Genome Resequencing Reveals Congenital Causes of Embryo and Nestling Death in Crested Ibis (Nipponia nippon)

Chun-Zheng Fu, Xuan-Min Guang, Qiu-Hong Wan, and Sheng-Guo Fang * MOE Key Laboratory of Biosystems Homeostasis & Protection, State Conservation Centre for Gene Resources of Endangered Wildlife, College of Life Sciences, Zhejiang University, Hangzhou, P.R. China

*Corresponding author: E-mail: sgfanglab@zju.edu.cn.

Accepted: June 14, 2019

Data deposition: Sequencing data of the crested ibis genomes of this study have been deposited at NCBI BioProject (http://www.ncbi.nlm.nih.gov/bioproject) under the accession number PRJNA508426.

Abstract

The crested ibis (Nipponia nippon) is endangered worldwide. Although a series of conservation measures have markedly increased the population size and distribution area of these birds, the high mortality of embryos and nestlings considerably decreases the survival potential of this bird species. High-throughput sequencing technology was utilized to compare whole genomes between ten samples from dead crested ibises (including six dead embryos and four dead nestlings aged 0–45 days) and 32 samples from living birds. The results indicated that the dead samples all shared the genetic background of a specific ancestral subpopulation. Furthermore, the dead individuals were less genetically diverse and suffered higher degrees of inbreeding compared with these measures in live birds. Several candidate genes (KLHL3, SETDB2, TNNT2, PKP1, AK1, and EXOSC3) associated with detrimental diseases were identified in the genomic regions that differed between the alive and dead samples, which are likely responsible for the death of embryos and nestlings. In addition, in these regions, we also found several genes involved in the protein catabolic process (UBE4A and LONP1), lipid metabolism (ACOT1), glycan biosynthesis and metabolism (HYAL1 and HYAL4), and the immune system (JAM2) that are likely to promote the normal development of embryos and nestlings. The aberrant conditions of these genes and biological processes may contribute to the death of embryos and nestlings. Our data identify congenital factors underlying the death of embryos and nestlings at the whole genome level, which may be useful toward informing more effective conservation efforts for this bird species.

Key words: inbreeding depression, deleterious gene, population genetic structure.

Introduction

The crested ibis (Nipponia nippon), a species of bird, endangered worldwide. It is classified as order Pelecaniformes, family Threskiornithidae (Birdlife International 2018). This bird was historically distributed widely across East Asia with numerous flocks and individuals (Ding 2004). During the 19th to mid-20th century, the number of crested ibises suffered a rapid decrease because of climate change, habitat loss, overhunting, and agrochemical use (Ding 2004; Li et al. 2014). In 1981, after the remaining five wild crested ibises in Japan were captured and subjected to captive breeding, this bird was considered to be nearly extinct in the wild (Ding 2010). Although these birds had not been observed for over 17 years in China, a 3-year intensive investigation eventually identified two additional breeding pairs and three chicks in Yangxian County (Shaanxi Province, China) in 1981 (Ding 2010; Li et al. 2014). Subsequently, a series of conservation measures were immediately taken by the Chinese government, such as habitat and foraging site conservation, legislation for a ban on hunting, and the establishment of a special protection agency (Shi and Cao 2001; Wu et al. 2017). These measures effectively increased the number of individuals and distribution areas of the created ibis. Over the past 30-plus years, the IUCN Red List categorization of the crested ibis has been upgraded from critically endangered (CE) to endangered (EN), with the population and distribution areas of the crested ibis improving from only one remnant population with seven birds in one area to more than ten areas with nearly 2,500
birds in three countries (supplementary fig. S1, Supplementary Material online) (Ding 2010; Wajiki et al. 2015; Li et al. 2016; Birdlife International 2018).

Even though the population size and distribution of the crested ibis have been increasing steadily, this species still faces many survival challenges. The limited number of founder individuals renders it susceptible to the loss of genetic variation, fixation of deleterious mutations, and inbreeding depression, which can negatively impact viability, disease resistance, fertility, embryo survival, and number of litters (Soulé 1980; Lehmkuhl 1984; Falconer 1989; Lande 1995, 1998; Fischer et al. 2003). Accordingly, the mean death rates of the embryos and nestlings aged 0–42 days of the crested ibis were ~37% in the Shannxi (China) Crested Ibis Feeding and Breeding Center during 1995–2000 and the mean death probability of the yearlings (0–1 years old) of a reintroduced crested ibis population in Ningshan (China) during 2008–2015 was ~61.6% (Xi et al. 2001; Li et al. 2018). In comparison, the mean death rate of nestlings aged 0–30 days of 12 species of birds in captive populations was 22.73%, ranging from 5.48% to 46.46% (Boakes et al. 2007). The high mortality of yearlings has significantly diminished the survival probability of the crested ibis population; hence, effective protection of yearlings is considered critical to population survival (Li et al. 2018). Conversely, previous research related to the crested ibis focused on external factors such as viruses, bacteria, and parasites that influence survival, and incorporated microsatellite markers, major histocompatibility complex loci, and defensins that only provide limited genomic information for the development of particular measures for species conservation (Zhang et al. 2004, 2006, 2015; He et al. 2006; Xi et al. 2007; Chen et al. 2013, 2015; Urano et al. 2013; Hao et al. 2014; Lan 2014; Taniguchi et al. 2014). The rapid developments in high-throughput sequencing technology increased the research in genetic disorders and multifactorial etiologies (Hennekam and Biesecker 2012). Although numerous human genetic disorders have been studied via whole genome resequencing, animal hereditary diseases are rarely investigated using this technology (Dey et al. 2018; Estrada et al. 2018; Fahey et al. 2018; Tang et al. 2018; Vardaran et al. 2018; Wang et al. 2018). Notably, as the whole crested ibis genome was assembled by Li et al. (2014), it is now feasible to study the intrinsic factors that affect embryonic and nestling death of the crested ibis (Li et al. 2014).

