Genetic Diversities among Founder Populations of the Endangered Avian Species, the Japanese Crested Ibis and the Oriental Stork in Japan

Yukio Taniguchi

2016

CONTENTS

GENERAL INTRODUCTION

The Japanese crested ibis *Nipponia nippon* and the Oriental stork *Ciconia boyciana* are endangered avian species, and their conservation is intensively promoted in Japan. Historically, both avian species were found in Japan, but habitat loss and overhunting caused a drastic decline in their numbers, resulting in their extinction. After their extinction in Japan, captive-breeding programs were implemented using relatively small numbers of individuals introduced from abroad as founders. As the size of the progeny population became larger, projects for the reintroduction of captive-bred individuals into the wild were launched.

It is necessary not only to increase the population size, but also to prevent loss of genetic diversity in these conservation projects. Genetic diversity refers to the genetic variability within species. Genetic diversity is important because it helps maintain the health of a population, by including alleles that may be valuable in resisting diseases and other stresses. Maintaining the diversity gives the population a buffer against change, providing the flexibility to adapt. If the environment changes, a population that has a higher variability of alleles will be better able to evolve to adapt to the new environment.

In population genetics, genetic diversity is frequently evaluated using pedigree information. Unfortunately, since the pedigree information among the founders of *N. nippon* and *C. boyciana* in Japan is not available, genetic relatedness within the founder population must be evaluated using molecular tools such as single nucleotide polymorphism (SNP) and short tandem repeat (STR) markers. However, information on the genome sequence or polymorphic markers in the both avian species remains sparse. Although the development of comprehensive polymorphic DNA markers is laborious

and expensive, next-generation sequencing with ultra-high-throughput might enable us to design an efficient, cost-effective approach for the discovery of genome-wide SNP and STR markers. Several methods of next-generation sequencing coupled with restriction enzyme digestion to reduce target complexity have been reported, such as reduced representation libraries (RRLs) (Kraus et al., 2011; Van Tassell et al., 2008), restriction-site-associated DNA sequencing (RAD-seq) (Baird et al., 2008), and complexity reduction of polymorphic sequences (CRoPS) (van Orsouw et al., 2007).

In Chapter 1, a large number of polymorphic markers in the Japanese crested ibis genome were developed using a combination of RRLs and next-generation sequencing. In many studies, RRLs have been prepared from pools of DNA samples from multiple individuals, thus allowing the detection of polymorphisms within a population but not for each individual. In this study, because the current Japanese crested ibis population originated from only 5 founder birds, 5 RRLs were independently prepared from each of the founder genomes to detect genome-wide polymorphic markers and their genotypes in each founder at the same time. Moreover, we investigated relative genetic similarities between the founders based on their genotypes at each marker.

The majority of SNPs and STRs are neutral makers. In addition to neutral makers, polymorphisms of functional genes could provide valuable information for the conservation of endangered species. The major histocompatibility complex (MHC) is a highly polymorphic genomic region that plays a central role in the immune system.

The MHC class I and class II genes encode glycoproteins that transport foreign peptides to the surfaces of cells for recognition by T-cell receptors on lymphocytes, which in turn triggers the adaptive immune response (Klein, 1986). Therefore, this genomic region is crucial for resistance and susceptibility to pathogenic disease.

Polymorphisms at MHC class I and class II genes facilitate binding of a diversity of pathogens, and these evolutionary selection pressures are thought to contribute to the high genetic variation in MHC loci (Zinkernagel and Doherty, 1979).

Moreover, the MHC genotype might influence patterns of mate choice, local adaptation, and expression of sexually selected ornaments (Von Schantz et al., 1997; Ekblom et al., 2007; Hale ML et al., 2009; Roberts, 2009). For these reasons, the diversity of MHCs is of major interest to the conservation of endangered species.

Polymorphism in the MHC is not restricted to allelic variation. The molecular evolution of the MHC involves frequent gene duplication and gene loss resulting in vast rearrangements and pronounced variation in gene number and genomic organization among organisms (Kulski et al., 2002; Kelley et al., 2005). In addition, copy number variations of MHC-IIB gene have been reported in many bird species (Hosomichi et al., 2006; Eimes et al., 2011; Strandh et al., 2012; Alcaide et al., 2014). These results suggest that the information of MHC genome structure is clearly required to exactly evaluate the genetic diversity in MHC region

In Chapters 2 and 3, the genomic organizations of MHC regions in *N. nippon* and *C. boyciana* were determined and then polymorphisms of MHC class II regions among their founder populations in Japan were investigated.

The information on genetic diversities and MHC structures presented here provides valuable insight for the conservation of *N. nippon* and *C. boyciana* and for future studies on the evolution of the avian MHC.

CHAPTER 1

Genome-Wide SNP and STR Discovery in the Japanese Crested Ibis and Genetic Diversity among Founders of the Japanese Population

Introduction

The Japanese crested ibis *Nipponia nippon* is an internationally conserved bird, listed as "Endangered" in the 2012 International Union for Conservation of Nature Red List of Threatened Species (http://www.iucnredlist.org).

