jadhao, R. G., Kumbhar, D. D., and Wakode, S. (2022). Understanding the histology and hematobiochemical aspects of Kadaknath an Indian black fowl with an appealing taste native to Dhar district Madhya Pradesh. *Br. Poult. Sci.* 64, 214. doi:10.1080/00071668.2022.2154637
- Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., et al. (2017). The K = 2 conundrum. *Mol. Ecol.* 26, 3594–3602. doi:10.1111/MEC.14187
- Jaturasitha, S., Srikanthai, T., Kreuzer, M., and Wicke, M. (2008). Differences in carcass and meat characteristics between chicken indigenous to northern Thailand (Black-Boned and Thai native) and imported extensive breeds (bresse and Rhode Island red). *Poult. Sci.* 87, 160–169. doi:10.3382/PS.2006-00398
- Jayakodi, M., Padmarasu, S., Haberer, G., Bonthala, V. S., Gundlach, H., Monat, C., et al. (2020). The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* 588, 284–289. doi:10.1038/S41586-020-2947-8
- Jena, P., Panigrahi, B., Panda, N., Mohapatra, L., Mallik, B., and Bagh, J. (2018). Reproductive performance and egg quality traits of Kadaknath in intensive managemental condition under hot and humid climate. *Int. J. Livest. Res.* 1. doi:10.5455/IJLR.20180131054034
- Johansson, A. M., and Nelson, R. M. (2015). Characterization of genetic diversity and gene mapping in two Swedish local chicken breeds. *Front. Genet.* 6, 44. doi:10.3389/fgene.2015.00044
- Khanna, A., Larson, D. E., Srivatsan, S. N., Mosior, M., Abbott, T. E., Kiwala, S., et al. (2022). Bam-readcount - rapid generation of basepair-resolution sequence metrics. *ArXiv* 7, 3722. doi:10.21105/joss.03722
- Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Neto, L. R., Cristobal, M. S., et al. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* 10. doi:10.1371/JOURNAL.PBIO.1001258
- Kirkpatrick, M., and Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics* 173, 419–434. doi:10.1534/genetics.105.047985
- Kokate, L. S., Kumar, S., Rahim, A., and Das, A. K. (2017). Estimating serological immune response against Newcastle disease vaccine in Aseel, Kadaknath and White Leghorn chicken by haemagglutination inhibition test. doi:10.37191/QUERY-UIJS
- Kong, L., Lu, S., Wang, Y., Fang, C., Wang, F., Nan, H., et al. (2018). Quantitative trait locus mapping of flowering time and maturity in soybean using next-generation sequencing-based analysis. *Front. Plant Sci.* 9. doi:10.3389/FPLS.2018.00995
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. doi:10.1111/1755-0998.12387
- Korlach, J., Gedman, G., Kingan, S. B., Chin, C. S., Howard, J. T., Audet, J. N., et al. (2017). De novo PacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *Gigascience* 6. doi:10.1093/GIGASCIENCE/GIX085
- Korneliusson, T. S., Albrechtsen, A., and Nielsen, R. (2014). Angsd: Analysis of next generation sequencing data. *BMC Bioinforma.* 15, 356. doi:10.1186/s12859-014-0356-4
- Kralik, G., Kralik, Z., Grčević, M., and Hanzek, D. (2016). "Quality of chicken meat," in *Animal Husbandry and Nutrition*, Editor B. Yücel and T. Taşkın. doi:10.5772/intechopen.72865
- Kriangwanich, W., Piboon, P., Sakorn, W., Buddhachat, K., Kochagul, V., Pringproa, K., et al. (2021). Consistency of dark skeletal muscles in Thai native black-bone chickens (*Gallus gallus domesticus*). *PeerJ* 9. doi:10.7717/PEERJ.10728
- Kronenberg, Z. N., Rhie, A., Koren, S., Concepcion, G. T., Peluso, P., Munson, K. M., et al. (2021). Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nat. Commun.* 12. doi:10.1038/S41467-020-20536-Y
- Krumsiek, J., Arnold, R., and Rattai, T. (2007). Gepard: A rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23, 1026–1028. doi:10.1093/BIOINFORMATICS/BTM039
- Lawal, R. A., and Hanotte, O. (2021). Domestic chicken diversity: Origin, distribution, and adaptation. *Anim. Genet.* 52, 385–394. doi:10.1111/AGE.13091
- Lee, T. H., Guo, H., Wang, X., Kim, C., and Paterson, A. H. (2014). SNPhylo: A pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* 15, 162. doi:10.1186/1471-2164-15-162
