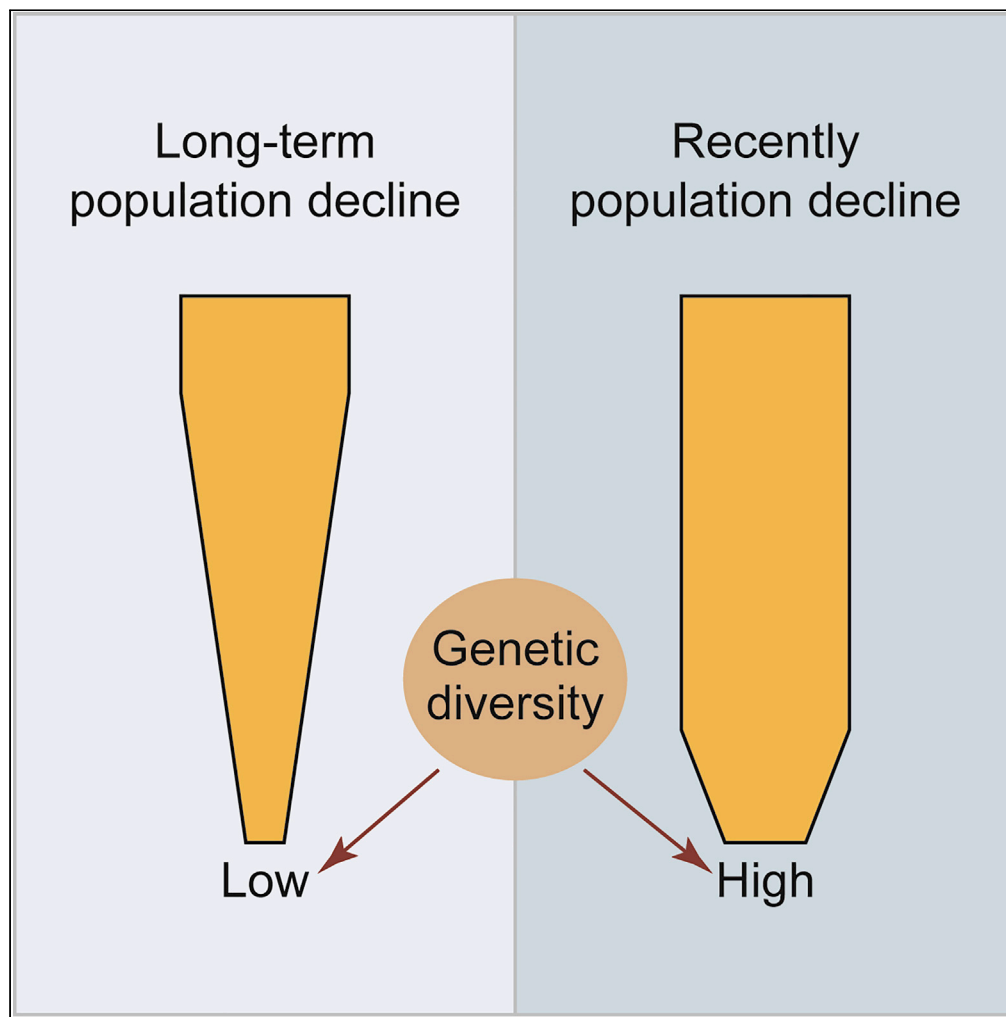


Article

Genomic status of yellow-breasted bunting following recent rapid population decline



Pengcheng Wang,
Rong Hou, Yang
Wu, Zhengwang
Zhang, Pinjia Que,
Peng Chen

quepinjia@gmail.com (P.Q.)
capricorncp@163.com (P.C.)

Highlights

A bird species harbors
high genetic diversity
following a rapid
population decline

Inbreeding levels are low,
too

However, the species'
genome has potentially
lethal mutations

Rapid population decline
of Yellow-breasted
Bunting likely began 147
generations ago

Wang et al., iScience 25,
104501
July 15, 2022 © 2022 The
Authors.
[https://doi.org/10.1016/
j.isci.2022.104501](https://doi.org/10.1016/j.isci.2022.104501)

Article

Genomic status of yellow-breasted bunting following recent rapid population decline

Pengcheng Wang,² Rong Hou,¹ Yang Wu,³ Zhengwang Zhang,³ Pinjia Que,^{1,4,*} and Peng Chen^{1,4,5,*}

SUMMARY

Global biodiversity is facing serious threats. However, knowledge of the genomic consequences of recent rapid population declines of wild organisms is limited. Do populations experiencing recent rapid population decline have the same genomic status as wild populations that experience long-term declines? Yellow-breasted Bunting (*Emberiza aureola*) is a critically endangered species that has been experiencing a recent rapid population decline. To answer the question, we assembled and annotated the whole genome of Yellow-breasted Bunting. Furthermore, we found high genetic diversity, low linkage disequilibrium, and low proportion of long runs of homozygosity in Yellow-breasted Bunting, suggesting that the populations following recent rapid declines have different genomic statuses from the population that experienced long-term population decline.

INTRODUCTION

Wild species are suffering population declines and some species have gone extinct (IUCN, 2021). 2.9 billion breeding adult birds have been lost since 1970 in North America, with particularly strong losses among grassland birds (53%) (Rosenberg et al., 2019). It is, therefore, urgent to understand the genomic consequences of the rapid population decline in wild organisms. Genetic status, including genetic diversity and inbreeding level, provides important information for planning conservation strategy (DeWoody et al., 2021). Endangered species with different genetic characteristics may have different survival fates. However, few at-risk species have had their genomic status evaluated. What is worse, it is still unclear what the genomic status of recent rapid population declines of wild species is.

With the development of genomic sequencing, the availability of population genomic data allows the exploration of genetic diversity, inbreeding level, mutation load, and historical effective population size fluctuations. Previous studies have found endangered species have low genetic diversity (e.g., *Crossoptilon mantchuricum*), high inbreeding levels (e.g., *Pavo muticus*), and damaging mutation loads (e.g., *Nipponia nippon*), but these species have experienced long-term population decline (Dong et al., 2021; Feng et al., 2019; Wang et al., 2021). Limited studies have explored the genomic status of species that have been subjected to recent rapid population declines.

Yellow-breasted Bunting (*Emberiza aureola*; Emberizidae; Passeriformes) is an important endangered species because it experienced a recent rapidly decline. It was one of the most abundant songbirds in northern Eurasia as recently as 1980s (Kamp et al., 2015). They breed from Finland in the west to Kamchatka in the east, and winter in South and Southeast Asia (International, 2021). Yellow-breasted Bunting was even considered a serious agricultural pest in the 20th century, as it often formed large flocks thousands-strong and fed on wheat and rice during the non-breeding season in China (Shaw, 1936; Zheng, 1956). However, this species is now considered to be critically endangered (International, 2021), with a global population decline of 84.3–94.7% from 1980 to 2013 (Kamp et al., 2015), which was mainly attributed to the excessive trapping at migration and wintering sites (IUCN, 2021; Kamp et al., 2015). As the speed and magnitude of the population decline is unprecedented, there was concern about its extinction. It is unclear whether the bunting retains the population's genetic capacity to recover.

To explore the genomic status of the Yellow-breasted bunting following its recent rapid population decline, and the species' genetic capacity to recover, we *de novo* assembled a whole genome of Yellow-breasted Bunting and sequenced whole genomes of ten individuals to estimate the genetic diversity,

¹Chengdu Research Base of Giant Panda Breeding, Sichuan Key Laboratory of Conservation Biology for Endangered Wildlife, Chengdu, 610081, P. R. China

²Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, P. R. China

³Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875, P. R. China

⁴Sichuan Academy of Giant Panda, Chengdu 610086, P. R. China

⁵Lead contact

*Correspondence: quepinjia@gmail.com (P.Q.), capricorncp@163.com (P.C.)
<https://doi.org/10.1016/j.isci.2022.104501>



inbreeding level, potential mutational load, and demographic history. The current study will provide genomic information for the conservation of Yellow-breasted Bunting.

RESULTS

Genome assembly and annotation

To estimate the whole genomic status and the amount of deleterious mutations of Yellow-breasted Bunting, we *de novo* assembled the genome and annotated its protein-coding regions. The length of the assembled Yellow-breasted Bunting genome was 1,149,406,387 bp (1.15 Gb) with the longest scaffold 37,240,703 bp long. The coverage of the Pacbio subreads was almost 162 x. The GC content in the Yellow-breasted Bunting genome was 43.16%. The scaffold N50 of the assembly was 8,222,108 bp (8.22 Mbp). The assembly covered 95.7% of BUSCOs and 91.94% of core gene, suggesting that the completeness of the assembly was good. The mapping rate and coverage of the Illumina reads were 98.84% and 99.87%, respectively, suggesting that the assembly has good consistency, integrity, and accuracy. Based on the alignment information, 153 scaffolds were identified as sex chromosome-linked scaffolds. The total length of these scaffolds was 106,560,505 bp, approximately 9.27% of the whole genome. The new assembly has been deposited in the Genome Warehouse in National Genomics Data Center (<https://ngdc.cncb.ac.cn/gwh>), Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bio-information, NGDC: GWHBJBD00000000.