The Japanese crested ibis once flew over much of Japan and northeastern Asia, but overhunting for the feathers and habitat loss devastated their numbers. After the Japanese crested ibis was extinct in Japan, captive-breeding programs have been continued with 5 birds (2 individuals introduced in 1999, 1 individual in 2000, and 2 individuals in 2007) donated by the Government of China, where a very small wild population survived (Liu, 1981). The current size of the captive-breeding population in Japan is approximately 180, mainly in the Sado Island. The idea of maintaining captive-bred animals for eventual release into the wild is a major aim of modern zoological collections (Durrell and Mallinson, 1987), and the Ministry of the Environment of Japan launched a project for tentative release of the Japanese crested ibis on Sado Island in 2008. In April 2012, 3 Japanese crested ibis chicks hatched on Sado Island and became the first of their species borne in the wild in 36 years (News of the week, Around the World, Sado Island, Japan, Back From the Brink, Science, 2012).

Conservation of small or captive populations requires particular concern for the loss of genetic diversity through genetic drift and inbreeding (Lande, 1988). Knowledge of genetic diversity and structure can be vital to the genetic management of captive

populations and reintroduction of captive-bred individuals into the wild. However, it is difficult to obtain precise knowledge of the genetic diversity and structure of the Sado captive population, because there is no pedigree information regarding kinship among the founders. Therefore, to improve management of the Japanese crested ibis toward national project goals, it is important to evaluate the genetic relatedness between the founders by using molecular tools such as single nucleotide polymorphism (SNP) and short tandem repeat (STR) markers. However, information on the genome sequence or polymorphic markers in the Japanese crested ibis remains sparse. Currently available genetic markers include only 26 microsatellites (Ji et al., 2004; He et al., 2006; Urano et al., 2013).

A general limitation of genome-wide polymorphic markers in non-model organisms has been a lack of extensive genomic sequence information from multiple individuals that represent a sufficient portion of the genetic variability of a given population or species. However, next-generation sequencing coupled with restriction enzyme digestion of target genomes to reduce target complexity, such as reduced representation libraries (RRLs) (Kraus et al., 2011; Van Tassell et al., 2008), restriction-site-associated DNA sequencing (RAD-seq) (Baird et al., 2008), and complexity reduction of polymorphic sequences (CRoPS) (van Orsouw et al., 2007), has provided an efficient approach to solving this problem (Davey et al., 2011). The Illumina HiSeq2000 sequencing system, released in 2010 provides the highest throughput available, and might enable us to design an efficient, cost-effective approach for the discovery of genome-wide SNP and STR markers.

The aim of this study was to develop a large number of polymorphic markers in the Japanese crested ibis genome by using a combination of RRLs prepared from 5 founder

genomes and next-generation sequencing. We also investigated relative genetic similarities between the founders based on their genotype at each marker.

Materials and Methods

Reduced representation library construction

Blood samples from the Japanese crested ibis were provided by the Sado Japanese Crested Ibis Conservation Center (Niigata, Japan). Protocols of sample collection were approved by the Animal Research Committee of Niigata University based on conservation project of the Ministry of the Environment of Japan. Genomic DNA samples were prepared from whole blood using the Wizard Genomic DNA Purification kit (Promega) according to the manufacturer's instructions with slight modification. Blood (60 µl) was washed with 3 mL PBS, and red and white blood cells were lysed with 6 mL Nuclei Lysis Solution (Promega). Aliquots of 80 µg of genomic DNA from each individual were digested with 800 units of *Hae*III or *Mbo*I (TAKARA) overnight at 37°C. Digested DNA was separated by 1.5% agarose gel, and digestion products of 250–350 bp were gel purified using the Wizard SV Gel and PCR Clean-Up system (Promega) according to the manufacturer's instructions. The *Hae*III- and *Mbo*I-digested fragments were combined and processed as 1 RRL for sequencing on the HiSeq2000 (Illumina). The RRLs were independently prepared from each of 5 founders.

Sequencing and data analysis

Sequencing was performed at the Hokkaido System Science Co. Ltd. (Sapporo, Japan). Briefly, the combined DNA fragments were end repaired and ligated with the sequence adaptor using the TruSeq DNA Sample Prep Kit (Illumina). The RRLs were

distinguished by adding sequence adaptors with different index sequences. The RRLs were pooled and sequenced in a single sequencing lane on the HiSeq2000 for 101 cycles in pair-end mode. Raw data files from the sequencing instrument were deposited in the DDBJ sequence read archive under accession number DRA000585.