- Leigh, J. W., and Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi:10.1111/2041-210X.12410
- Li, H., and Barrett, J. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi:10.1093/BIOINFORMATICS/BTR509
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324
- Li, D., Sun, G., Zhang, M., Cao, Y., Zhang, C., Fu, Y., et al. (2020). Breeding history and candidate genes responsible for black skin of Xichuan black-bone chicken. *BMC Genomics* 21. doi:10.1186/S12864-020-06900-8
- Li, J., Zhang, Y., Ma, R., Huang, W., Hou, J., Fang, C., et al. (2022a). Identification of ST1 reveals a selection involving hitchhiking of seed morphology and oil content during soybean domestication. *Plant Biotechnol. J.* 20, 1110–1121. doi:10.1111/PBI.13791
- Li, M., Sun, C., Xu, N., Bian, P., Tian, X., Wang, X., et al. (2022b). De novo assembly of 20 chicken genomes reveals the undetectable phenomenon for thousands of core genes on microchromosomes and subtelomeric regions. *Mol. Biol. Evol.* 39. doi:10.1093/MOLBEV/MSAC066
- Lindenbaum, P. (2015). *JVarkit: Java-based utilities for bioinformatics*. doi:10.6084/M9.FIGSHARE.1425030.V1
- Malarmathi, M., Nagarajan, M., Mani, S., Karuppusamy, S., Vasudevan, G., Vadivel, R. B., et al. (2023). In-vitro characterization of <em>chIFITMs</em> of Aseel and Kadaknath chicken breeds against Newcastle disease virus infection. *Biol. life Sci.* 2023041180. doi:10.20944/PREPRINTS202304.1180.V1
- Mariadassou, M., Suez, M., Sathyakumar, S., Vignal, A., Arca, M., Nicolas, P., et al. (2021). Unraveling the history of the genus Gallus through whole genome sequencing. *Mol. Phylogenet. Evol.* 158. doi:10.1016/J.YMPEV.2020.107044
- Martin, S. H., Singh, K. S., Gordon, I. J., Omufwoko, K. S., Collins, S., Warren, I. A., et al. (2020). Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLoS Biol.* 18. doi:10.1371/JOURNAL.PBIO.3000610
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi:10.1101/gr.107524.110
- Meisner, J., and Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* 210, 719–731. doi:10.1534/GENETICS.118.301336
- Mir, N. A., Rafiq, A., Kumar, F., Singh, V., and Shukla, V. (2017). Determinants of broiler chicken meat quality and factors affecting them: A review. *J. Food Sci. Technol.* 54, 2997–3009. doi:10.1007/S13197-017-2789-Z
- Monson, M. S., Van Goor, A. G., Persia, M. E., Rothschild, M. F., Schmidt, C. J., and Lamont, S. J. (2019). Genetic lines respond uniquely within the chicken thymic transcriptome to acute heat stress and low dose lipopolysaccharide. *Sci. Rep.* 9. doi:10.1038/S41598-019-50051-0
- Muroya, S., Tanabe, R. I., Nakajima, I., and Chikuni, K. (2000). Molecular characteristics and site specific distribution of the pigment of the silky fowl. *J. Vet. Med. Sci.* 62, 391–395. doi:10.1292/JVMS.62.391
- Musmann, S. M., Douglas, M. R., Chafin, T. K., and Douglas, M. E. (2020). Admixpipe: Population analyses in ADMIXTURE for non-model organisms. *BMC Bioinforma.* 21. doi:10.1186/S12859-020-03701-4
- Nie, C., Almeida, P., Jia, Y., Bao, H., Ning, Z., Qu, L., et al. (2019). Genome-wide single-nucleotide polymorphism data unveil admixture of Chinese indigenous chicken breeds with commercial breeds. *Genome Biol. Evol.* 11, 1847–1856. doi:10.1093/GBE/EVZ128
- Patterson, M. D., Marschall, T., Pisanti, N., Van Iersel, L., Stougie, L., Klau, G. W., et al. (2015). WhatsHap: Weighted haplotype assembly for future-generation sequencing reads. *J. Comput. Biol.* 22, 498–509. doi:10.1089/CMB.2014.0157
- Peters, J., Lebrasseur, O., Irving-Pease, E. K., Paxinos, P. D., Best, J., Smallman, R., et al. (2022). The biocultural origins and dispersal of domestic chickens. *Proc. Natl. Acad. Sci. U. S. A.* 119. doi:10.1073/PNAS.2121978119/-DCSUPPLEMENTAL
- Qanbari, S., Rubin, C. J., Maqbool, K., Weigend, S., Weigend, A., Geibel, J., et al. (2019). Genetics of adaptation in modern chicken. *PLoS Genet.* 15. doi:10.1371/JOURNAL.PGEN.1007989
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. Available at: <http://bioinformatics.oxfordjournals.org/content/26/6/841.short> (Accessed July 9, 2014).