In the genome of Yellow-breasted Bunting, 18.86% genome was identified as repeat sequences (Table S1). Based on the combined evidence, the Yellow-breasted Bunting genome was predicted to have 18,348 genes, among which 84.30% of genes assigned functions (Figures S1 and S2). The average transcript length and CDS lengths were 23,435.04 bp and 1,531.24 bp, respectively. There are 0.02% of the genome was predicted to be non-coding RNA (Table S2).

Yellow-breasted bunting has high genetic diversity

To perform population genetic analysis, we re-sequenced 10 Yellow-breasted Buntings (Table S3). The mean coverage and mapping rates of the 10 Yellow-breasted Buntings were 20.59 (standard deviation (SD) = 1.31) and 99.37% (SD = 0.10%), respectively (Table S3). After filtering, 18,938,565 high-quality SNPs were used in the population genetic analysis. There were no paired samples with relationship closer than 3rd-degree (Figure S3), so the 10 samples were all used in the population genetic analysis.

To understand the genomic status of Yellow-breasted Bunting among Passeriformes, we compared its genomic heterozygosity with 17 Passeriformes birds. The 17 Passeriformes birds included Florida scrub jay (*Aphelocoma coerulescens*), Noisy scrubbird (*Atrichornis clamosus*), Chestnut-collared longspur (*Calcarius ornatus*), Seychelles magpie robin (*Copsychus sechellarum*), Ribbon-tailed drongo (*Dicrurus megarhynchus*), Painted honeyeater (*Grantiella picta*), Akiapola'au (*Hemignathus wilsoni*), Philippine fairy-bluebird (*Irena cyanogastra*), Kirtland's warbler (*Setophaga kirtlandii*), Loggerhead shrike (*Lanius ludovicianus*), Bali myna (*Leucopsar rothschildi*), Yellowhead (*Mohoua ochrocephala*), Velvet myiagra (*Myiagra hebetior*), Inaccessible island finch (*Nesospiza acunhae*), Stitchbird (*Notiomystis cincta*), White-necked rockfowl (*Picathartes gymnocephalus*), and Golden-crowned babbler (*Sterrhoptilus dennistouni*). The result tells us the Yellow-breasted Bunting has high heterozygosity. The mean heterozygosity of the 10 samples is 0.46×10^{-2} (0.44×10^{-2} – 0.47×10^{-2}) (Table S4). The vulnerable chestnut-collared longspur has higher heterozygosity than Yellow-breasted Bunting (Figure 1, Table S4). Excepting the chestnut-collared longspur, all other Passeriformes assessed had lower heterozygosity than the Yellow-breasted Bunting, no matter their IUCN threatened status (Figure 1, Table S4).

Yellow-breasted bunting is not severely inbred

To understand the inbreeding pattern of Yellow-breasted Bunting, we constructed the linkage disequilibrium (LD) decay curve and calculated the amount of runs of homozygosity (ROH) in the species. The LD decay curve becomes smooth at a pairwise distance of less than 100 kb (Figure 2A). The maximum correlation coefficient between any two pairs of SNPs in a 100 bp windows is 0.18. The average total length of long and short ROH in the 10 samples is 1143.68 kb (SD = 2458.15) and 6153.12 kb (SD = 2165.16), respectively. The average percentage of long and short ROH on autosomes is 0.11% (SD = 0.25%) and 0.59% (SD = 0.22%), respectively (Figure 2B). The total length of short ROH is significantly larger than that of long ROH

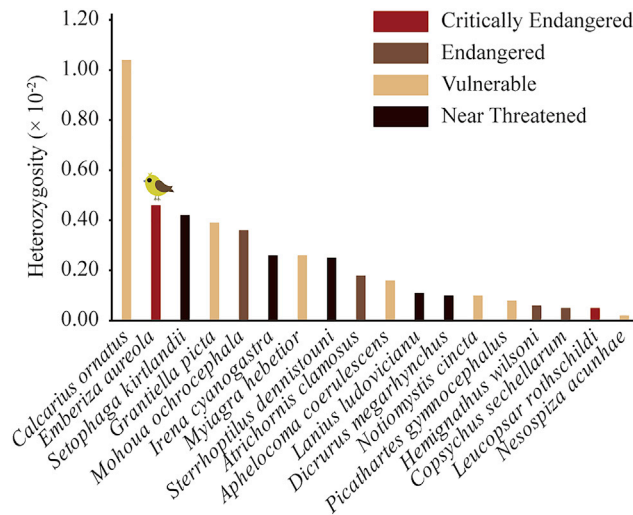


Figure 1. Heterozygosity of Yellow-breasted Bunting (indicated by yellow bird icon) compared to that of 17 other Passeriformes bird species

Heterozygosity equals the number of heterozygous SNPs divided by the length of the genome without gaps. The red, brown, sepia, and black colors correspond to critically endangered, endangered, vulnerable, and near-threatened status, respectively.

(KS test, $p < 0.01$). Both the LD decay curve and the percentage of ROH on autosomes suggest inbreeding is not serious in Yellow-breasted Bunting (Figure 2).

Potential mutational load

To recover the genetic load in Yellow-breasted Bunting, we identified the synonymous, missense, and loss of function (LOF) mutations in each sample. There were 130,622 synonymous, 61,977 missenses, 813 LOF mutations in Yellow-breasted Bunting. The frequency of homozygous sites is less than that of heterozygous sites for all three kinds of mutations (Figure 3A). Among 1000 random SNPs in coding sites, there was no significant difference in the relative frequency of homozygous sites between synonymous and missense mutations (Figure 3A). However, the relative frequency of homozygous sites of LOF mutations was significantly less than that of synonymous mutations (Figure 3A), suggesting that most LOF mutations are potential homozygous lethal mutations. There were 743 genes harboring LOF mutations, and they were significantly enriched in six KEGG pathways (Figure 3B). The six pathways are oocyte meiosis (gga04114), metabolic pathways (gga01100), NOD-like receptor signaling pathway (gga04621), apoptosis (gga04210), ECM-receptor interaction (gga04512), and C-type lectin receptor signaling pathways (gga04625).

Yellow-breasted bunting began to decline 100–200 generations ago

To reconstruct the demographic history of the Yellow-breasted bunting, we employed three methods: PSMC, stairway plot, and fastsimcoal. The complementary methods provide a detailed demographic history of Yellow-breasted Bunting, suggesting that Yellow-breasted Bunting experienced a bottleneck, and that the most recent population decline started 100–200 generations ago (Figures 4 and 5C). PSMC and the stairway plot both suggest Yellow-breasted Bunting began to expand about 2×10^6 years ago (Figure 4). After a period of population stability, a decline began about 4×10^5 years ago (Figures 4A and 4B). Recovery likely began about 2×10^5 years ago (Figure 4A). The best fit model inferred by fastsimcoal is consistent with the results of PSMC and stairway plot (Figure 5C). The best fit model supports Yellow-breasted Bunting having experienced a bottleneck and the recently decline of the population started at about 147 (95% confidence interval (95% CI) = 146–149) generations ago. The average current population size of Yellow-breasted Bunting is 5119488 (95% CI = 5117928–5121048). Before the recent population decline, the population size of Yellow-breasted Bunting was 7320650 (95% CI = 7319004–7322296). Nearly 30.07% of the population disappeared. The Akaike Information Criterion (AIC) value of the best model is 49761.58, and its average likelihood from 100 replications was a significant difference from other models ($p < 0.01$).

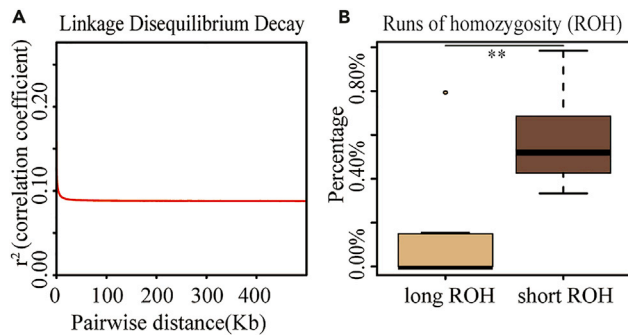


Figure 2. Linkage disequilibrium (LD) decay and runs of homozygosity (ROH) in Yellow-breasted Bunting

(A) Linkage disequilibrium decay of Yellow-breasted Bunting. The r^2 is a correlation coefficient between any pair of SNPs.

(B) Box plot of runs of homozygosity (ROH) in Yellow-breasted Bunting. The percentage is the total length of ROH divided by the length of autosomes \times 100%. Long ROH are defined as those longer than 1.5 Mb and short ROH as shorter than 1 Mb. A Kolmogorov-Smirnov test suggests the total length of short ROH is significantly larger than that of long ROH. ** means the difference is significant.