Primary data analysis was also performed at Hokkaido System Science. After adaptor trimming with the cutadapt program (http://code.google.com/p/cutadapt/) and discarding reads containing N bases, filter-passed sequence reads from the founder RRLs were divided into 3 groups by their 5′-terminal sequences (both-end *Hae*III, both-end *Mbo*I, and others). Sequence reads within a group were clustered by the clustering program "SEED" (http://manuals.bioinformatics.ucr.edu/home/seed), and consensus sequences of 300 bp (read-pair) were generated. These consisted of forward and reverse 101-bp reads with internal 98 bases of N. Parameters in the program "SEED" were as follows: --shift was 0 and other parameters were defaults. Consensus sequences with depth ≥10 were used as reference sequences for mapping of read pairs from each founder. Consensus sequences used for mapping were deposited in the DDBJ sequence read archive under accession number DRZ002863. Mapping was performed by the short read aligner program "bowtie" (http://bowtie-bio.sourceforge.net/index.shtml). Parameters in the program "bowtie" were as follows: -I was 300, -X was 300, -v was 3, and --best option was specified.

SNP discovery, genotyping, and haplotyping

The 123,506 predicted SNPs whose depth was ≥100 were extracted from the mapping results. Putative SNPs were selected by the following filtering processes: (1) Alleles with a depth of 1 in a founder sample at predictive SNP positions were ignored. (2)

SNPs with a depth of more than 300 reads in any individual were filtered out. (3) If the read pairs for an allele were more than 5% of the total depth from a founder, the corresponding alleles were considered present. Then, SNPs with 3 or more alleles in any individual were discarded. (4) After the predictive SNPs with 2 alleles were identified, predictive SNPs for which the depth ratio between the 2 alleles in any individual was more than 3 were also removed. We used the putative SNPs with depth ≥20 in each founder for the following genetic analysis.

For haplotyping, all consensus sequences with polymorphisms at more than 1 position were extracted. Then, a set of SNPs with the same depth within 202-bp consensus sequences was treated as a haplotype (if the difference in depth between SNPs was <4 due to sequence error, they were assumed to have the same depth).

Analysis for genetic similarities between founders

The numbers of single founder-specific alleles and the heterozygous and homozygous loci were counted in each founder across the putative SNPs with depths ≥20 in each of the founders. Then, the number of loci with the same genotype was also computed. Principal component analysis (PCA) and multidimensional scaling (MDS) were performed using the princomp function with cor=T option and the cmdscale function with default option by setting 1 correlation as the distance measure in R (http://www.R-project.org), respectively, where counts of the major allele for each locus were used to calculate correlation matrix between founders. Hierarchical clustering was also carried out using the R package pvclust (http://www.is.titech.ac.jp/~shimo/prog/pvclust/) to evaluate stability in the clustering results through multiscale bootstrap resampling. We applied “average” and “correlation”

options for the method of agglomerative clustering and the distance measure, respectively, and computed approximately unbiased (AU) p-value and bootstrap probability (BP) value based on 10,000 bootstrap replications.

STR discovery

All consensus sequences containing 8 or more di-nucleotide tandem repeats, 5 or more tri-nucleotide tandem repeats, or 4 or more tetra-nucleotide tandem repeats were extracted. Consensus sequences that were identical other than those in the repeat-sequence region were grouped by self-mapping with the short read aligner program “bowtie”. Using the mapping results described above, we counted the number of read pairs corresponding to each STR allele.

Results

Sequencing strategy

To discover genome-wide SNP and STR markers in the Japanese crested ibis, we generated sequences from RRLs using the next-generation HiSeq2000 sequencer (Illumina). We simultaneously obtained genotype data on each of the markers by sequencing the RRLs prepared independently from the founder genomes.

At the time of this study, the HiSeq2000 produced sequences of 150–200 million DNA fragments with 100-bp read length in 1 sequencing lane. This allowed us to analyze DNA fragments from 0.5 to 1 million loci per genome, with 5 DNA samples and a sequencing depth of approximately 30 (5 sample × 1 million loci × 30 depth = 150 million). Pair-end sequencing with 100-bp read length of 0.5–1 million fragments resulted in 100–200 Mb, which was estimated to represent 8–13% of the Japanese

crested ibis genome, supposing the genome size is 1.5 Gb based on several entries in the eukaryotic genome size databases (Gregory et al., 2007).

To prepare RRL with the desired number of fragments, we digested genomic DNA with *Hae*III or *Mbo*I and extracted fragments in the 250–350 bp size range from agarose gels. In a preliminary experiment, the number of size-selected DNA fragments by digestion with *Hae*III and *Mbo*I were estimated to be 0.34 and 0.44 million, respectively, from the yield of isolated DNA fragments. We also chose DNA fragments in the 250–350 bp size range to prevent decreasing sequence data by overlapping forward and reverse 100-bp sequence reads from a restriction fragment. Restriction fragments generated by digestion with *Hae*III and *Mbo*I were combined and processed as a single RRL for sequencing.

The RRLs were independently prepared from each of the 5 founder genomes. Each RRL was distinguished by adding sequence adaptors with different index sequences. The RRLs were pooled and sequenced on a single sequencing lane on a HiSeq2000 instrument for 101 cycles in pair-end mode.