The captive crested ibis population of Deqing County (Zhejiang, China), established in 2008, constitutes the first southern population of the crested ibis in China and has five breeding couples introduced from the captive crested ibis population of Zouzhi County (Shaanxi, China) as founders. In order to improve the genetic diversity of this population, ten and eight novel individuals were introduced from the captive population of Japan to this population in 2009 and 2011, respectively, and then four novel individuals from the captive population of Beijing were introduced into this population in 2014. By the end of the breeding season in 2017, after 9 years of assisted breeding with artificial incubation, the size of this captive population has gradually increased to 209 individuals. However, the death rate of embryos and nestlings aged 0–45 days of this captive population is high, ranging from 0% to 61.8% from 2009 to 2017 with a mean death rate of 38.8%.

To better understand potential genetic factors underlying the high embryo and nestling mortality in the Deqing population, in this study, we sequenced the whole genome of ten samples from six dead embryos and four dead nestlings aged 0–45 days, together with 32 samples from surviving crested ibises that were all incubated and fed under the same experimental conditions. These whole-genomes were then compared with explore genomic differences between the two groups. The findings are expected to substantially contribute to increasing the size and viability of the crested ibis population and will provide a better understanding at the whole genome scale of factors influencing the conservation of this endangered species.

**Materials and Methods**

**Sample Collection**

All samples used in this research were collected from the Xiazhu Lake Crested Ibis Breeding Base in Deqing County (Zhejiang, China) from 2009 to 2017. During the crested ibis breeding season, we collected eggs every morning for artificial incubation and recorded the exact time that the eggs started incubation. The hatching conditions and brooding conditions used were the same as those reported in a previous study (Xi et al. 2001). After 28 days, if the chick had not hatched, we collected a tissue sample if the nestling had died from unknown reasons within 45 days posthatching, and stored it at −20°C. We also collected a tissue sample if the nestling had died at an age of 45 days, and stored it at −20°C. The peripheral venous blood samples from living birds were drawn via wing venipuncture when birds reached an age of 45 days, and stored at −20°C. Samples were collected under the supervision of the Animal Ethics Committee of Zhejiang University and permission was acquired from the Xiazhu Lake Crested Ibis Breeding Base in Deqing County (Zhejiang, China). We tried to collect as many samples of individuals that died at an early age as possible; however, owing to the problem of sample collection and genomic degradation, ten samples including six dead embryos and four dead nestlings aged 0–45 days (DEN: Dead embryos and nestlings) were amenable for whole genome resequencing. Of those, five dead embryos and two dead nestlings were collected in 2017, one dead nestling was collected in 2016, one dead embryo was collected in 2014, and one dead nestling was collected in 2013. About 32 samples from living crested
ibises (AL: Alive individuals) were accessible for whole genome resequencing, including 12 birds that were born in 2017, 6 birds that were born in 2016, 2 birds that were born in 2009, and 5 founder breeding couples and 2 second spouses (the first spouses of 2 founder breeding couples died). Pedigree information on this captive population is missing (supplementary table S1, Supplementary Material online).

DNA Extraction and Sequencing

Genomic DNA was extracted using the DNeasy Tissue & Blood Kit (Generay, China). DNA degradation and contamination were detected by 1% agarose gel electrophoresis and its concentration was measured using the Qubit DNA Assay Kit with a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA). Sequencing libraries were constructed using the TrueSeq Nano DNA HT Sample preparation Kit (Illumina, San Diego, CA) following manufacturer’s instructions. The generated libraries, with insert sizes of ~350 bp, were sequenced using the Illumina HiSeq X Ten system to generate 150-bp paired-end reads at the Beijing Novogene Co., Ltd (Beijing, China). Approximately 1,386.27 Gb of raw data were generated in total. The Q30 ranged from 88.50% to 94.47% and GC content comprised 43.35–47.49% for all samples (supplementary table S1, Supplementary Material online).

Sequence Quality Checking and Read Mapping

We removed low quality paired reads based on the following criteria: 1) ≥10% unidentified nucleotides, 2) >50% bases having phred quality <5, 3) >10 nt aligned to the adapter, allowing ≤10% mismatches, and 4) putative PCR duplicates (i.e., read1 and read2 of two paired-end reads were completely identical). These low-quality reads were mainly caused by base-calling duplicates and adapter contamination. The cleaned data accounted for ~99.26% of the raw data and the amount of high-quality sequence data of all samples ranged from 30.32 to 47.41 Gb (supplementary table S1, Supplementary Material online). We mapped the remaining high-quality reads to the N. nippon reference genome (GCF_000708225.1) using BWA-MEM (0.7.10-r789) with default parameters (Li and Durbin 2009; Li et al. 2014). We used SAMtools1.8 (Li et al. 2009), Picard2.18.7 (http://broadinstitute.github.io/picard/), and GATK4.0 (DePristo et al. 2011) (https://software.broadinstitute.org/gatk/) successively for sorting and merging the preliminary alignment results, assigning read group information, and performing local realignment, respectively. The cleaned reads of all samples appropriately mapped to the crested ibis reference genome. The effective sequencing depth exceeded 22.70× (22.70–33.98×) and the genome coverage surpassed 95.24% (95.24–96.92%) for all samples (supplementary table S1, Supplementary Material online).

Variant Calling

We used Picard2.18.7 and GATK4.0 (DePristo et al. 2011) with the HaplotypeCaller tool and followed the standard pipeline to perform SNP calling. GATK4.0 was used to identify insertions and deletions (InDels) with a fragment size <50 bp and it was also used to perform variant quality score recalibration, which utilized the top10 of the best-quality InDels as the training set to model InDels. We used the linear regression function in R3.5.1 to test for assessing differences in the number of SNPs and InDels between the AL and DEN samples, because regression has been shown to be robust to violations of the normality assumption (Knief and Forstmeier 2018).