DISCUSSION

The current project assembled the whole genome of Yellow-breasted Bunting and combined it with population genomic data to infer the genomic status of the recent rapid population decline of Yellow-breasted Bunting, and to determine whether the species could recover its abundant population size. The results suggest Yellow-breasted Bunting has rich genetic diversity, and low levels of inbreeding, but potentially homozygous lethal mutations. Demographic history shows Yellow-breasted Bunting has experienced a bottleneck and the last population decline began about 147 generations ago.

Genomic status following recent rapid population decline differs from long-term decline

Wild populations that experienced long-term declines may have low genetic diversity, high levels of inbreeding, and genetic load. In small populations, genetic drift would eliminate genetic diversity (Eilstrand and Elam, 1993; Feng et al., 2019). Individuals in small populations have much more opportunity to be inbred compared to that in large populations. In turn, inbreeding leads small populations to accumulate deleterious mutations (Charlesworth and Willis, 2009). Empirical studies have found that wild populations that experienced long-term populations have low genetic diversity, high frequency of inbreeding, and serious mutation loads. Such species include Brown Eared Pheasant (*Crossoptilon mantchuricum*) (Wang et al., 2021) and Crested Ibis (*Nipponia nippon*) (Feng et al., 2019). Long-term declines make populations have small population size, which may facilitate genetic drift toward low genetic diversity, high levels of inbreeding, and genetic load.

Compared with populations that experienced long-term declines, those experiencing recent rapid declines may have high genetic diversity, low levels of inbreeding, and low mutation load. The populations declining recently and rapidly may have relatively large population sizes, in which case genetic drift has less power to erase the genetic variations. Over short time frames, the population may accumulate most genetic characteristics from its parental population facing genetic drift.

The Yellow-breasted Bunting population began to decline about 100–200 generations ago (Figure 5). Even though genetic drift could act on the Yellow-breasted Bunting, the species has high genetic diversity (Figure 1). During most of the recent decline, the Yellow-breasted Bunting may have had a large population size, which makes genetic drift have less power to eliminate the genetic diversity. The short bottleneck, nearly 147 generations, may be the other reason that explains Yellow-breasted Bunting has high genetic diversity. The declining population may need a long time to lose genetic diversity. The Yellow-breasted Bunting has low levels of inbreeding (Figure 2), which may be indirect evidence to support a large recent population size because individuals have less chance of being inbred with each other in larger populations. The low level of inbreeding will help the population have only a slight mutation load. The result suggests Yellow-breasted Bunting has potential homozygous lethal mutations (Figure 3), which may be from a historical bottleneck (Figure 4).

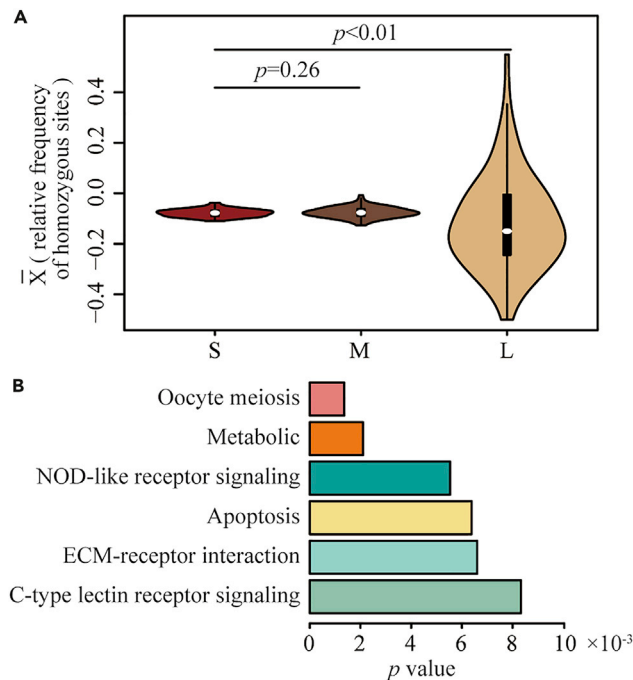


Figure 3. Mutation load and enriched KEGG pathways of the genes that harbor loss-of-function (LOF) mutations in Yellow-breasted Bunting

(A) Average difference between homozygous site frequency and heterozygous site frequency. The y-axis is the relative frequency of homozygous sites and negative values mean the frequency of heterozygous sites is larger than that of homozygous sites. S, synonymous mutations; M, missense mutations; L, loss of function mutations. Kolmogorov-Smirnov comparison was used to calculate the p value.

(B) The significant enriched KEGG pathways. Fisher's exact test was used to calculate the p value.

The yellow-breasted bunting has the capability to recover

Genetic diversity is an important metric in conservation and biodiversity, because it is intimately tied to fitness and adaptive potential (DeWoody et al., 2021). In taxonomically diverse organisms, such as fruit flies (*Drosophila melanogaster*) (Frankham, 1995), field mice (*Peromyscus leucopus*) (Lacy et al., 2013), and monkeyflower (*Mimulus guttatus*) (Willis, 1993), genetic diversity is positively related to fitness (Chapman et al., 2009). In threatened desert tortoises (*Gopherus agassizii*), individual heterozygosity predicts translocation success (Scott et al., 2020). Besides fitness, adaptive potential is correlated with genetic diversity. For example, the genetic diversity in the coding region and conserved non-coding region is low in Brown Eared Pheasant, a species with low genome-wide diversity (Wang et al., 2021). Moreover, neutral diversity in the genome provides a genetic template for the future evolution of adaptive diversity (Harrison et al., 2014). The rich genetic diversity in the Yellow-breasted Bunting population (Figure 1) suggests it has high fitness and adaptive potential relative to other endangered species, such as Painted Honeyeater (*Grantiella picta*) and Velvet Myiagra (*Myiagra hebetior*) (Figure 1), which could contribute to its recovery. For example, the genome-wide heterozygosity of crested ibis is 0.04×10^{-2} (Li et al., 2014), much less than the heterozygosity of Yellow-breasted Bunting (0.46×10^{-2}). Only seven crested ibis remained in the world in 1981, but the species has recovered to more than 5,000 individuals over the last 40 years (Valchuk, 2001). The rich genetic diversity in Yellow-breasted Bunting and the successful conservation efforts of crested ibis give us confidence that the Yellow-breasted Bunting has the genetic ability to recover.

Inbreeding can accelerate the extinction of threatened species because it can accelerate the spread of potential homozygous lethal mutations through inheritance (Frankham, 1995; Willis, 1993). In wild populations of the Glanville fritillary butterfly (*Melitaea cinxia*) and in laboratory cultures of fruit flies (*Drosophila simulans*), extinction risk increased significantly with increasing inbreeding (Saccheri et al., 1998; Wright et al., 2007). We did not find genetic signals of recent inbreeding in Yellow-breasted Buntings because the species has short LD extension and ROH (Figure 2) (Ceballos et al., 2018).

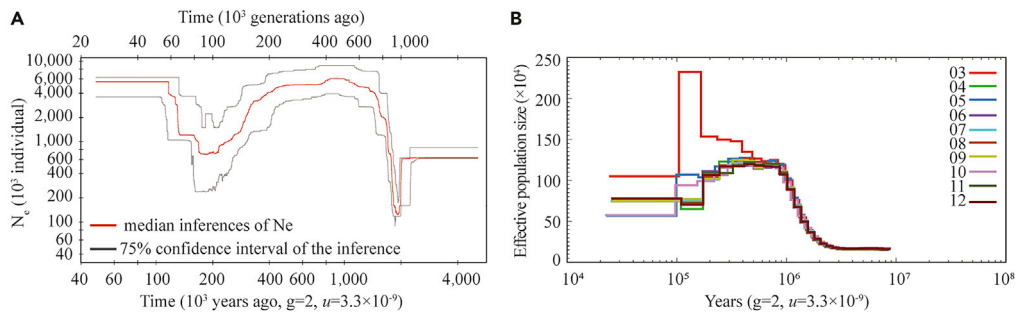


Figure 4. The fluctuations of historical effective population size (N_e)

(A) Historical N_e inferred by stairway plot. The red line represents the median inferred N_e . The grey line represents the 75% confidence interval.

(B) Historical N_e inferred by pairwise sequentially Markovian coalescent (PSMC). The numbers before the colored lines are the numbers in the sample IDs in Table S3. The generation time was assumed to be two years. The mutation rate was set as 3.3×10^{-9} per site per generation.

In Yellow-breasted Bunting, the average difference between homozygous and heterozygous frequency in LOF sites is significantly lower than that in synonymous sites (Figure 3), suggesting that the homozygous LOF sites are under purifying selection. The homozygous LOF mutations may reduce the survival rate. Although the Yellow-breasted Bunting may have potential homozygous lethal mutations, low levels of inbreeding may restrain the negative influence of these mutations. If we protect the remaining wild population, and conserve its habitat, the Yellow-breasted Bunting could recover, despite having potential homozygous lethal mutations. Low levels of inbreeding could reduce the negative influence of potential homozygous lethal mutations on survival. Certainly, we need to test the effects of the potential homozygous lethal mutations on fitness in the future.