Table 1. Summary of DNA sequence

	Total	founder individual				
		A	B	C	D	E
No. of 101bp reads	339,597,768	51,920,408	63,860,862	78,009,738	73,383,930	72,422,830
No. of bases (Mb)	34,300	5,244	6,450	7,879	7,412	7,315
% of ≧ Q30 Bases		92.3	92.3	92.2	92.2	92.2
No. of trimmed reads	316,436,996	48,357,306	59,497,380	72,706,332	68,382,502	67,493,476

% of ≧ Q30 Bases : proportion of high quality bases (≧Q30) in filter-passed bases

Illumina sequencing results

We sequenced the 5 RRLs in pair-end mode, generating 339 million 101-bp reads (Table 1). The proportion of high-quality bases (≥Q30) over all sequence reads was >92% in every sample. After adaptor trimming and discarding reads containing N bases,

the remaining 316 million reads were used for analysis. The number of reads in each founder was 48–72 million (Table 1).

Table 2. No. of Created consensus sequences and putative SNPs

	Total	group		
		Both-End HaeIII	Both-End MboI	Others
No. of read-pairs	158,218,498	85,131,252	9,024,997	64,062,249
No. of consensus sequence	31,418,852	4,175,097	952,879	26,290,876
No. of consensus sequence (depth≧10)	1,754,793	465,471	249,515	1,039,807
No. of consensus sequence (mapping depth≧100 in total of 5 birds)	532,712	294,989	13,353	224,370
No. of putative SNP	52,512	28,764	321	23,427
No. of putative SNP (mapping depth≧20 in each of 5 birds)	32,157	16,334	224	15,599

Read pairs combined with paired forward and reverse 101-bp reads were divided into 3 groups by their 5′-terminal sequence (both-end *Hae*III, both-end *Mbo*I, and others) (Table 2). The read-pairs within a group were clustered and consensus sequences were created (Table 2). In total, 31,418,852 consensus sequences were created. The number of consensus sequences with depths (counts of read pairs clustered to identical sequence) ≥10 was 465,471, 249,515, and 1,039,807 in both-end *Hae*III, both-end *Mbo*I, and others, respectively. Though different groups could contain a set of overlapping sequences, estimation from the number of consensus sequences with ≥10 depth would mean that the sequence information generated here represented at least 6–10% of the Japanese crested ibis genome (0.46 million for *Hae*III to 0.71 million for *Hae*III+*Mbo*I, multiplied by 202 base yielded 0.09–0.14 Gb).

SNP prediction

Because no reference genome sequence is available for the Japanese crested ibis, we searched putative SNPs by mapping read pairs from each founder to consensus

sequences (depth ≥10) and filtering (see Materials and Methods for criteria). Approximately 70% of read pairs from each founder were mapped (Table 3), resulting in 532,712 consensus sequences with depth ≥100 in all 5 founders (Table 2). Out of the 123,506 predictive SNPs in these consensus sequences, 52,512 (42.5%) putative SNP markers were detected, fulfilling the criteria (Table 2, the list of all putative SNP sites is provided in supplementary Table S1). Further, the number of the putative SNPs with depth ≥20 in each founder was 32,157 (Table 2), and these putative SNPs were used for the collection of genotype data. The list of all genotype data is provided in supplementary Table S2.

Table 3. Mapping results

	founder individual				
	A	B	C	D	E
No. of read-pairs	24,178,653	29,748,690	36,353,166	34,191,251	33,746,738
No. of mapped read-pairs	16,154,580	20,825,091	26,120,640	23,684,263	22,881,668
% of mapped read-pairs	66.8	70.0	71.9	69.3	67.8

As the 4,842 consensus sequences contained multiple putative SNP sites within a 202-bp sequence; their haplotypes were deduced from mapping data (the list of all consensus sequences containing multiple SNP sites is provided in supplementary Table S3). Of these, haplotypes could be determined in 4,080 (84.3%) consensus sequences, but not in 762 (15.7%) consensus sequences. The deduced haplotype numbers were 2–4 in most loci (Table 4).

Table 4. Number of haplotype on consensus sequence containing multiple SNP sites

Number of haplotype per locus	2	3	4	5	6
No. of locus	2,750	1,054	258	13	5

Genetic similarities between founders

The genotype data on 32,157 putative SNPs in each of the founders were used to

analyze genetic similarities between them. Single founder-specific allele numbers were 2,087, 1,367, 2,305, 1,676, and 1,003 in founders A, B, C, D, and E, respectively (Table 5). Proportions of heterozygous genotypes and of SNPs whose genotypes were common in pair-wise combination were calculated. The proportion of heterozygous genotypes in each founder was 0.49–0.56 (Table 5). The proportion of SNPs whose genotypes were common in 2 founders was 48.5–59.4% (Table 6). Founders B and E had the highest proportion of common genotypes. We performed PCA using the 32,157 SNPs and used the first 2 principal components (PCs 1 and 2) to visualize the degree of relative genetic similarities among the 5 founders, where PC1 accounted for 32.7% of the variation, while PC2 accounted for an additional 23.1%. This analysis revealed that each individual was located in a relatively dispersed position, although founders B and D were plotted relatively closer (Fig. 1). The results of MDS and hierarchical clustering (Figs. 2 and 3) were similar to the result from PCA, and AU and BP values by a bootstrap procedure indicated that the dendrogram topology was stable (Fig. 3). The results of common SNP genotype and the multivariate analyses were slightly inconsistent owing to the difference between genotype sharing and allele sharing, but seemed to indicate that the genomes of founders B, D, and E shared significant similarities.