Runs of Homozygosity

To reduce the impact of SNPs on the sex chromosome, we used LASTZ1.02 (with parameters T = 2, C = 2, H = 2,000, Y = 3,400, L = 6,000, K = 2,200 –identity = 90%) to align the crested ibis genome with the chicken ZW chromosome sequence and removed regions >100 bp exhibiting synteny with the ZW sequence of chicken (Harris 2007). The BCFtools1.8/RoH approach, which is based on a hidden Markov model, was utilized to detect runs of homozygosity (ROH) on the autosomal genome of the 42 samples. We used this method with default parameters: bcf tools roh -s sample file.vcf (Narasimhan et al. 2016). To estimate the homozygosity of each sample, we counted the number and calculated the total length of ROHs. The genomic inbreeding coefficient, Froh, was used to measure individual inbreeding, and was defined as the percentage of the autosomal genome in ROH as follows:

\[ \text{Froh} = \frac{\sum \text{Lroh}}{\text{Lauto}} \]

where \( \sum \text{Lroh} \) is the total length of all ROHs on the autosomes of a sample and \( \text{Lauto} \) is the total length of the autosomal genome (McQuillan et al. 2008; Ceballos et al. 2018). We used the linear regression function in R3.5.1 to test for differences in the number of ROHs, the total length of ROHs, and Froh between AL and DEN.

Population Genetic Structure

The program ADMIXTURE1.3.0 (Alexander et al. 2009) was implemented to infer the population genetic structure. The number of clusters was varied from \( K = 2 \) to \( K = 5 \) and the lowest cross-validation error was used to determine the optimal \( K \) value. The neighbor-joining method was performed in VCF-kit (http://vcf-kit.readthedocs.io/en/latest/) to reconstruct a phylogenetic tree of the 42 samples, which was visualized using FigTree v1.4.3. EIGENSOFT v6.0.1 was utilized to perform the principal component analysis and the Tracey–Widom test was conducted to ascertain the significance of eigenvectors (Patterson et al. 2006).
Identification of Genomic Regions with Differences between AL and DEN

In order to identify genomic regions with differences between the AL and DEN samples, we calculated the difference in sequence diversity (as the ratio $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}})$), the population differentiation statistic ($F_{ST}$), and Tajima’s $D$ values between AL and DEN samples in a genome-wide sliding window approach with a window size of 30 kb and a step size of 15 kb. The windows with simultaneously the lowest 5% $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}})$ ratios (i.e., $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}}) < -0.4645$ or $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}}) > 0.60822$) and the highest 5% $F_{ST}$ values (i.e., $F_{ST} > 0.134475$) were selected as representative regions (part A and part B) with differences between AL and DEN. These regions were further confirmed by testing whether there was a significant difference in values of Tajima’s $D$ between AL and DEN (Li et al. 2013; Qi et al. 2013; Qiu et al. 2015). We used the linear regression function in R 3.5.1 to test for differences in values of Tajima’s $D$ of the representative regions with difference between AL and DEN. KOBAS 3.0 was used to perform functional enrichment analysis with the candidate genes located in the genomic outlier regions using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations (Wu et al. 2006; Xie et al. 2011).

**Results**

**SNPs and InDels**

Using SAMtools, 1.29 million SNPs were identified, with the majority (811,567; 62.70%) being located in intergenic regions, and 20,691 located in exonic regions (supplementary table S2, Supplementary Material online). The average number of SNPs across all individuals was 766,748. The number of SNPs per individual was significantly lower in the DEN samples (estimate $\pm$ SE: 748,737 $\pm$ 8,991) than in the AL samples (estimate $\pm$ SE: 772,377 $\pm$ 5,026) ($P = 0.03$, fig. 1A and supplementary table S3, Supplementary Material online). The sequence diversity ($\pi$) of all samples was 4.65 $\times 10^{-4}$, and the sequence diversity of DEN (4.60 $\times 10^{-4}$) was also lower than that of AL (4.64 $\times 10^{-4}$). We obtained a total of 114,794 insertions and 102,935 deletions among all samples. The average number of InDels across all individuals was 123,085, and the number of InDels per individual in the DEN samples (estimate $\pm$ SE: 121,046 $\pm$ 1,103) was significantly lower than that in the AL samples (estimate $\pm$ SE: 123,722 $\pm$ 617) ($P = 0.04$, fig. 1B and supplementary table S3, Supplementary Material online).

**Runs of Homozygosity**

The mean number of ROHs across all individuals was 2,979 and the number of ROHs per individual in the DEN samples (estimate $\pm$ SE: 3,080 $\pm$ 94) was significantly higher than that in the AL samples (estimate $\pm$ SE: 2,711 $\pm$ 52) ($P = 0.0014$, fig. 2A and supplementary table S4, Supplementary Material online). The mean value of the total length of ROHs across all individuals was 219,219,330 bp, and the total length of ROHs per individual in the DEN samples (estimate $\pm$ SE: 258,397,906 $\pm$ 10,482,167 bp) was significantly higher than that in the AL samples (estimate $\pm$ SE: 206,976,024 $\pm$ 5,859,710 bp) ($P = 0.0001$, fig. 2B and supplementary table S4, Supplementary Material online). The average Froh was 0.20 for all samples. The Froh of the DEN group (estimate $\pm$ SE: 0.231 $\pm$ 0.009) was significantly higher than that of the AL group (estimate $\pm$ SE: 0.184 $\pm$ 0.005) ($P = 0.0001$, fig. 3 and supplementary table S4, Supplementary Material online).