The factors linked with rapid population decline

Natural selection, climate change, and human activities are three factors that affect the abundance of wild species (Feng et al., 2019; Murray et al., 2017; Zhao et al., 2012). Natural selection reduces genetic diversity through linkage selection (Buffalo, 2021); however, the LD extension in Yellow-breasted Bunting is short (Figure 2A). The short LD extension suggests natural selection has a lower probability of reducing genetic diversity through linkage selection in Yellow-breasted Bunting. The high genetic diversity of Yellow-breasted Bunting (Figure 1) also suggests natural selection may be not the main factor in prior population declines.

Climate change could act species' abundance and cause extinction, and is in fact inferred as the one driver of long-term fluctuations in giant panda population size (Zhao et al., 2012). Demographic history suggests the Yellow-breasted Bunting experienced a bottleneck between 10^5 and 4×10^5 years ago (Figure 4) in the Middle Pleistocene. This period was characterized by climate oscillations, suggesting that climate change may have caused the bunting bottleneck. The Yellow-breasted Bunting began to decline about 100–200 generations ago (Figure 5). Anthropogenic global climate change has only been progressing since the late 19th century (Zanna et al., 2019), so it is unlikely that climate change directly caused the recent population decline of Yellow-breasted Bunting.

Interestingly, our study suggests that the time of decline in Yellow-breasted Bunting was much earlier than the prevailing view that the decline has mainly occurred since the 1980s (Zheng, 1956). Human activities have dramatically changed the global environment in many aspects as the industrial revolution, including climate warming, pollution, and fragmentation or loss of habitat (Steffen et al., 2016). All these processes could lead to population declines or extinction. As the industrial revolution, the global extinction rate is estimated to be at least 100 times greater than that in the fossil record (Hassan et al., 2005). Because the time that the Yellow-breasted Bunting began to decline is similar to the time of the industrial revolution, it is reasonable to infer that the decline may be directly or indirectly caused by human activities.

Human activities may be one driver of the population decline. The Yellow-breasted Bunting mainly breeds in wet meadows with scrub nearby. This includes wetlands in Russia. The species winters in South-east Asia and South China (11). In its place, cropland and pastures have increased by 358.1 million ha during 1850–1990

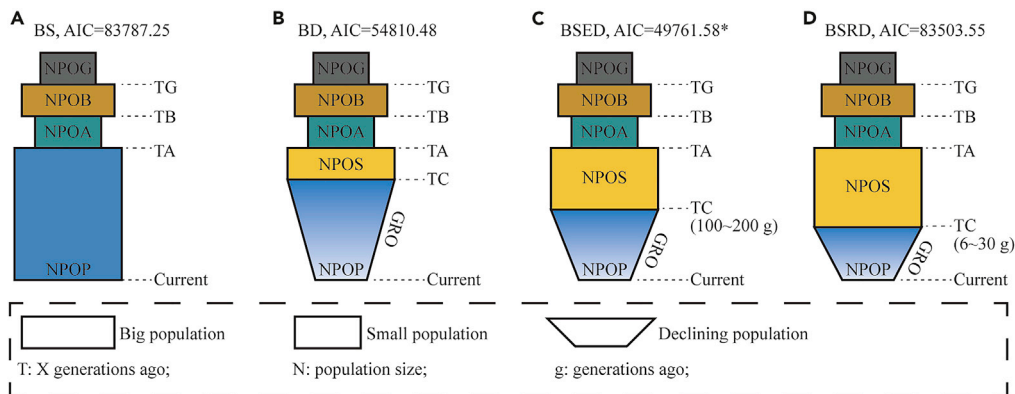


Figure 5. The demographic history models estimated by fastsimcoal

(A) Bottleneck and stable model (BS).

(B) Bottleneck and decline model (BD).

(C) Bottleneck, ancestry stable, and early decline model (BSED).

(D) Bottleneck, ancestry stable, and recently decline model (BSRD). NPOA, NPOP, NPOS, NPOG, and NPOB represent the effective population size (N_e) of the different historical stages. TG, TB, TA, and TC represent the time that N_e began to change; Current, 2019 A. D.; GRO, population decline rate; g, generations ago; AIC, Akaike Information Criterion.

(Goldewijk, 2001). Hence, habitat loss caused by human activity could be the driver of the recent population decline. In addition, unsustainable trapping of the species is one driver of the recent decline. The Yellow-breasted Bunting has historically been trapped for food by the indigenous people in South China for at least 2,100 years (Yuren, 1996). China's population has increased by about one billion people in the past 200 years (Poston and Yaukey, 2013), which could exacerbate anthropogenic pressure causing the Yellow-breasted Bunting decline. Furthermore, environmental pollution and climate change as the industrial revolution also could play an important role during the population decline (Gils et al., 2016; Kubelka et al., 2018; Zheng, 1956).

However, we still know little about the specific mechanisms of population decline in Yellow-breasted Bunting. We need to conduct many more studies, e.g., to determine the effects of heavy metal pollution on fitness, and the effects of habitat change coupled with climate warming, to infer the specific reason for population decline in the future.

Conservation implications

With a *de novo* genome assembly and population genomic data from 10 individuals, we show that Yellow-breasted Buntings have high genetic diversity and low levels of inbreeding, giving confidence in the potential to recover the species. Increasing population size over the long term may be the most important management goal, especially when the population has the mutation load and inbreeding is associated with inbreeding depression. It will be critical to increase its habitat size in future conservation efforts. Helping the local community to increase the paddy and wheat fields without the use of pesticides may be an effective method to increase the habitat size and population size of Yellow-breasted Bunting. It is critical to understand the fluctuations of the local population size and genetic variation with non-invasive sampling during conservation management.

On the other hand, protecting Yellow-breasted Bunting from illegal trapping is key to conservation. The Chinese government has placed Yellow-breasted Bunting in the highest protection classes (Ministry of Agriculture and Rural affairs of the People's Republic of China, 2021), and imposed a ban on wildlife trade and consumption Zhang (2020). As a result, some positive signs have emerged. The number of criminal cases involving Yellow-breasted Bunting in the last two years has been significantly lower than before. A number of small breeding populations also recovered considerably in Russia (Heim et al., 2021). Ongoing and increased collaboration will help gain a better understanding of genomic status across its whole population and better advice for conservation efforts.

Limitations of the study

The current project used 10 samples to investigate the genetic status of Yellow-breasted Bunting, which did not cover all its populations. In future, evaluating the genetic status of all the breeding populations is necessary. What is more, we need to monitor the changes in genetic status for lasting years.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- METHOD DETAILS
 - Sampling
 - DNA/RNA extraction and sequencing
 - Genome assembly and evaluation
 - Genome annotation
 - Identifying sex chromosome-linked scaffolds
 - Read mapping and SNP calling
 - Family relationship inference
 - Estimating genetic diversity
 - Estimating inbreeding pattern
 - Estimating potential inbreeding depression
 - Reconstructing yellow-breasted bunting demographic history

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104501>.

ACKNOWLEDGMENTS

We especially thank Zhiwei Tian who provided much help with field sampling. We give our best thanks to Yang Liu from State Key Laboratory of Biocontrol, School of Ecology, Sun Yat-sen University for his advice on writing the article. We also appreciate Xuming Zhou from the Chinese Academy of Sciences for providing us with computer cluster access. We give our best thanks to the editor and two anonymous reviewers. This research was supported by Chengdu Research Base of Giant Panda Breeding (NO. 2020CPB-C09, NO. 2020CPB-C08, NO. CPB2018-02).

AUTHOR CONTRIBUTIONS

Conceptualization: PCW, P.J.Q.

Methodology: PCW, YW.

Investigation: PCW, RH.

Funding acquisition: P.J.Q, PC.

Writing – original draft: PCW, P.J.Q, ZWZ.

Writing – review & editing: PCW, P.J.Q, PC.

DECLARATION OF INTERESTS

Authors declare that they have no competing interests.