- Rout, P. K., Pani, P. K., and Naithani, S. (1992). Genetic susceptibility of indigenous chicks to subgroup A Rous sarcoma virus inoculated via the chorioallantoic membrane. *Vet. Immunol. Immunopathol.* 33, 89–102. doi:10.1016/0165-2427(92)90037-Q
- Roux, P. F., Boutin, M., Désert, C., Djari, A., Esquerré, D., Klopp, C., et al. (2014). Re-sequencing data for refining candidate genes and polymorphisms in QTL regions affecting adiposity in chicken. *PLoS One* 9. doi:10.1371/JOURNAL.PONE.0111299

- Rubin, C. J., Zody, M. C., Eriksson, J., Meadows, J. R. S., Sherwood, E., Webster, M. T., et al. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464, 587–591. doi:10.1038/nature08832
- Sahu, N., Kumar, V., Dhruw, S., and Kumar Patre, S. (2019). Woman empowerment and uplifting of tribal farmers through Kadaknath poultry farming in Dantewada district of Chhattisgarh. *J. Entomol. Zool. Stud.* 7, 315–319.
- Sanders, A. D., Falconer, E., Hills, M., Spierings, D. C. J., and Lansdorp, P. M. (2017). Single-cell template strand sequencing by Strand-seq enables the characterization of individual homologs. *Nat. Protoc.* 12(12), 1151–1176. doi:10.1038/nprot.2017.029
- Sehrawat, R., Sharma, R., Ahlawat, S., Sharma, V., Thakur, M. S., Kaur, M., et al. (2021). First report on better functional property of black chicken meat from India. *Indian J. Anim. Res.* 55, 727–733. doi:10.18805/IJAR.B-4014
- Shanmathy, M., Tyagi, J. S., Gopi, M., Mohan, J., Beulah, P., and Ravi Kumar, D. (2018). Comparative assessment on performance of Aseel and Kadaknath in hot and humid conditions in tropics. *Int. J. Curr. Microbiol. Appl. Sci.* 7, 2156–2165. doi:10.20546/ijcmas.2018.705.251
- Sharma, A., Gupta, S., Patil, A. B., and Vijay, N. (2022a). Birth and death in terminal complement pathway. *Mol. Immunol.* 149, 174–187. doi:10.1016/j.molimm.2022.07.006
- Sharma, R., Sehrawat, R., Ahlawat, S., Sharma, V., Parmar, A., Thakur, M. S., et al. (2022b). An attempt to valorize the only black meat chicken breed of India by delineating superior functional attributes of its meat. *Sci. Rep.* 12, 1–12. doi:10.1038/s41598-022-07575-9
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29, 308–311. doi:10.1093/NAR/29.1.308
- Shinomiya, A., Kayashima, Y., Kinoshita, K., Mizutani, M., Namikawa, T., Matsuda, Y., et al. (2012). Gene duplication of endothelin 3 is closely correlated with the hyperpigmentation of the internal organs (fibromelanosis) in silky chickens. *Genetics* 190, 627. doi:10.1534/genetics.111.136705
- Skotte, L., Korneliusen, T. S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics* 195, 693–702. doi:10.1534/GENETICS.113.154138
- Slater, A. E. (1945). The poultry industry in India. *Worlds. Poult. Sci. J.* 1, 46–51. doi:10.1079/WPS19450020
- Sohn, J. I., and Nam, J. W. (2018). The present and future of de novo whole-genome assembly. *Brief. Bioinform.* 19, 23–40. doi:10.1093/BIB/BBW096
- Sohn, J., Nam, K., Hong, H., Kim, J. M., Lim, D., Lee, K. T., et al. (2018). Whole genome and transcriptome maps of the entirely black native Korean chicken breed Yeosan Ogye. *Gigascience* 7, 1–14. doi:10.1093/GIGASCIENCE/GIY086