**Population Genetic Structure of the Captive Crested Ibis Population of Deqing**

We inferred that the captive crested ibis population of Deqing was derived from three ancestral subpopulations by running the ADMIXTURE 1.3 program from $K = 2$ to $K = 5$ and calculating the cross-validation error (fig. 4A and B). We also found that all DEN samples possessed the genetic background from ancestral subpopulation 1 (fig. 4C). We grouped the 42 samples on the basis of the contribution percentage of the three ancestral subpopulations (fig. 5A and supplementary table S5, Supplementary Material online). The genomic components of three samples, all of which were DEN, were completely derived from ancestral subpopulation 1, whereas eight and two samples (all AL) were entirely derived from ancestral subpopulation 2 and 3, respectively. The results of neighbor joining tree and principal component analysis were consistent with the results of population genetic structure analysis with $K = 3$ (fig. 5B and Q).

**Identification of Genomic Regions That Differ between the AL and DEN Samples**

The windows with simultaneously the lowest 5% $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}})$ ratios and the highest 5% $F_{ST}$ values, and the windows with simultaneously the lowest 5% $\ln(\pi_{\text{AL}}/\pi_{\text{DEN}})$ ratios and the highest 5% $F_{ST}$ values were selected as representative regions (part A and part B) with differences between AL and DEN, respectively (fig. 6 and supplementary fig. S2A and B, Supplementary Material online). These regions were further confirmed by significantly different values in Tajima’s $D$ between the AL and DEN ($P < 2.2e - 16$ and $P < 2.2e - 16$ of part A and part B, respectively). We identified 858 (15.78 Mb in total) and 269 windows (4.65 Mb in total) in the part A and part B of the representative regions with differences between AL and DEN, respectively, with these windows containing 107 and 31 nonsynonymous SNPs in 117 and 57 protein-coding genes, respectively. Of those, 56 and 21 genes contained the nonsynonymous SNPs, respectively. To obtain further information regarding the functional roles of these genes in the crested ibis genome, we used KOBAS to perform GO and KEGG pathway analysis.
A total of 110 GO terms and 4 KEGG pathways were significantly enriched in part A of the representative regions at $P < 0.05$ (supplementary tables S6 and S7, Supplementary Material online). Among these, five significantly enriched GO terms ($P < 0.05$, GO: 0030163; GO: 0006511; GO: 0019941; GO: 0051603; GO: 0044257) included four genes ($UBE4A$, $ANAPC2$, $USP40$, and $LONP1$) that were related to the protein catabolic process. In addition, $ADAM11$ (with one nonsynonymous SNP) and $JAM2$, which fall within the same GO term of binding (GO: 0005488), may play a role as tumor suppressors. The four significantly enriched pathways are associated with lipid metabolism, glycan biosynthesis and metabolism, and the immune system, respectively. Specifically, $ACOT1$ (with three nonsynonymous SNPs) pertains to the pathway of lipid metabolism; $HYAL1$ (with one nonsynonymous SNP) and $HYAL4$, which can catabolize hyaluronic acid, belong to the pathway of glycan biosynthesis and metabolism, and $JAM2$, in the immune system pathway, plays a crucial role in the migration of natural killer cells.

Moreover, 58 GO terms and 2 KEGG pathways were significantly enriched in part B of the representative regions at $P < 0.05$ (supplementary tables S8 and S9, Supplementary Material online). Among the GO terms, four genes were associated with diseases that likely result in lethality: $KLHL3$, $SETDB2$ (with one nonsynonymous SNP), $TNNT2$, and $PKP1$, all within the GO term “intracellular” (GO: 0005622). In addition, $AK1$, which pertains to the pathway of thiamine metabolism and $EXOSC3$, which belongs to the pathway of RNA degradation, are also associated with severe, potential lethal diseases.

**Discussion**

**Population Genetic Structure of the Crested Ibis**

Our findings suggest that the captive crested ibis population of Deqing derived from three ancestral subpopulations. This population was founded by five breeding couples that were...
introduced from Zhouzhi County in 2008, merged with ten and eight novel individuals from the captive population of Japan in 2009 and 2011, respectively, then acquired four individuals from a captive population of Beijing in 2014 to improve its genetic diversity. This population thus represents the entire global population of the crested ibis to some extent. Thus, all living individuals of the global crested ibis population may be descendants of the three ancestral subpopulations. However, individuals of the Sado captive crested ibis population were reported to derive from only two ancestral subpopulations (Urano et al. 2013). This outcome may have been reached because the Sado captive population was founded by only five individuals from China and the results were obtained using microsatellite markers with limited polymorphism. In addition, we also found that all DEN individuals possessed the genetic background of a specific subpopulation, with the genomic components of three DEN samples being completely inherited from this ancestral subpopulation. This result suggested that the genome of this ancestral subpopulation may contain some recessive deleterious mutations, with homozygosity of these mutations potentially constituting one of the reasons causing the death of affected embryos and nestlings. In subsequent breeding projects, we should try to introduce more novel individuals from other populations into the captive crested ibis population of Deqing to increase its genetic diversity. In addition, we should try to collect accurate pedigree information by careful observation, and perform paternity testing on newborn birds. Based on the complete pedigree data, we would thus be able to minimize inbreeding with appropriate human intervention and reduce the risks of species extinction.

Inbreeding Depression as a Cause of Embryo and Nestling Death

Mating among individuals or organisms that have close affinities with each other through common ancestry is termed inbreeding, which often leads to inbreeding depression, that is, a reduction in fitness (Falconer 1989; Charlesworth and Willis 2009). Research from birds and mammals in the wild revealed that inbreeding depression has a marked impact on birth weight, survival, proliferation, and disease resistance (Keller and Waller 2002). Furthermore, by evaluating inbred wild species under natural conditions, it was discovered that wild animals commonly displayed moderate to high degrees of inbreeding depression in fitness traits (Cmokrak and Roff 1999). Moreover, the disease prevalence of human gout, depression, peptic ulcer, and cancer increased with higher inbreeding coefficients, and 23–48% of disease morbidity can be ascribed to inbreeding (Rudan et al. 2003).