Received: January 31, 2022

Revised: May 4, 2022

Accepted: May 26, 2022

Published: July 15, 2022

REFERENCES

- Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28, 45–48. <https://doi.org/10.1093/nar/28.1.45> %.
- Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* 6, 11. <https://doi.org/10.1186/s13100-015-0041-9>.
- Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2004). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. <https://doi.org/10.1093/bioinformatics/bth457>.
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. <https://doi.org/10.1093/nar/27.2.573>.
- Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. *Genome Res.* 14, 988–995. <https://doi.org/10.1101/gr.1865504>.
- Blanco, E., Parra, G., and Guigó, R. (2007). Using geneid to identify genes. *Curr. Protoc. Bioinform.* 18. <https://doi.org/10.1002/0471250953.bi0403s18>.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., Wu, Y., Zhao, L., Liu, J., Guo, J., et al. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* 49, W317–W325. <https://doi.org/10.1093/nar/gkab447> %.
- Buffalo, V. (2021). Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's paradox. *Elife* 10, e67509. <https://doi.org/10.7554/eLife.67509>.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421. <https://doi.org/10.1186/1471-2105-10-421>.
- Cantarel, B.L., Korff, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., Sánchez Alvarado, A., and Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18, 188–196. <https://doi.org/10.1101/gr.6743907>.
- Ceballos, F.C., Joshi, P.K., Clark, D.W., Ramsay, M., and Wilson, J.F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* 19, 220–234. <https://doi.org/10.1038/nrg.2017.109>.
- Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>.
- Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J., and Sheldon, B.C. (2009). A quantitative review of heterozygosity–fitness correlations in animal populations. *Mol. Ecol.* 18, 2746–2765. <https://doi.org/10.1111/j.1365-294X.2009.04247.x>.
- Charlesworth, D., and Willis, J.H. (2009). The genetics of inbreeding depression. *Nat. Rev. Genet.* 10, 783–796. <https://doi.org/10.1038/nrg2664>.
- Chen, N. (2004). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* 25, 4. <https://doi.org/10.1002/0471250953.bi0410s25>.
- Ministry of Agriculture and Rural affairs of the People's Republic of China (2021). Updated list of nationally key protected wild animals, containing 980 species in 8 categories. http://www.moa.gov.cn/xw/zwdt/202102/t20210205_6361251.htm.
- Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff, SnpEff. *Fly (Austin)* 6, 80–92. <https://doi.org/10.4161/fly.19695>.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. <https://doi.org/10.1093/gigascience/giab008>.
- DeWoody, J.A., Harder, A.M., Mathur, S., and Willoughby, J.R. (2021). The long-standing significance of genetic diversity in conservation. *Mol. Ecol.* 30, 4147–4154. <https://doi.org/10.1111/mec.16051>.
- Dong, F., Kuo, H.-C., Chen, G.-L., Wu, F., Shan, P.-F., Wang, J., Chen, D., Lei, F.-M., Hung, C.-M., Liu, Y., and Yang, X.-J. (2021). Population genomic, climatic and anthropogenic evidence suggest the role of human forces in endangerment of green peafowl (*Pavo muticus*). *Proc. Biol. Sci.* 288, 20210073. <https://doi.org/10.1098/rspb.2021.0073>.
- Ellstrand, N.C., and Elam, D.R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annu. Rev. Ecol. Systemat.* 24, 217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>.
- Excoffier, L., and Foll, M. (2011). fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* 27, 1332–1334. <https://doi.org/10.1093/bioinformatics/btr124>.
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genet.* 9, e1003905. <https://doi.org/10.1371/journal.pgen.1003905>.
- Feng, S., Fang, Q., Barnett, R., Li, C., Han, S., Kuhlwillm, M., Zhou, L., Pan, H., Deng, Y., Chen, G., et al. (2019). The genomic footprints of the fall and recovery of the Crested Ibis. *Curr. Biol.* 29, 340–349.e7. <https://doi.org/10.1016/j.cub.2018.12.008>.
- Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A.H., Xie, D., Chen, G., Guo, C., Faircloth, B.C., et al. (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature* 587, 252–257. <https://doi.org/10.1038/s41586-020-2873-9>.
- Frankham, R. (1995). Inbreeding and extinction: a threshold effect. *Conserv. Biol.* 9, 792–799. <https://doi.org/10.1046/j.1523-1739.1995.09040792.x>.
- Gils, J.A.v., Lisovski, S., Lok, T., Meissner, W., Ożarowska, A., de Fouw, J., Rakhimberdiev, E., Soloviev, M.Y., Piersma, T., and Klaassen, M. (2016). Body shrinkage due to Arctic warming reduces red knot fitness in tropical wintering range. *Science* 352, 819–821. <https://doi.org/10.1126/science.aad6351>.
- Goldewijk, K.K. (2001). Estimating global land use change over the past 300 years: the HYDE Database: The HYDE Database. *Global Biogeochem. Cycles* 15, 417–433. <https://doi.org/10.1029/1999GB001232>.
- Gonzalez-Garay, M.L. (2016). Introduction to isoform sequencing using pacific biosciences technology (Iso-Seq). In *Transcriptomics and Gene Regulation*, J. Wu, ed. (Springer Netherlands), pp. 141–160. https://doi.org/10.1007/978-94-017-7450-5_6.
- Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., and Bateman, A. (2005). Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 33, D121–D124. <https://doi.org/10.1093/nar/gki081>.
- Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 9, R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
- Harrisson, K.A., Pavlova, A., Telonis-Scott, M., and Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol. Appl.* 7, 1008–1025. <https://doi.org/10.1111/eva.12149>.
- Hassan, R., Scholes, R., and Ash, N. (2005). *Ecosystems and Human Well-Being: Current State and Trends* (Island Press).

- Heim, W., Chan, S., Hölzel, N., Ktitorov, P., Mischenko, A., and Kamp, J. (2021). East Asian buntings: ongoing illegal trade and encouraging conservation responses. *Conserv. Sci. Prac.* 3, e405. <https://doi.org/10.1111/csp2.405>.
- Howrigan, D.P., Simonson, M.A., and Keller, M.C. (2011). Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genom.* 12, 460. <https://doi.org/10.1186/1471-2164-12-460>.
- Hu, J., Fan, J., Sun, Z., and Liu, S. (2019). NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics* 36, 2253–2255. <https://doi.org/10.1093/bioinformatics/btz891>.
- International, B. (2021). Species Factsheet: *Emberiza Aureola*. <https://doi.org/10.1007/s11516-006-0014-3>. <http://www.birdlife.org/on/09/11/2021>.
- IUCN (2021). The IUCN Red List of Threatened Species. Version 2021-2. <https://www.iucnredlist.org>.
- Kamp, J., Oppel, S., Ananin, A.A., Durnev, Y.A., Gashev, S.N., Hölzel, N., Mishchenko, A.L., Pessa, J., Smirenski, S.M., Strelnikov, E.G., et al. (2015). Global population collapse in a superabundant migratory bird and illegal trapping in China. *Conserv. Biol.* 29, 1684–1694. <https://doi.org/10.1111/cobi.12537>.
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
- Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics* 5, 59. <https://doi.org/10.1186/1471-2105-5-59>.
- Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* 15, 356. <https://doi.org/10.1186/s12859-014-0356-4>.
- Kronenberg, Z.N., Rhie, A., Koren, S., Concepcion, G.T., Peluso, P., Munson, K.M., Hiendleder, S., Fedrigo, O., Jarvis, E.D., Phillippy, A.M., et al. (2019). Extended haplotype phasing of de novo genome assemblies with FALCON-Phase. Preprint at bioRxiv. <https://doi.org/10.1101/327064> %J bioRxiv.
- Kubelka, V., Šálek, M., Tomkovich, P., Végvári, Z., Végvári, Z., Freckleton, R.P., Székely, T., and Székely, T. (2018). Global pattern of nest predation is disrupted by climate change in shorebirds. *Science* 362, 680–683. <https://doi.org/10.1126/science.aat8695>.
- Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5, R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
- Lacy, R.C., Alaks, G., and Walsh, A. (2013). Evolution of *Peromyscus leucopus* mice in response to a captive environment. *PLoS One* 8, e72452. <https://doi.org/10.1371/journal.pone.0072452>.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature* 475, 493–496. <https://doi.org/10.1038/nature10231>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, S., Li, B., Cheng, C., Xiong, Z., Liu, Q., Lai, J., Carey, H.V., Zhang, Q., Zheng, H., Wei, S., et al. (2014). Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biol.* 15, 557. <https://doi.org/10.1186/s13059-014-0557-1>.
- Liu, X., and Fu, Y.-X. (2015). Exploring population size changes using SNP frequency spectra. *Nat. Genet.* 47, 555–559. <https://doi.org/10.1038/ng.3254>.
- Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. <https://doi.org/10.1093/nar/25.5.955>.
- Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* 20, 2878–2879. <https://doi.org/10.1093/bioinformatics/bth315>.
- Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics* 26, 2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Mulder, N., and Apweiler, R. (2007). InterPro and InterProScan. In *Comparative Genomics*, N.H. Bergman, ed. (Humana Press), pp. 59–70. https://doi.org/10.1007/978-1-59745-515-2_5.
- Murray, G.G.R., Soares, A.E.R., Novak, B.J., Schaefer, N.K., Cahill, J.A., Baker, A.J., Demboski, J.R., Doll, A., Da Fonseca, R.R., Fulton, T.L., et al. (2017). Natural selection shaped the rise and fall of passenger pigeon genomic diversity. *Science* 358, 951–954. <https://doi.org/10.1126/science.aao0960>.
- Nawrocki, E.P., and Eddy, S.R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935. <https://doi.org/10.1093/bioinformatics/btt509>.
- Parra, G., Bradnam, K., Ning, Z., Keane, T., and Korf, I. (2008). Assessing the gene space in draft genomes. *Nucleic Acids Res.* 37, 289–297. <https://doi.org/10.1093/nar/gkn916> %.
- Poston, D.L., Jr., and Yaukey, D. (2013). *The Population of Modern China* (Springer Science & Business Media).
- Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C., Panjabi, A., Helft, L., Parr, M., and Marra, P.P. (2019). Decline of the north American avifauna. *Science* 366, 120–124. <https://doi.org/10.1126/science.aaw1313>.
- Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., and Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature* 392, 491–494. <https://doi.org/10.1038/33136>.
- Scott, P.A., Allison, L.J., Field, K.J., Averill-Murray, R.C., and Shaffer, H.B. (2020). Individual heterozygosity predicts translocation success in threatened desert tortoises. *Science* 370, 1086–1089. <https://doi.org/10.1126/science.abb0421>.
- Shaw, T.-h. (1936). *The Birds of Hopei Province* (Fan Memorial Institute of Biology).
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34, W435–W439. <https://doi.org/10.1093/nar/gkl200>.
- Steffen, W., Crutzen, P.J., and McNeill, J.R. (2016). *The Anthropocene: are humans now overwhelming the great forces of nature? In The New World History* (University of California Press), pp. 440–459.
- Tiwari, S., Ramachandran, S., Bhattacharya, A., Bhattacharya, S., and Ramaswamy, R. (1997). Prediction of probable genes by Fourier analysis of genomic sequences. *Bioinformatics* 13, 263–270. <https://doi.org/10.1093/bioinformatics/13.3.263>.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. <https://doi.org/10.1093/bioinformatics/btp120>.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. <https://doi.org/10.1038/nbt.1621>.
- Valchuk, O. (2001). *Threatened Birds of Asia: The BirdLife International Red Data Book* (BirdLife International).
- Wang, P., Burley, J.T., Liu, Y., Chang, J., Chen, D., Lu, Q., Li, S.-H., Zhou, X., Edwards, S., and Zhang, Z. (2021). Genomic consequences of long-term population decline in Brown Eared Pheasant. *Mol. Biol. Evol.* 38, 263–273. <https://doi.org/10.1093/molbev/msaa213>.
- Willis, J.H. (1993). Effects of different levels of inbreeding on fitness components in *mimulus*