Table 5. Single founder−specific alle and SNP genotype

	founder individual				
	A	B	C	D	E
No. of specific allele	2,087	1,367	2,305	1,676	1,003
major homo	12,871	14,504	13,613	15,684	15,506
minor homo	1,185	462	1,450	785	610
hetro	18,101	17,191	17,094	15,688	16,041
heterozygosity	56.3	53.5	53.2	48.8	49.9

Table 6. Number of common SNP genotype in pair-wise						
	founder					
	A	B	C	D	E	
A		16715	15611	15931	19034	
B	52.0		16403	18682	19095	
C	48.5	51.0		16112	16410	
D	49.5	58.1	50.1		17002	
E	59.2	59.4	51.0	52.9		
The numbers (above the diagonal) and the frequency (below the diagonal) were shown.						

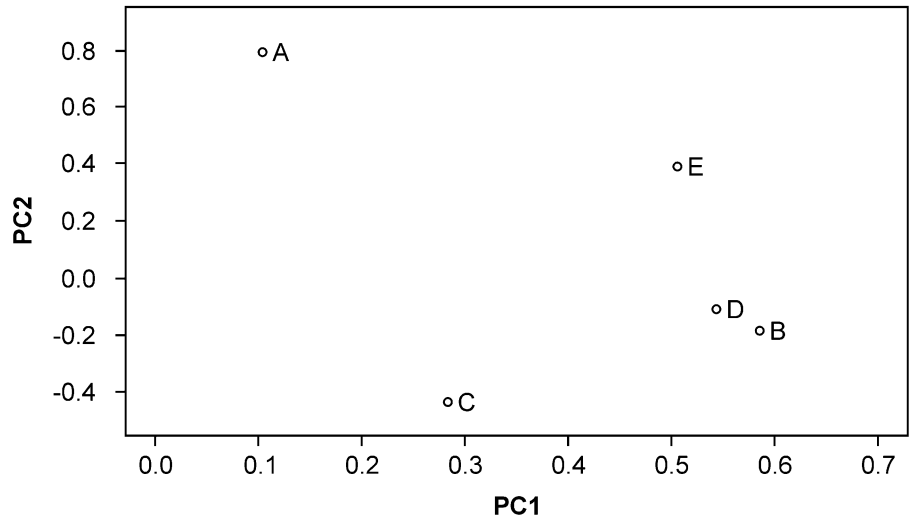

Figure 1. Principal component analysis for the 5 founders by using genotyping data of 32,157 putative SNPs

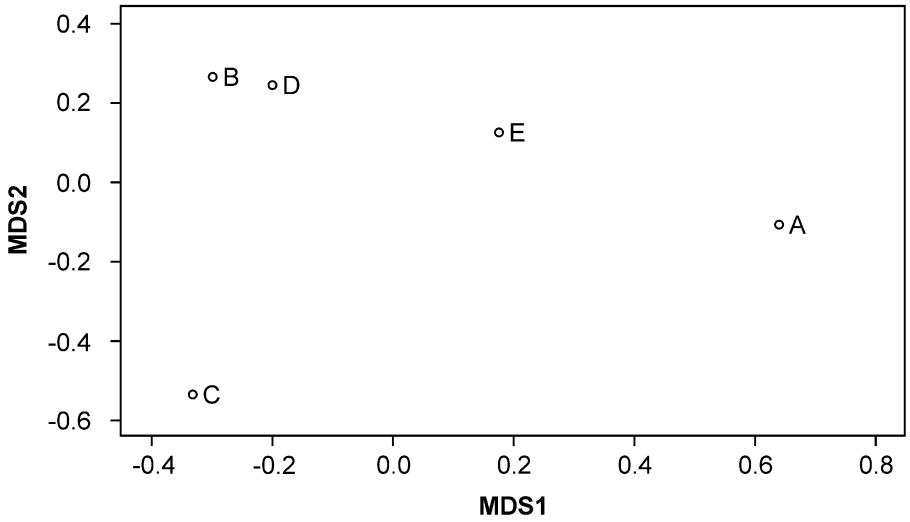

Figure 2. Multidimensional scaling analysis for the 5 founders by using genotyping data of 32,157 putative SNPs.