Researchers have found that the crested ibis shows a high inbreeding coefficient and low genetic diversity (Zhang et al. 2004; Urano et al. 2013; Li et al. 2014; Wajiki et al. 2015). The crested ibis experienced a severe bottleneck and all the birds of this species in the world at present are descendants of the two breeding couples that were found in Yangxian County (Shaanxi Province, China) in 1981 (Ding 2010). The limited number of founders and small population size thus account for the present state of high inbreeding coefficient and low genetic diversity of this species, which render it more susceptible to inbreeding depression. Even though the size of the captive crested ibis population of Deqing has increased gradually to reach 209 individuals, it is still far lower than the number suggested by Franklin (1980) that is necessary to maintain sufficient genetic variation within the population over a long period (Franklin 1980). Small populations are especially sensitive to the effects of deleterious mutations and inbreeding depression (Fischer et al. 2003). Our study identified that the number of SNPs and InDels in the DEN samples were significantly lower than in the AL samples \( (P < 0.05) \), and the sequence diversity of DEN was also lower than that of AL. In addition, the number and total length of ROHs, along with Froh values in DEN samples were significantly higher than those in AL. These results all suggest that in comparison with AL, the DEN samples contained less genetic diversity and suffered higher degrees of inbreeding. Therefore, inbreeding depression may be the reason for the death of crested ibis embryos and nestlings.

Even though the effect of inbreeding depression appears to vary among species, meta-analyses of 119 zoo populations indicated that inbreeding depression has a significant impact on neonatal survival (Boakes et al. 2007). With inbreeding, the
neonatal mortality of the Speke’s gazelle (*Gazella spekei*) increased (Boakes et al. 2007). Inbreeding populations of the domestic Pacific blue shrimp *Penaeus (Litopenaeus) stylirostris* also exhibit higher mortality than interbreeding populations (Goyard et al. 2008). In addition, inbreeding has adverse effects on fitness-related traits such as body mass and horn growth of the Alpine ibex (*Capra ibex*) (Brambilla et al. 2014). Moreover, inbred individuals may have genetic diseases, such as the chondrodystrophy observed in the California condor (*Gymnogyps californianus*), and may be more susceptible to infectious diseases, as avian malaria for the crow (*Corvus brachyrhynchos*) (Ralls and Ballou 2004; Townsend et al. 2018). These results all suggest that it is important to minimize inbreeding and protect genetic diversity within populations for preserving the population’s fitness, especially in small populations with limited founder individuals.

**Congenital Factors That Cause the Death of Embryos and Nestling**

Among the significantly enriched GO terms and KEGG pathways in part A of the representative regions with difference, five significantly enriched GO terms were related to protein catabolic process. This process is extremely important for embryonic survival and development as the incubated eggs lack carbohydrate, and most of the glucose that sustains embryo vigor and survival is derived from protein metabolism by glycogenesis and gluconeogenesis (Christensen et al. 2001; De Oliveira et al. 2008). The two genes *UBE4A* and *LONP1* play particular roles in the protein catabolic process. *UBE4A* participates in diverse biochemical processes that are necessary for growth and differentiation (Contino et al. 2004). *LONP1* can degrade proteins that are damaged by oxidation for metabolic purposes (Bezawork-Geleta et al. 2015). In addition, *ADAM11* and *JAM2*, which belong to the GO term of binding, may act as tumor suppressors of breast cancer and colorectal cancer, respectively (Katagiri et al. 1995; Zhao et al. 2016).

The significantly enriched pathways are associated with lipid metabolism, glycan biosynthesis and metabolism, and the immune system in part A of the representative regions with difference. In the midincubation stage until 2 or 3 days prior to piping, lipid metabolism comprises the most active metabolism in embryos. In particular, the biosynthesis of fatty

---

**Fig. 4.** — Population genetic structure of the captive crested ibis population of Deqing. (A) Population structure plots with $K = 2$ to $K = 5$. (B) Broken line graph of cross-validation error. (C) Samples grouped according to dead and alive state while $K = 3$. 

---

**Genome Resequencing Reveals Congenital Causes of Embryo and Nestling Death in Crested Ibis** 

*Genome Biol. Evol.* 11(8):2125–2135 doi:10.1093/gbe/evz149 Advance Access publication July 12, 2019 2131
Acids and β-oxidation are very active during this embryonic developmental stage (De Oliveira et al. 2008). Specifically, ACOT1, which pertains to the pathway of lipid metabolism, can sustain free fatty acid homeostasis (Hunt et al. 2006). During the last stage of incubation, the majority of threats that bird embryos encounter are related to energy metabolism, and the processes of glycogen synthesis and degradation are crucial for their survival (De Oliveira et al. 2008). Embryos metabolize glycogen to maintain the activity of muscle and generate heat. If the stored glycogen becomes depleted, this will affect embryo vigor and survival (Christensen et al. 1999; De Oliveira et al. 2008). HYAL1 and HYAL4 can catabolize hyaluronic acid, a kind of glycosaminoglycan (Laurent and Fraser 1992; Gakunga et al. 1997; Csoka et al. 1999).

In addition, in the last-incubation stage, embryos are immunocompetent and can generate innate and adaptive immune responses to causative agents (Ribatti 2008; Schilling et al. 2018). JAM2 belongs to the immunoglobulin superfamily and plays a crucial role in the migration of natural killer cells and cytotoxic T cells (Johnson-Léger et al. 2002; Ueki et al. 2008). These genes, GO terms, and KEGG pathways are likely to promote normal embryo development and healthy nestling growth. Conversely, the aberrant conditions of these genes, GO terms, and pathways in DEN samples may contribute to the eventual death of embryos and nestlings.