guttatus. *Evolution* 47, 864–876. <https://doi.org/10.1111/j.1558-5646.1993.tb01240.x>.

Wright, L.I., Tregenza, T., and Hosken, D.J. (2007). Inbreeding, inbreeding depression and extinction. *Conserv. Genet.* 9, 833–843. <https://doi.org/10.1007/s10592-007-9405-0>.

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 35, W265–W268. <https://doi.org/10.1093/nar/gkm286> %.

Yuren, G. (1996). Ginseng from heaven—rice birds yellow-breasted buntings. *Daziran* 1, 34–35.

Zanna, L., Khatiwala, S., Gregory, J.M., Ison, J., and Heimbach, P. (2019). Global reconstruction of historical ocean heat storage and transport. *Proc. Natl. Acad. Sci. U S A* 116, 1126–1131. <https://doi.org/10.1073/pnas.1808838115> %.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., Storz, J.F., Antunes, A., Greenwold, M.J., Meredith, R.W., et al. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346, 1311–1320. <https://doi.org/10.1126/science.1251385>.

Zhang, M. (2020). Decision of the Standing Committee of the National People's Congress

on Banning Illegal Wildlife Trade, the Consumption of Wild Animals and Protecting People's Health and Safety (The National People's Congress of the People's Republic of China).

Zhao, S., Zheng, P., Dong, S., Zhan, X., Wu, Q., Guo, X., Hu, Y., He, W., Zhang, S., Fan, W., et al. (2012). Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nat. Genet.* 45, 67–71. <https://doi.org/10.1038/ng.2494>.

Zheng, Z. (1956). Sparrow and sparrow damage. *Chin. Sci. Bull.* 2, 60–63.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
The blood of Yellow-breasted Buntings (<i>Emberiza aureola</i>).	Daqinghe Bird Rescue Center, Tangshan, Hebei, China.	N/A
Deposited data		
The reference genomes and the fragment genome libraries of Florida scrub jay (<i>Aphelocoma coerulescens</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253992
The reference genomes and the fragment genome libraries of noisy scrubbird (<i>Atrichornis clamosus</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253955
The reference genomes and the fragment genome libraries of chestnut-collared longspur (<i>Calcarius ornatus</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253923
The reference genomes and the fragment genome libraries of Seychelles magpie robin (<i>Copsychus sechellarum</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12254010
The reference genomes and the fragment genome libraries of ribbon-tailed drongo (<i>Dicrurus megarhynchus</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253804
The reference genomes and the fragment genome libraries of painted honeyeater (<i>Grantiella picta</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253950
The reference genomes and the fragment genome libraries of Akiapola'au (<i>Hemignathus wilsoni</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN10867508
The reference genomes and the fragment genome libraries of Philippine fairy-bluebird (<i>Irena cyanogastra</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253784
The reference genomes and the fragment genome libraries of Kirtland's warbler (<i>Setophaga kirtlandii</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253801
The reference genomes and the fragment genome libraries of loggerhead shrike (<i>Lanius ludovicianus</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253815
The reference genomes and the fragment genome libraries of Bali myna (<i>Leucopsar rothschildi</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253829
The reference genomes and the fragment genome libraries of yellowhead (<i>Mohoua ochrocephala</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253963
The reference genomes and the fragment genome libraries of velvet myiagra (<i>Myiagra hebetior</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253791
The reference genomes and the fragment genome libraries of Inaccessible Island finch (<i>Nesospiza acunhae</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12254004

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
The reference genomes and the fragment genome libraries of stitchbird (<i>Notiomystis cincta</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253957
The reference genomes and the fragment genome libraries of white-necked rockfowl (<i>Picathartes gymnocephalus</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253907
The reference genomes and the fragment genome libraries of golden-crowned babbler (<i>Sterrhoptilus dennistouni</i>)	National Center for Biotechnology Information search database (NCBI)	SAMN12253785
chicken genome	Ensembl	ftp://ftp.ensembl.org/pub/release-98/fasta/gallus_gallus/dna/
Raw sequencing data	This paper	NGDC: PRJCA009403

Software and algorithms

FALCON	Kronenberg et al. (2019)	https://github.com/PacificBiosciences/FALCON
Nextpolish	Hu et al. (2019)	https://github.com/Nextomics/NextPolish
RepeatMasker v. 4.1.2	N/A	http://www.repeatmasker.org/RepeatModeler/
RepeatModeler v. 2.0.3	N/A	http://www.repeatmasker.org/RepeatModeler/
SNAP v. 20131129	Korf (2004)	http://korflab.ucdavis.edu/software.html
Augustus v. 3.4.0	Stanke et al., (2006)	http://bioinf.uni-greifswald.de/augustus/
Maker v. 3.1.4	Cantarel et al., (2008)	http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/Main_Page
MUMmer v. 3.23	Kurtz et al. (2004)	http://mummer.sourceforge.net/
Burrows-Wheeler Alignment Tool (BWA) v. 0.7.17	Li and Durbin (2009)	http://bio-bwa.sourceforge.net/
Samtools v. 1.13	Li et al. (2009)	http://samtools.sourceforge.net/
GATK v. 4.1.4.0	McKenna et al. (2010)	https://gatk.broadinstitute.org/hc/en-us/sections/360007279452-4-1-4-0
vcftools v. 0.1.13	Danecek et al., 2021	http://vcftools.sourceforge.net/
KING v. 2.2.7	Manichaikul et al. (2010)	https://www.kingrelatedness.com/
Haploview v. 4.2	Barrett et al. (2004)	https://www.broadinstitute.org/haploview/haploview
PLINK	Howrigan et al. (2011)	https://www.cog-genomics.org/plink/
SnPEff v. 5.0	Cingolani et al. (2012)	http://pcingola.github.io/SnpEff/
Pairwise Sequentially Markovian Coalescent (PSMC)	Li and Durbin (2011)	https://github.com/lh3/psmc
Stairway plots	Liu and Fu (2015)	https://github.com/xiaoming-liu/stairway-plot-v22
fastsimcoal26	Excoffier and Foll (2011)	http://cmpg.unibe.ch/software/fastsimcoal26/
easySFS	https://github.com/isaacovercast/easySFS	https://github.com/isaacovercast/easySFS
ANGSD v. 0.910	Korneliussen et al. (2014)	http://www.popgen.dk/angsd/index.php/ANGSD

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Peng Chen (capricorncp@163.com).

Materials availability

All the samples in the study could be shared with researchers who aim to conserve Yellow-breasted Buntings. Further requests for the samples will be fulfilled by the first author, Pengcheng Wang (wpc@nynu.edu.cn).

Data and code availability

All the raw sequencing data and the assembly have been deposited in the National Genomics Data Center (NGDC, <https://ngdc.cncb.ac.cn/?lang=en>), NGDC: PRJCA009403.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

METHOD DETAILS

Sampling

On 15 September 2019, local policemen found Yellow-breasted Buntings were illegally trapped by mist net at Tangshan, East China, and brought them to Daqinghe Bird Rescue Center. Two buntings died soon after they arrived at the rescue center. These two Yellow-breasted Buntings were refrigerated immediately, and their tissues were sampled for *de novo* assembly and to annotate the genome. To analyze the population genomic status of Yellow-breasted Bunting, trace brachial vein blood was sampled from the remaining 10 rescued Yellow-breasted Bunting. After sampling and physical exam, the 10 healthy Yellow-breasted Buntings were released to the wild. All samples were kept in a -80°C freezer in College of life sciences, Nanjing Normal University. Because the Yellow-breasted Bunting is a critically endangered species, the current project did not collect samples from all populations. All sample collection was done by an experienced veterinarian and sample information can be found in [Table S3](#).