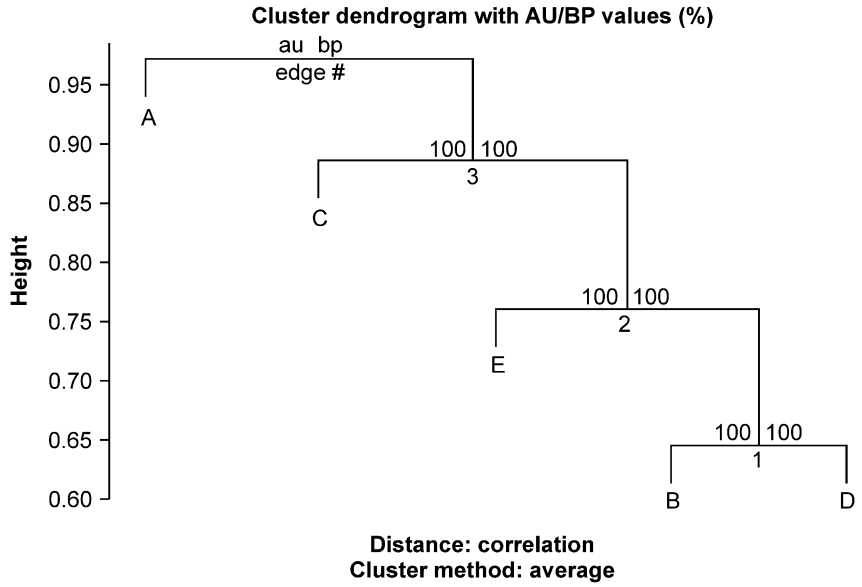

Figure 3. Hierarchical clustering of the 5 founders based on the similarities of their genotype patterns at 32,157 putative SNPs. Values at branch nodes represent AU values (left), BP values (right), and cluster labels (bottom).

STR prediction

To detect STR markers, we extracted all consensus sequences containing 2-, 3-, or 4-nucleotide tandem repeats; we detected 162 putative STR markers, of which 155 STRs were 2 allelic and only 7 STRs were 3 allelic (all putative and all genotyped STRs are listed in supplementary Tables S4 and S5, respectively). The numbers of single founder-specific alleles at the 86 STR markers genotyped in every founder were 8, 4, 8, 2, and 2 in founders A, B, C, D, and E, respectively (Table S5).

Discussion

The Japanese crested ibis *Nipponia nippon* is a critically threatened species and an internationally conserved bird. The Ministry of the Environment of Japan has been engaged in a captive breeding and hopes to release the Japanese crested ibis on Sado

Island. Whereas genetic management is critical for these projects, information on the genome sequence or polymorphic markers remains sparse. Currently available genetic markers include only 26 microsatellites (Ji et al., 2004; He et al., 2006; Urano et al., 2013).

Several methods of next-generation sequencing coupled with restriction enzyme digestion to reduce target complexity have been developed for the discovery of genome-wide genetic markers, such as RRL (Kraus et al., 2011; Van Tassell et al., 2008), RAD-seq (Baird et al., 2008), and CRoPS (van Orsouw et al., 2007). Next-generation sequencing of RRLs has been effective in identification of SNPs in species with reference genome sequences, such as mallard (Kraus et al.,2011), pig (Ramos et al., 2009; Wiedmann et al., 2008), and cattle (Van Tassell et al., 2008), and in species without reference genome sequences, such as the turkey (Kerstens et al., 2009) and great tit (van Bers et al., 2010). In many studies, RRLs have been prepared from pools of DNA samples from multiple individuals, thus allowing the detection of polymorphisms within a population but not for each individual.

Because the current Japanese crested ibis population originated from only 5 founder birds, we aimed to detect genome-wide polymorphic markers and their genotype in each founder at the same time. We developed an approach using next-generation sequencing and RRL. Five RRLs were prepared from each of 5 founder genomes and distinguished by ligating a sequence adapter containing a different index sequence. The 5 RRLs were pooled and sequenced on a single sequencing lane on the Illumina HiSeq2000 sequencing instrument.

Sequence information, including 316 million 101-bp reads (more than 31 Gb) from a single sequencing lane on the Illumina HiSeq2000, was sufficient for the discovery of

genome-wide genetic markers, providing an extremely cost-effective approach.

In this study, 52,512 putative SNPs were detected by creating consensus sequences by clustering sequence reads, mapping sequence reads from each founder to the consensus sequences, and filtering the predicted SNPs obtained by mapping. Of these, the 32,157 putative SNPs whose depth was ≥20 in each founder were selected to analyze genetic similarities. As the 4,842 consensus sequences contained multiple putative SNP sites within a 202-bp sequence, their haplotypes were deduced from mapping data (supplementary Table S3). Haplotypes could be determined in 4,080 (84.3%) consensus sequences but could not be determined in 762 (15.7%) consensus sequences, suggesting that these sequences represented multiple loci or sequence errors. These results suggest that putative SNPs include a considerable number of false SNPs. However, if 30% of putative SNPs were false, the remaining 70% could provide a sufficient number of markers for genetic management of the Japanese crested ibis population.

Approximately 52,000 putative SNPs (28,764 in both-end *Hae*III) were found in 530,000 of 202-bp consensus sequences (294,989 in both-end *Hae*III) (Table 2). A rough estimation based on this frequency suggested that the whole genome of the Japanese crested ibis contained approximately 700,000 SNP sites. Because approximately 50% of SNP sites were homozygous in a single individual (Table 5), the number of heterozygous SNPs in a single individual was found in approximately 350,000 sites (the SNP map might have an average density of one SNP per 2000 bp). This may be an overestimation because the putative SNPs detected here apparently included a considerable number of false SNPs.