Among the genes associated with the significantly enriched GO terms in part B of the representative regions with difference, four are related to detrimental diseases. Mutations in KLHL3 can cause pseudohypoaldosteronism II.
cause pontocerebellar hypoplasia and spastic paraplegia, pathway of RNA degradation. Mutation in this gene can
pertains to the pathway of thiamine metabolism, and its mu-
result of thiamine deficiency, hatchlings of chickens may die
might carry genetic defects of thiamine metabolism (Jeffrey
fant death syndrome (SIDS) exhibit higher mortality from thi-
energy metabolism, and interorganellar signal transduction
significantly enriched KEGG pathways in part B of the repre-
Genome Resequencing Reveals Congenital Causes of Embryo and Nestling Death in Crested Ibis
Fig. 6.—Genomic regions with difference between the alive (AL) and
dead embryo and nestling (DEN) samples of crested ibis. Distribution of
the population statistic (Fst) values and the In(\text{AL}/\text{DEN}) ratios of 30-kb win-
dows with 15-kb steps. Blue and red dots highlight part A and part B of the
representative regions with difference between AL and DEN, respectively.
(OMIM: 145260) (Louis-Dit-Picard et al. 2012; Wu and Peng
2013). SETDB2 plays oncogenic roles in gastric cancer
(Nishikawaji et al. 2016). TNNT2 encodes a part of the troponin
complex; this protein has been reported as an effective prog-
nostic factor for all-cause mortality of patients with acute coro-
nary syndrome (Arnadottir et al. 2017). Mutations in PKP1 can
cause ectodermal dysplasia-skin fragility syndrome, an autosomal
recessive genodermatosis (Hernandez-Martin et al. 2013).

The thiamine metabolism pathway represents one of the
significantly enriched KEGG pathways in part B of the repre-
sentative regions with difference. Thiamine is a vital com-
pound for multicellular organisms that participates in several
metabolic processes including immune system activation, en-
ergy metabolism, and interorganellar signal transduction
(Manzetti et al. 2014). Families with “near-miss” sudden in-
fant death syndrome (SIDS) exhibit higher mortality from thi-
amine deficiency, with thiamine deficient infants being more
likely to die because of SIDS. Such “near-miss” SIDS families
might carry genetic defects of thiamine metabolism (Jeffrey
et al. 1985). In addition, rats deficient in thiamine exhibit neu-
ropathological damage of brainstem structures (Butterworth
et al. 1985). Moreover, researchers have shown that as a
result of thiamine deficiency, hatchlings of chickens may die
from polyneuritis (Poin et al. 1962; Charles et al. 1972). AK1
pertains to the pathway of thiamine metabolism, and its mu-
tation can cause chronic hemolytic anemia (Bianchi 1999).
In addition, EXOSC3 is included in the significantly enriched
pathway of RNA degradation. Mutation in this gene can
cause pontocerebellar hypoplasia and spastic paraplegia,
which is a serious autosomal recessive neurodegenerative
disease with prenatal onset (Wan et al. 2012; Eggens et al.
2014; Halevy et al. 2014). This disease usually causes patient
death within 1 year after birth (Schwabova et al. 2013). Thus,
these genes, GO terms, and KEGG pathways may comprise
causal factors in embryos and nestlings’ death.

Human genetic disorders like the Hirschsprung disease,
spontaneous coronary artery dissection, late-onset
Alzheimer’s disease, and neuromyelitis optica have been stud-
ied through whole genome resequencing (Estrada et al. 2018;
Fahey et al. 2018; Tang et al. 2018; Vardarajan et al. 2018). The
animal hereditary disease ascites syndrome for the chicken was
also studied through this technology (Dey et al. 2018). Whole
genome resequencing technology might be utilized to study
other threatened bird species with high mortality of embryos
and nestlings, such as the Congo peafowl (Afropavo congensis)
and Guam Micronesian kingfisher (Halcyon cinnamomina cin-
namomina), as the mortality of nestlings aged 0–30 days of
these two birds exceeds 40% (Boakes et al. 2007). Knowing
the congenital factors that cause the nestling death of these
species can help acquire more knowledge of detrimental dis-
eases and better protect these species.

In summary, high-throughput sequencing technology was
utilized to resequence the whole crested ibis genome of AL
and DEN individuals in this study. The DEN group exhibited
a genetic background that was probably enriched for deleteri-
ous genes (KLHL3, SETDB2, TNNT2, PKP1, AK1, and EXOSC3)
and it also showed higher extent of inbreeding, ultimately
resulting in the death of embryos and nestlings. Moreover,
in the genomic regions that differ between the alive and dead
duplicates, we also identified several genes involved in the pro-
tein catabolic process (UBE4A and LONP1), lipid metabolism
(ACOT1), glycan biosynthesis and metabolism (HYAL1 and
HYAL4), and the immune system (JAM2) that are likely to
promote the normal development of embryos and nestlings.
The aberrant conditions of these genes and biological pro-
cesses may also contribute to the eventual death of embryos
and nestlings. Even though additional experiments are re-
quired to verify our results, our study provides analyses at
the whole genome level of the congenital factors underlying
the death of crested ibis embryos and nestlings. These data
establish a foundation for more effectively managing the cap-
tive breeding populations of the crested ibis, with the goal
toward ensuring the survival of this endangered species.

Supplementary Material
Supplementary data are available at Genome Biology and
Evolution online.

Acknowledgments
This work was supported by the National Key Program
(2016YFC0503200) from the Ministry of Science and
Technology of China, a special grant from the State Forestry
Central Universities of China.

Literature Cited
Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19(9):1655–1664.
Arnadottir A, Klein CF, Æresen K. 2017. Head-to-head comparison of cardiotroponin T and troponin I in patients without acute coronary syndrome: a systematic review. Biomarkers 22:701–708.
Bezawork-Geleta A, Brodie EJ, Dougan DA, Truscott KN. 2015. LON is the master protease that protects against protein aggregation in human mitochondria through direct degradation of misfolded proteins. Sci Rep. 5:17397.