- Sudrajad, P., Subiharta, S., Adinata, Y., Lathifah, A., Lee, J. H., Lenstra, J. A., et al. (2020). An insight into the evolutionary history of Indonesian cattle assessed by whole genome data analysis. *PLoS One* 15, e0241038. doi:10.1371/JOURNAL.PONE.0241038
- Sun, Y. Y., and Sun, L. (2016). A teleost bactericidal permeability-increasing protein kills gram-negative bacteria, modulates innate immune response, and enhances resistance against bacterial and viral infection. *PLoS One* 11. doi:10.1371/JOURNAL.PONE.0154045
- Thakur, M. S., Parmar, S. S., Kumar, A., Tomar, S. S., Sarkhel, B. C., Das, G., et al. (2015). Evaluation of Kadaknath chicken for coccidial resistance by oocyst count, lesion scoring and oocyst index in eimeria tenella infection. *J. Anim. Res.* 5, 579. doi:10.5958/2277-940X.2015.00097.2
- Tian, Y., Xie, M., Wang, W., Wu, H., Fu, Z., and Lin, L. (2007). Determination of carnosine in black-bone silky fowl (*Gallus gallus domesticus* brisson) and common chicken by HPLC. *Eur. Food Res. Technol.* 226, 311–314. doi:10.1007/S00217-006-0528-1/FIGURES/1
- Tian, Y., Zhu, S., Xie, M., Wang, W., Wu, H., and Gong, D. (2011). Composition of fatty acids in the muscle of black-bone silky chicken (*Gallus gallus domesticus* brisson) and its bioactivity in mice. *Food Chem.* 126, 479–483. doi:10.1016/j.foodchem.2010.11.024
- Tixier-Boichard, M., Bed'Hom, B., and Rognon, X. (2011). Chicken domestication: From archeology to genomics. *C. R. Biol.* 334, 197–204. doi:10.1016/j.crvi.2010.12.012
- Tripathi, S. B., Jatav, G., Malik, A. A., Joshi, S., Singh, V. K., Negi, M. S., et al. (2022). AFLP markers based genetic diversity and population structure analysis of Kadaknath: An indigenous black meat poultry breed of India. *Anim. Biotechnol.* 33, 1045–1055. doi:10.1080/10495398.2020.1865390
- Tu, Y., Sun, Y., Tian, Y., Xie, M., and Chen, J. (2009). Physicochemical characterisation and antioxidant activity of melanin from the muscles of Taihe Black-bone silky fowl (*Gallus gallus domesticus* brisson). *Food Chem.* 114, 1345–1350. doi:10.1016/j.foodchem.2008.11.015
- Ulfah, M., Kawahara-Miki, R., Farajallah, A., Muladno, M., Dorshorst, B., Martin, A., et al. (2016). Genetic features of red and green junglefowls and relationship with Indonesian native chickens Sumatera and Kedu Hitam. *BMC Genomics* 17. doi:10.1186/S12864-016-2652-Z
- Wang, Z., Qu, L., Yao, J., Yang, X., Li, G., Zhang, Y., et al. (2013). An EAV-HP insertion in 5' flanking region of SLCO1B3 causes blue eggshell in the chicken. *PLoS Genet.* 9, e1003183. doi:10.1371/JOURNAL.PGEN.1003183
- Wang, M. S., Li, Y., Peng, M. S., Zhong, L., Wang, Z. J., Li, Q. Y., et al. (2015). Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Mol. Biol. Evol.* 32, 1880–1889. doi:10.1093/MOLBEV/MSV071
- Wang, M. S., Thakur, M., Peng, M. S., Jiang, Y., Frantz, L. A. F., Li, M., et al. (2020a). 863 genomes reveal the origin and domestication of chicken. *Cell Res.* 30, 693. doi:10.1038/S41422-020-0349-Y