In the whole-genome sequencing of a single giant panda individual (an endangered species), 2.7 million heterozygous SNPs were detected (1 SNP per 750 bp) (Li et al.,

2010). This is approximately 1.95 times higher than that estimated for humans (1 SNP per 1450 bp) (Wang et al. (2008). In thoroughbred horses, which are derived from a few founders, 0.75 million heterozygous SNPs were detected (1 SNP per 3,000 bp) (Wade et al., 2009). The number of heterozygous SNPs in a single Japanese crested ibis might be much lower than that in pandas and humans, and comparable to or lower than that in thoroughbred horses.

In contrast to SNP markers, which are usually biallelic, STR markers are expected to be multiallelic (3 or more alleles). We extracted consensus sequences containing short tandem repeats and detected 162 putative STR markers. Of these, 155 STRs were biallelic and only 7 STRs were triallelic. We detected no STRs with 4 or more alleles. Moreover, deduced haplotype numbers on consensus sequences containing multiple putative SNP sites were 2–4 in most cases (Table 4). The allele numbers in several tens of STRs previously developed were 2-5 in Chinese population and 2 or 3 in Japanese population (Ji et al., 2004; He et al., 2006; Urano et al., 2013). The only 2 haplotypes in mitochondria DNA control region were detected in Chinese wild and captive populations (Zhang etal. 2004). Our results obtained using a large number of genome-wide markers supported lower genetic diversity in the Japanese crested ibis populations previously estimated from a small number of markers. It was reasonable that the genetic diversity in Japanese population was somewhat lower than that in Chinese population because 5 founders of Japanese population originated from China.

Unfortunately, because of the absence of a reference genome sequence for the Japanese crested ibis, we could not determine whether putative SNPs and STRs represented polymorphisms at a single locus or multiple loci associated with repeated sequences or gene families. In addition, information about the locations of putative

SNPs and STRs on chromosomes or linkage between markers remains unknown. To determine whether the putative SNPs or STRs in this study were true heritable genetic markers, further analysis is necessary.

Although validation of SNP and STR markers has not yet been performed, we thought that the genotype data on putative SNPs or STRs in each of the 5 founders could be useful for analyzing the relative genetic similarities between them. The proportion of heterozygous genotypes in each founder was 0.49–0.56 (Table 5). The proportion of SNPs whose genotypes were common in 2 founders was 48.5–59.4% (Table 6). PCA and MDS indicated that each individual was located in a relatively dispersed position, except for founders B and D plotted closely (Figs. 1 and 2). These results suggest that genome similarities were not high. Whereas no pair having closely related genome composition was observed, smaller numbers of 202-bp read-pair haplotypes and STR alleles suggested that the genetic diversity of the population in total was much lower than that expected when they were unrelated (i.e., 10 of maximum haplotype or allele number in 5 birds). Lower genetic diversity in a population might be reflected by a smaller number of alleles and haplotypes at any locus and/or longer linkage disequilibrium, rather than by total number of SNPs in whole genome or proportion of heterozygous genotypes.

The comparison of genotypes at each putative SNP revealed that each of the 5 founders had 1200–2000 potential single-founder specific alleles. The loss of these specific alleles will directly reduce genetic diversity in the population. Therefore, it is important that single-founder specific alleles are passed on to some descendants, increasing the allele frequencies in the population.

The availability of a large number of SNPs and STRs predicted here provides

sufficient markers to study the Japanese crested ibis population structure and to develop methods for parentage testing, individual identification, and genetic management. Further analysis of a large number of accurately inferred polymorphic markers will also facilitate the construction of linkage maps of the Japanese crested ibis genome.

In conclusion, this study provided important insight into protocols for genetic management of the captive breeding population of Japanese crested ibis in Japan and will help in extending the national project for reintroduction of captive-bred individuals into the wild.

We proposed a simple, efficient, and cost-effective approach for the simultaneous detection of genome-wide polymorphic markers and their genotype data for species lacking a reference genome sequence. Our proposed approach might be useful for an extremely small population such as an endangered species or a population originating from a small number of dominant founders.

Supporting Information

Table S1 Mapping results of 52512 predictive SNPs

Table S2 Mapping results and genotypes 32157 putative SNPs

Table S3 Haplotypes in consensus sequences containing multiple SNP sites

Table S4 Mapping results of 162 putative STRs

Table S5 Mapping results and genotypes of 86 putative STRs

Supporting Information are not shown here because of extremely large excel files.

Please refer the article of "Taniguchi et al., (2013) PLoS ONE 8(8): e72781"

CHAPTER 2

Structure and Polymorphism of the Major Histocompatibility Complex Class II Region in the Japanese Crested Ibis, *Nipponia Nippon*

Introduction

The major histocompatibility complex (MHC) is a highly polymorphic genomic region that plays a central role in the immune system of all jawed vertebrates. The MHC class I and class II genes encode glycoproteins that transport foreign peptides to the surfaces of cells for recognition by T-cell receptors on lymphocytes, which in turn triggers the adaptive immune response (Klein, 1986). Therefore, this genomic region is crucial for resistance and susceptibility to pathogenic disease. Polymorphisms at MHC class I and class II genes facilitate binding of a diversity of pathogens, and these evolutionary selection pressures are thought to contribute to the high genetic variation in MHC loci (Zinkernagel and Doherty, 1979). The MHC class II molecule is a heterodimer consisting of an α and a β chain, which are encoded by MHC-IIA and -IIB genes, respectively.