Bianchi P. 1999. A case of complete adenylate kinase deficiency due to a nonsense mutation in AK-7 gene (Arg 107 -> Stop, CGA -> TGA) associated with chronic haemolytic anaemia. Br J Haematol. 105(1):75–79.
Birdlife International [Internet]. 2018. Species factsheet: Nipponia nippon. [cited 2018 07/03/2018]. Available from: http://www.birdlife.org.
Boakes EH, Wang J, Amos W. 2007. An investigation of inbreeding depression and purging in captive pedigree populations. Heredity 98(3):172–182.
Brambilla A, Biebach I, Bassano B, Bogliani G, von Hardenberg A. 2014. Direct and indirect causal effects of heterozygosity on fitness-related traits in Alpine ibex. Proc Biol Sci. 282(1798):20141873–20141873.
Butterworth RF, Giguere JF, Besnard AM. 1985. Activities of thiamine-dependent enzymes in two experimental models of thiamine-deficiency encephalopathy: 1. The pyruvate dehydrogenase complex. Neurochem Res. 10(10):1417–1428.
Ceballos FC, Joshi PK, Clark DW, Ramays M, Wilson JF. 2018. Runs of homozygosity: windows into population history and trait architecture. Nat Rev Genet. 19(4):220–234.

Charles OW, Roland DA, Edwards HM. 1972. Thiamine deficiency identification and treatment in commercial turkeys and Coturnix quail. Poult Sci. 51(2):419–423.
Charlesworth D, Willis JH. 2009. The genetics of inbreeding depression. Nat Rev Genet. 10(11):783–796.
Chen LC, et al. 2015. Genomic organization of the crested ibis MHC provides new insight into ancestral avian MHC structure. Sci Rep. 5:7963.
Chen SL, et al. 2013. Phylogenetic and pathogenic analyses of two virulent Newcastle disease viruses isolated from Crested ibis (Nipponia nippon) in China. Virus Genes. 46(3):447–453.

Christer sen VL, Donaldson WE, Nestor KE. 1999. Length of the plateau and pipping stages of incubation affects the physiology and survival of turkeys. Br Poult Sci. 40(2):297–303.

Christer sen VL, Wineland MJ, Fasenko GM, Donaldson WE. 2001. Egg storage effects on plasma glucose and supply and demand tissue glycogen concentrations of broiler embryos. Poult Sci. 80(12):1729–1735.

Contino G, et al. 2004. Expression analysis of the gene encoding for the U-box-type ubiquitin ligaseUBE4A in human tissues. Gene 328:69–74.

Cronk r P, Roff DA. 1999. Inbreeding depression in the wild. Heredity 83(3):260–270.

Csoa AB, Scherer SW, Stern R. 1999. Expression analysis of six paralogous human hyaluronidase genes clustered on chromosomes 3p21 and 7q31. Genomics 60:356–361.

De Oliver he JE, Un C, Fer kert PR. 2008. Important metabolic pathways in poultry embryos prior to hatch. Worlds Poult Sci J. 64(4):488–499.

DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43(5):491–498.

Dey S, et al. 2018. Whole genome resequencing identifies the CPQ gene as a determinant of ascites syndrome in broilers. PLoS One 13(1):e0189544.

Ding C. 2004. Research on the crested ibis. Shanghai: Shanghai Scientific and Technological Publishing House.

Ding C. 2010. Crested ibis. Chin Birds. 1(2):156–162.

Eggens VRC, et al. 2014. EXOSC3 mutations in pontocerebellar hypoplasia type 1: novel mutations and genotype-phenotype correlations. Orphanet J Rare Dis. 9(1):23.

Estrada K, et al. 2018. A whole-genome sequence study identifies genetic risk factors for neuromyelitis optica. Nat Commun. 9(1):1929.

Fahey J, Ramalison M, White A. 2018. Identifying genetic causes of spontaneous coronary artery dissection by whole genome sequencing in related individuals. J Am Coll Cardiol. 71:123–123.

Falconer DS. 1989. Introduction to quantitative genetics. Burnt Mill, Harlow, Essex, England. New York: Longman Wiley.

Fischer M, Hock M, Paschke M. 2003. Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. Conserv Genet. 4(3):325–336.

Franklin IR. 1980. Conservation biology— an evolutionary-ecological perspective. Sunderland: Massachusetts Sinauer Associates.

Gakunga P, et al. 1997. Hyaluronan is a prerequisite for ductal branching morphogenesis. Development 124(20):3987–3997.

Goyard E, et al. 2008. Cross breeding of different domesticated lines as a simple way for genetic improvement in small aquaculture industries: heterosis and inbreeding effects on growth and survival rates of the Pacific blue shrimp Penaeus (Litopenaeus) stylirostris. Aquaculture 278(1–4):43–50.

Halevy A, et al. 2014. Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. J Neurol. 261(11):2165–2169.

Hao H, et al. 2014. Genomic characterisation of two virulent Newcastle disease viruses isolated from crested ibis (Nipponia nippon) in China. Gene 553(2):84–89.

Harris RS. 2007. Improved pairwise alignment of genomic DNA. PhD Thesis: The Pennsylvania State University.

He LP, Wan QH, Fang SG, Xi YM. 2006. Development of novel microsatellite loci and assessment of genetic diversity in the endangered Crested ibis, Nipponia nippon. Conserv Genet. 7(1):157–160.

Hennekam RCM, Bieseker LG. 2012. Next-generation sequencing demands next-generation phenotyping. Hum Mutat. 33(5):884–886.

Hernandez-Martin A, et al. 2014. Novel EXOSC3 mutation causes complicated hereditary spastic paraplegia. Hum Mutat. 35(3):278–286.

Hunt MC, Rautanen A, Westin MAK, Svensson LT, Alexson SEH. 2006. Analysis of the mouse and human acyl-CoA thioesterase (ACOT) gene clusters shows that convergent, functional evolution results in a reduced number of human peroxisomal ACOTs. FASEB J. 20(11):1855–1864.