DNA/RNA extraction and sequencing

DNA and RNA extraction was completed using the phenol-chloroform method and trizol method, respectively. After passing a quality inspection, the DNA from fresh muscle was used to construct Illumina (Insert size was 350 bp, target data volume was 120 Gb) and Pacbio (target data volume was 120 Gb) genome libraries for genome assembly. The qualified RNA was used to construct Illumina (Insert size was 250–300 bp, target data volume was 10 Gb for each tissue) and Pacbio (target data volume was 30 Gb for each tissue) transcriptome libraries for genome annotation. The 10 qualified DNA samples from blood were used to construct Illumina DNA libraries for whole genome sequencing (Insert size was 350 bp, target data volume was 30 Gb for each sample). The Illumina DNA and RNA libraries were sequenced on Illumina HiSeq and Illumina NovaSeq 6000 platforms with paired-end 150 bp sequencing strategies, respectively. The Pacbio libraries were sequenced on Pacbio Sequel sequencing platform. The DNA/RNA extraction, genome/transcriptome library construction and sequencing were completed by the Novogene sequencing company, Beijing.

Genome assembly and evaluation

Nearly 186 Gb Pacbio subreads were used to assemble the genome in FALCON software ([Kronenberg et al., 2019](#)). After removing the reads whose N proportion was larger than 10%, and any low-quality reads (quality score less than 5), we retained reads with more than 30 bp, nearly 128 Gb of clean Illumina reads to polish the genome in the Nextpolish software ([Hu et al., 2019](#)). After polishing, we calculated the Scaffold N50 of the assembly. The Benchmarking Universal Single-Copy Orthologs (BUSCO) Aves dataset ([Simão et al., 2015](#)) and Core Eukaryotic Genes Mapping Approach (CEGMA) core gene dataset ([Parra et al., 2008](#)) were used to evaluate the completeness of the genome. To further evaluate the completeness of the genome and the sequencing uniformity, we mapped the clean Illumina reads to the assembled genome with BWA software ([Li and Durbin, 2009](#)) and calculated the mapping rate and coverage with Samtools ([Danecek et al., 2021](#)).

Genome annotation

For the Yellow-breasted Bunting genome, we performed repeat annotation, gene coding region annotation, and non-coding RNA annotation. To identify genome repeats, we used homology alignment and *de*

de novo predict methods. The Repbase database and RepeatMasker software was used for homology prediction (Bao et al., 2015; Chen, 2004). The LTR_FINDER, RepeatScout, and RepeatModeler with default parameters were used to *de novo* predict the transposable elements (Chen, 2004; Xu and Wang, 2007). Tandem repeats were predicted by the Tandem Repeats Finder (Benson, 1999). The DNA-level repeat identification was performed in RepeatMasker (Chen, 2004). We used the program tRNAscan-SE to predict the tRNAs (Lowe and Eddy, 1997). The software with Rfam database was used to identify ncRNAs (Griffiths-Jones et al., 2005; Nawrocki and Eddy, 2013).

After identifying genome repeats and non-coding RNA, we combined *de novo* prediction, homology sequences-based prediction, and RNA sequencing-assisted prediction to annotate the gene structure in MAKER pipeline (Cantarel et al., 2008). The protein sequences of turkey (*Meleagris gallopavo*), chicken (*Gallus gallus*), and zebra finch (*Taeniopygia guttata*) were downloaded from the Ensembl database (<http://ftp.ensembl.org/pub/release-100/fastaf/>), and the protein sequences of Song Sparrow (*Melospiza melodia*) were downloaded from NCBI database (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/013/398/205/GCA_013398205.1_ASM1339820v1/). These protein sequences were aligned to the Yellow-breasted Bunting genome using TblastN v. 2.2.26 with the E-value less than 1×10^{-5} (Camacho et al., 2009). We then used GeneWise v. 2.4.1 (Birney et al., 2004) to predict gene structure based on the matching proteins and homologous genome sequences. The Augustus v. 3.2.3 with Geneid v. 1.4, Genescan v. 1.0, GlimmerHm v. 3.04 and SNAP (2013-11-29) were used to *de novo* predict gene structure (Blanco et al., 2007; Korf, 2004; Majoros et al., 2004; Stanke et al., 2006; Tiwari et al., 1997).

The raw Illumina transcriptome sequences were filtered in the Trimmomatic software with default parameters (Bolger et al., 2014). There were 13.46 Gb, 14.19 Gb, 14.22 Gb, and 11.85 Gb of clean Illumina transcriptome sequences of muscle, brain, heart and liver, respectively, used to correct the error of the corresponding Pacbio subreads. After error correction, the Pacbio transcriptome sequences from muscle, brain, heart and liver were each used to produce the consensus sequences in the isoseq pipeline (Gonzalez-Garay, 2016). TopHat v. 2.0.11 was used to align the RNA sequences from different tissues to the genome (Trapnell et al., 2009). The transcripts were then used as input for Cufflinks v. 2.2.1 to predict the gene structure (Trapnell et al., 2010). The consensus gene structures from the three methods were produced by EvidenceModeler v 1.1.1 with PASA (Program to Assemble Spliced Alignment) (Haas et al., 2008).

To assign the gene functions, we aligned the protein sequences to the Swiss-Prot using Blastp (E-value $\leq 1e-5$) (Bairoch and Apweiler, 2000; Camacho et al., 2009). The InterProScan70 v. 5.31 was used to annotate the motifs and domains using the public databases, including Pfam, ProDom, PRINTS, SMRT, PANTHER and PROSITE (Mulder and Apweiler, 2007). The Gene Ontology (GO) IDs of each gene was assigned based on the corresponding InterPro entry (Mulder and Apweiler, 2007). We also used the NR database to predict the gene function in the BLAST software (Camacho et al., 2009) and clustered the genes to KEGG pathway (Kanehisa and Goto, 2000).

Identifying sex chromosome-linked scaffolds

Because sex chromosomes have different recombination rates compared to that of autosomes, we needed to identify the sex chromosome-linked scaffolds. We used the chicken genome as reference because it has relatively continuous sex chromosomes among published avian genomes. We aligned the draft Yellow-breasted Bunting genome to the chicken genome (ftp://ftp.ensembl.org/pub/release-98/fastaf/gallus_gallus/dna/) with MUMmer v. 3.23 (Kurtz et al., 2004). After alignment, the longest consistent alignment was retained. The alignment results were filtered such that the minimum alignment length retained was 500, and minimum alignment identity was 50. After filtering, the position of the longest alignment delineated the location of each scaffold.

Read mapping and SNP calling

The raw Illumina sequences were trimmed using Trim Galore v. 0.6.6 (<https://github.com/FelixKrueger/TrimGalore/releases/tag/0.6.6>). We trimmed the adapter contents and removed five base pairs from both ends of each read. The FastQC v. 0.11.9 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used to evaluate the quality of the clean reads. After trimming, the clean reads were mapped to the draft Yellow-breasted Bunting genome with a Burrows–Wheeler alignment-maximal exact matches (BWA-MEM) algorithm in the Burrows-Wheeler Alignment Tool (BWA) v. 0.7.17 (Li and Durbin, 2009). After alignment, we used Samtools v. 1.13 (Li et al., 2009) to calculate the coverage for each site of the genome,

and then calculated the mean coverage of each sample. Samtools v. 1.13 was also used to estimate the mapping rate of each sample.

After mapping, the GATK v. 4.1.4.0 package (McKenna et al., 2010) was used to sort the reads and mark duplicates in the 'bam' file. The HaplotypeCaller, GenomicsDBImport, and GenotypeGVCFs tools in GATK v. 4.1.4.0 package were used to identify SNPs and indels according to the GATK best practices recommendations. vcftools v. 0.1.13 (Danecek et al., 2021) was used to extract autosome-linked SNPs and the quality metrics of each raw SNP. We used a density curve of each parameter as a reference to set filtering criteria as follows: QUAL < 30.0, AC < 2, FS > 60.0, QD < 2.0, MQ < 40.0, SOR > 3.0, MQRankSum < -4.0, ReadPosRankSum < -2.0, DP < 50, and DP > 361. Raw SNPs that fit any one of these filtering criteria were removed. The explanation of these parameters can be found from the GATK guide book (McKenna et al., 2010). The bi-allelic sites that on autosome were used for following analyses.

Family relationship inference

It is necessary to subsample non-closely related samples from the population genetic analysis. However, it is difficult to determine relationships from the pedigree of wild samples. Estimating kinship relationships among the samples based on genotype data can help confirm their relationships. We used KING v. 2.2.7 (Manichaikul et al., 2010) to estimate the relationship among samples. Following the KING manual, we set the threshold of the close relationship as kinship coefficient larger than 0.0442.