- Wang, X., Gao, L., Jiao, C., Stravoravdis, S., Hosmani, P. S., Saha, S., et al. (2020b). Genome of *Solanum pimpinellifolium* provides insights into structural variants during tomato breeding. *Nat. Commun.* 11. doi:10.1038/S41467-020-19682-0
- Wang, H., Cahaner, A., Lou, L., Zhang, L., Ge, Y., Li, Q., et al. (2021). Genetics and breeding of a black-bone and blue eggshell chicken line. 1. Body weight, skin color, and their combined selection. *Poult. Sci.* 100. doi:10.1016/j.psj.2021.101035
- Wang, X., Luan, Y., and Yue, F. (2022). EagleC: A deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. *Sci. Adv.* 8. doi:10.1126/SCIADV.ABN9215
- Webster, M. T., Kamgari, N., Perloski, M., Hoepfner, M. P., Axelsson, E., Hedhammar, Å., et al. (2015). Linked genetic variants on chromosome 10 control ear morphology and body mass among dog breeds. *BMC Genomics* 16, 1–17. doi:10.1186/S12864-015-1702-2/FIGURES/4
- Weiss, J. (2003). Bactericidal/permeability-increasing protein (BPI) and lipopolysaccharide-binding protein (LBP): Structure, function and regulation in host defence against gram-negative bacteria. *Biochem. Soc. Trans.* 31, 785–790. doi:10.1042/BST0310785
- Weissensteiner, M. H., Bunikis, I., Catalán, A., Francoijs, K. J., Knief, U., Heim, W., et al. (2020). Discovery and population genomics of structural variation in a songbird genus. *Nat. Commun.* 11, 1–11. doi:10.1038/s41467-020-17195-4
- Weist, P., Jentoft, S., Tørresen, O. K., Schade, F. M., Pampoulie, C., Krumme, U., et al. (2022). The role of genomic signatures of directional selection and demographic history in the population structure of a marine teleost with high gene flow. *Ecol. Evol.* 12. doi:10.1002/ECE3.9602
- Xue, Z., Wang, L., Tian, Y., Yang, Y., Li, P., Yang, G., et al. (2023). A genome-wide scan to identify signatures of selection in Lueyang black-bone chicken. *Poult. Sci.* 102, 102721. doi:10.1016/j.psj.2023.102721
- Zhang, D., Liu, C., Cheng, H., Kan, G., Cui, S., Meng, Q., et al. (2010). Quantitative trait loci associated with soybean tolerance to low phosphorus stress based on flower and pod abscission. *Plant Breed.* 129, 243–249. doi:10.1111/J.1439-0523.2009.01682.X
- Zhang, T., Du, W., Lu, H., and Wang, L. (2018). Genetic diversity of mitochondrial DNA of Chinese black-bone chicken. *Braz. J. Poult. Sci.* 20, 565–572. doi:10.1590/1806-9061-2018-0739
- Zhang, P., Cao, Y., Fu, Y., Zhu, H., Xu, S., Zhang, Y., et al. (2022). Revealing the regulatory mechanism of lncRNA-LMEP on melanin deposition based on high-throughput sequencing in xichuan chicken skin. *Genes (Basel)* 13, 2143. doi:10.3390/GENES1312143/S1
- Zhao, Y. P., Fan, G., Yin, P. P., Sun, S., Li, N., Hong, X., et al. (2019). Resequencing 545 ginkgo genomes across the world reveals the evolutionary history of the living fossil. *Nat. Commun.* 10. doi:10.1038/S41467-019-12133-5
- Zhu, W. Q., Li, H. F., Wang, J. Y., Shu, J. T., Zhu, C. H., Song, W. T., et al. (2014). Molecular genetic diversity and maternal origin of Chinese black-bone chicken breeds. *Genet. Mol. Res.* 13, 3275–3282. doi:10.4238/2014.APRIL.29.5