Jeffrey HE, Mc Cleary BV, Hensley WJ, Read DJC. 1985. Thiamine-deficiency—a neglected problem of infants and mothers—possible relationships to sudden infant death syndrome. Aust N Z J Obstet Gynaecol. 25(3):198–202.

Johnson-Leger CA, Aurand-Lions M, Beltramini N, Fasel N, Imhof BA. 2002. Junctional adhesion molecule-2 (JAM-2) promotes lymphocyte transendothelial migration. Blood 100(7):2479–2486.

Katagiri T, Harada Y, Emi M, Nakamura Y. 1995. Human metalloprotease/alpha-1-antiplasmin (Arg 107 - > Stop, CGA - > TGA) gene (Arg 107 - > Stop, CGA - > TGA) associated with chronic haemolytic anaemia. Br J Haematol. 105(1):75–79.

Knief U, Forstmeier W. 2018. Violating the normality assumption may be the lesser of two evils. bioRxiv. 49893.
Genome Resequencing Reveals Congenital Causes of Embryo and Nestling Death in Crested Ibis

Li MZ, et al. 2013. Genomic analyses identify distinct patterns of selection

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760.

Laurent TC, Fraser JR. 1992. Hyaluronan. FASEB J. 6(7):2397–2404.

Lehmkuhl JF. 1984. Determining size and dispersion of minimum viable populations for land management planning and species conservation. Environ Manage. 8(2):167–176.

Li G, Yin W, Li X. 2016. The research status and forecast of Threskiornithidae in China. Chin J Wildlife. 2016;234–238.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760.

Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079.

Li MZ, et al. 2013. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nat Genet. 45(12):1431–1438.

Li SB, et al. 2014. Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. Genome Biol. 12(15):557.

Li YF, et al. 2018. Survival rates of a reintroduced population of the Crested Ibis Nipponia nippon in Ningshan County (Shaanxi, China). Bird Conserv Int. 28(1):145–156.

Louis-Dit-Picard H, et al. 2012. Amprolium.5. Studies on thiamine uptake, and transport. Biochemistry 53(5):821–835.

McQuillan R, et al. 2008. Runs of homozygosity in European populations. Ann Hum Genet. 83(3):359–372.

Narasingh V, et al. 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics 32(11):1749–1751.

Polin D, Porter CC, Wynosky ER. 1962. Amprolium.5. Studies on thiamine deficiency in laying chickens and their eggs. J Nutr. 76(1):59–68.

Xi Y, Lu B, Fujihara N. 2001. Captive rearing and breeding of the crested Ibis. Jpn Poult Sci. 38(3):213–224.

Wu JM, Mao XZ, Cai T, Luo JC, Wei LP. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. Nucleic Acids Res. 34(Web Server):W720–W724.

Xi Y, Lu B, Fujihara N. 2001. Captive rearing and breeding of the crested Ibis. Jpn Poult Sci. 38(3):213–224.

Xi YM, Wood C, Lu BZ, Zhang YM. 2007. Prevalence of a septicaemia disease in the crested Ibis (Nipponia nippon) in China. Avian Dis. 51(2):614–617.

Wu JM, Mao XZ, Cai T, Luo JC, Wei LP. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. Nucleic Acids Res. 34(Web Server):W720–W724.

Shi D, Cao Y. 2001. The crested Ibis of China. Beijing: China Forestry Publishing House.

Soulé ME. 1980. Thresholds for survival: maintaining fitness and evolutionary potential. Conservation Biology: An Evolutionary-ecological Perspective 9: 151–169.

Shi D, Cao Y. 2001. The crested Ibis of China. Beijing: China Forestry Publishing House.

Souce ME. 1980. Thresholds for survival: maintaining fitness and evolutionary potential. Conservation Biology: An Evolutionary-ecological Perspective 9: 151–169.

Sanz CS, et al. 2018. Uncovering the genetic lesions underlying the most severe form of Hirschsprung disease by whole-genome sequencing. Eur J Hum Genet. 26(6):818–826.

Taniguchi Y, et al. 2014. Structure and polymorphism of the major histocompatibility complex class I region in the Japanese crested ibis, Nipponia nippon. PLoS One 9(9):e108506.

Towend AK, et al. 2018. Low heterozygosity is associated with vector-borne disease in cows. Ecosphere 9(10):e02407.

Wu J, et al. 2017. Breeding ecology of a reintroduced population of the Japanese crested ibis. Zool Sci. 30(6):432–438.

Wardarajan BN, et al. 2018. Whole genome sequencing of Caribbean Hispanic families with late-onset Alzheimer’s disease. Ann Clin Transl Neurol. 5(4):406–417.

Zhang X, et al. 2015. Disease-causing mutations in KLHL3 impair its effect on WNK4 degradation. FEBS Lett. 587(12):1717–1722.

Wu J, et al. 2017. Breeding ecology of a reintroduced population of crested Ibis (Nipponia nippon) in Henan Province, China. Pak J Zool. 49:2027–2035.

Wu JM, Mao XZ, Cai T, Luo JC, Wei LP. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. Nucleic Acids Res. 34(Web Server):W720–W724.

Xie C, et al. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 39:316–322.

Zhang B, Fang SG, Xi YM. 2004. Low genetic diversity in the endangered crested Ibis Nipponia nippon and implications for conservation. Bird Conserv Int. 14(3):183–190.

Zhang B, Fang SG, Xi YM. 2006. Major histocompatibility complex variation in the endangered crested ibis Nipponia nippon and implications for reintroduction. Biochem Genet. 44(3–4):110–123.

Zhang X, et al. 2015. Ascariis spp. and Capillaria caudinflata infections in captive-bred crested ibis (Nipponia nippon) in China. Zoo Biol. 34(1):80–84.

Zhao HS, et al. 2016. Effect of junctional adhesion molecule-2 expression on cell growth, invasion and migration in human colorectal cancer. Int J Oncol. 48(3):929–936.

Associate editor: Laurence Hurst