Estimating genetic diversity

To quantify genetic diversity of Yellow-breasted Buntings relative to other Passeriformes, we estimated the heterozygosity of the 10 Yellow-breasted Bunting and of 17 other Passeriformes birds. Heterozygosity refers to the number of heterozygous SNPs divided by the length of the genome without gaps. The heterozygosity of SNPs from each sample was extracted using vcftools v. 0.1.13 (Danecek et al., 2021). The 17 Passeriformes birds included Florida scrub jay (*Aphelocoma coerulescens*), Noisy scrubbird (*Atrichornis clamosus*), Chestnut-collared longspur (*Calcarius ornatus*), Seychelles magpie robin (*Copsychus sechellarum*), Ribbon-tailed drongo (*Dicrurus megarhynchus*), Painted honeyeater (*Grantiella picta*), Akiapola'au (*Hemignathus wilsoni*), Philippine fairy-bluebird (*Irena cyanogastra*), Kirtland's warbler (*Setophaga kirtlandii*), Loggerhead shrike (*Lanius ludovicianus*), Bali myna (*Leucopsar rothschildi*), Yellowhead (*Mohoua ochrocephala*), Velvet myiagra (*Myiagra hebetior*), Inaccessible island finch (*Nesospiza acunhae*), Stitchbird (*Notiomystis cincta*), White-necked rockfowl (*Picathartes gymnocephalus*), and Golden-crowned babbler (*Sterrhoptilus dennistouni*). The reference genomes and the fragment genome libraries of these 17 birds were downloaded from the National Center for Biotechnology Information search database (NCBI) (Feng et al., 2020). We used same pipeline as for SNPs of Yellow-breasted Bunting to locate and filter SNPs of these 17 birds.

Estimating inbreeding pattern

To infer the inbreeding status of Yellow-breasted Bunting, we calculated the linkage disequilibrium (LD) pattern and runs of homozygosity (ROH). Haploview v. 4.2 was used to calculate the correlation coefficient (r^2) of any pair of SNPs on autosomes (Barrett et al., 2004). The greater the R^2 value, the greater the degree of linkage disequilibrium. We set the maximum distance between any pair of SNPs to 500 kb and the minor allele frequency to 0.01. To plot the LD decay curve, we calculated the average R^2 in each 100 bp window. ROH always form due to inbreeding; the offspring inherit the identical-by-decent haplotype from parents (Ceballos et al., 2018). Thus, the number and length of ROH in the populations can help us infer the inbreeding pattern (Ceballos et al., 2018). In inferring the number and length of ROH, we first used vcftools v. 0.1.13 (Danecek et al., 2021) to convert vcf file to PLINK ped format for each scaffold and then used plink v. 1.90 (Howrigan et al., 2011) to infer ROH with the following parameters: ROH of at least 50 SNPs, at least 3 heterozygous SNPs, at least one SNP per 50 kb, a minimum of 300 kb length for an ROH, maximum distance between adjacent SNPs of 1000 kb per ROH, each scanning window having length of 50 SNPs, and 10 missing SNPs. Recombination can break the ROH. Thus, long and short ROH combined can reflect ancestral and recent inbreeding, respectively (Ceballos et al., 2018). We calculated the number and the total length of short ROH (<1Mb) and long ROH (>1.5 Mb) of each sample. We also calculated the percentage of the two kinds of ROH within full autosomes. Kolmogorov-Smirnov comparison was used to test whether the percentage has significant difference between two ROH groups.

Estimating potential inbreeding depression

To infer whether the Yellow-breasted Bunting genome contains signs of historical inbreeding depression, we used SnpEff v. 5.0c to identify the synonymous, missense, and lose of function (LOF) mutations in the 10 samples (Cingolani et al., 2012). After that, we calculated the average difference between the frequency of homozygous sites and heterozygous sites in 1000 random SNPs within the coding region (relative frequency of homozygous sites), according to the following formula:

$$\bar{X}_C = \left(\sum_{i \in C} f_{i,hom} - f_{i,heter} \right) / n$$

$f_{i,hom}$ is the frequency of homozygous mutations in each site. $f_{i,heter}$ is the frequency of heterozygous mutations in each site. C is the set of synonymous, missense, and LOF mutations. n is the number of different sets of mutations among the 1000 random SNPs. To test whether the relative frequency of homozygous sites is significantly different between missense and synonymous mutations, and between LOF and synonymous mutations, we performed 100 bootstrap replications of the coding sites. If the difference was not significant, this indicated the missense or LOF mutations were similar to neutral mutations; if the relative frequency of homozygous sites of missense or LOF is significantly larger than that of the synonymous mutations, then the Yellow-breasted Bunting likely has homozygous non-lethal mutations; if the relative frequency of homozygous sites of missense or LOF is significantly smaller than that of synonymous mutations, the Yellow-breasted Bunting has potential homozygous lethal mutations. We used a Kolmogorov-Smirnov comparison to test whether the difference was significant. To explore the potential influence of deleterious mutations on biological function, we performed KEGG pathway enrichment analysis for the genes that harbor LOF mutations in KOBAS (Bu et al., 2021).

Reconstructing yellow-breasted bunting demographic history

We aimed to detect the historical factors that caused Yellow-breasted Bunting's rapid decline. To do so, we used three methods to infer the species' effective population size (N_e) over time: Pairwise Sequentially Markovian Coalescent (PSMC) modeling (Li and Durbin, 2011), Stairway plots (Liu and Fu, 2015), and fastsimcoal (Excoffier et al., 2013; Excoffier and Foll, 2011). PSMC models use a hidden Markov model and local density of heterozygotes to infer the time of the most recent common ancestor between two alleles in a diploid genome sequence (Li and Durbin, 2011). We followed the instruction of PSMC to analysis the historical effective population size fluctuations, but we adjusted some parameters as follows: minimum accepted mapping quality for an alignment was 1, minimum read depth was 6, maximum read depth was 46, maximum number of iterations was 30, and the initial theta/rho ratio was 5. While PSMC models provide detailed information on the fluctuation of N_e , stairway plots can produce more accurate inference for recent N_e changes (Liu and Fu, 2015). Hence, we also used stairway plots to infer the historical N_e fluctuations using the SNP frequency spectrum (SFS) (Liu and Fu, 2015). The ANGSD v. 0.910 was used to produce the folded SFS from the bam files (Korneliussen et al., 2014). The folded SFS file was used as input file for stairway plot v. 2.1.1 (Liu and Fu, 2015).

To simulate the demographic history of Yellow-breasted Bunting in recent centuries, we used fastsimcoal (Excoffier and Foll, 2011), a complementary approach to PSMC and stairway plots. Fastsimcoal is an ancestral recombination graph (ARG) model based simulation program that can simulate complex demographic scenarios with SFS and then select the best fit model based on the Akaike Information Criterion (AIC) and the distribution of the likelihood of the models (Excoffier and Foll, 2011). Because linkage between any two sites will reduce the recombination rates, we filtered the linkage SNPs before running fastsimcoal. The plink software was used to filter the SNPs that their linkage disequilibrium coefficient larger than 0.2 in 1000 kb windows (Chang et al., 2015). After filtering, the 14,808 no-missing bi-allelic sites were used to produce folded SFS file in easySFS (<https://github.com/isaacovercast/easySFS>). The new SFS file was used for fastsimcoal26 (Excoffier and Foll, 2011).

We devised four demographic models based on the results from PSMC and the stairway plot. For each model, we performed 100 replicates of the optimum procedure, each with 100 expectations conditional maximization (ECM) cycles and 1,000,000 simulations per cycle. The replicate with the highest likelihood was considered the best model of each demographic model. To identify the best fit model, we re-ran each demographic model 100 times with the best parameter from the initial 100 replications. If the average likelihood of one model is significant different with other models, and its AIC value is less than other

models, the model was considered as the best fit demographic model. The Welch two sample t-test was used to determine whether the likelihood between two models is significant. The AIC value was calculated based on the likelihood value and the number of demographic parameters.

The four models included (the schematic diagram is shown in [Figure S4](#)): (a) Bottleneck and stable model (BS), in which Yellow-breasted Bunting has a constant population size after a bottleneck; (b) Bottleneck and decline model (BD), in which after experiencing a bottleneck, Yellow-breasted Bunting had a constant population and then declined continuously; (c) Bottleneck, which had stable ancestry and early decline model (BSED), the model is similar to model BD, but with time of population decline between 100 and 200 generations ago; (d) Bottleneck, stable ancestry and recent decline model (BSRD), similar to the BD model, but with population decline at 6–30 generations ago. The 200 generations ago of the Yellow-breasted Bunting was probably the beginning of the industrial revolution. In inferring the historical N_e fluctuations, we set the mutation rate as 3.3×10^{-9} per site per generation ([Zhang et al., 2014](#)) and assumed the generation time to be two years. After we identified the best fit model, we did 1000 bootstraps to estimate the demographic parameters and their confidence interval.