Polymorphism in the MHC is not restricted to allelic variation. The molecular evolution of the MHC involves frequent gene duplication and gene loss that result in vast rearrangements and pronounced variation in gene number and genomic organization among organisms (Kulski et al., 2002; Kelley et al., 2005). In birds, two MHC-IIB lineages (*DAB1* and *DAB2*) have been characterized (Burri et al., 2008). Phylogenetic reconstructions and simulations using 63 MHC-IIB exon 3 sequences from six avian orders have suggested that a unique duplication event preceding the major avian radiations gave rise to ancestral MHC-IIB lineages that were each likely

lost once later during avian evolution (Burri et al., 2010). However, to obtain deeper insights into the long-term evolutionary history of the avian MHC, more data from other exons, other genes and ultimately genomic structures are required

The domestic chicken (*Gallus gallus*, order Galliformes) has been most intensively studied and its MHC (also known as MHC-B or B-complex) has a remarkable structure referred to as a "minimal essential MHC" (Kaufman et al., 1999). In contrast to the human MHC (human leukocyte antigen, HLA), which spans approximately 7.6 Mb and contains 421 gene loci on a contiguous region (Horton et al., 2004), the chicken MHC-B consists of only 19 genes spanning 92 kb (Kaufman et al., 1999). In addition to MHC-B, chicken MHC class I and class II genes are present in a separate and unlinked cluster called the MHC-Y region (Miller et al., 1994; Delany et al., 2009).

The overall MHC-B structures of five other galliform species, Japanese quail (*Coturnix coturnix japonica*), domestic turkey (*Meleagris gallopavo*), golden pheasant (*Chrysolophus pictus*), black grouse (*Tetrao tetrix*), and greater prairie-chicken (*Tympanuchus cupido*) are largely similar to that of chicken, whereas gene number, order, and orientation in these structures vary among the species (Hosomichi et al., 2006; Chaves et al., 2009; Ye et al., 2012; Wang et al., 2012; Eimes et al., 2013).

However, studies of non-galliform species, such as duck (*Anas platyrhynchos*, order Anseriformes) and zebra finch (*Taeniopygia guttata*, order Passeriformes) have suggested that the chicken minimal essential MHC is not typical among birds (Ren et al., 2011; Balakrishnan et al., 2010; Ekblom et al., 2011). For example, the chicken MHC-B contains two MHC-IIB genes (*BLB1* and *BLB2*) but no MHC-IIA genes (*BLA*), whereas the duck MHC possesses one MHC-IIA and five MHC-IIB genes (Ren et al., 2011). Because taxonomic and genomic sampling of avian MHC regions is limited, it is

unclear whether the minimal essential MHC represents the ancestral condition or whether it is a highly derived condition unique to the Galliformes.

The Japanese crested ibis (*Nipponia nippon*, order Pelecaniformes) is an internationally conserved bird, listed as "Endangered" in the 2012 International Union for Conservation of Nature Red List of Threatened Species (http://www.iucnredlist.org). The range of *N. nippon* formerly included much of Japan and northeastern Asia, but habitat loss and overhunting for its feathers have caused a drastic decline in its numbers and resulted in its extinction in Japan. Captive-breeding programs have been conducted using five birds as founders (two individuals introduced in 1999, one introduced in 2000, and two introduced in 2007) provided by the Chinese government. The current size of the captive-breeding population in Japan is approximately 210 birds, most of which are on Sado Island. The Ministry of the Environment of Japan launched a project to release *N. nippon* on Sado Island in 2008; in April 2012, three chicks hatched there, the first of this species born in the wild in 36 years (News of the week, Around the World, Sado Island, Japan, Back From the Brink, Science, 2012).

Molecular ecology studies have shown that, in addition to adaptive immune responses, the MHC genotype influences patterns of mate choice, local adaptation, and expression of sexually selected ornaments (Roberts, 2009; Ekblom et al., 2007; Hale ML et al., 2009; Von Schantz et al., 1997). For these reasons, the diversity of MHCs is of major interest to the conservation of endangered species.

The Japanese crested ibis belongs to an avian lineage that is highly divergent from that of chicken, duck, or zebra finch. Characterizing the MHC class II region in *N. nippon* may provide valuable information about the primordial avian MHC. In addition, insight into the genetic diversity of this genome region could be vital to conservation of the *N.*

nippon population. In this study, we investigated the genomic organization of the *N. nippon* MHC class II region and polymorphisms among the 5 founders of the Japanese population.

Materials and Methods

Samples

Blood samples from *N. nippon* (5 founders, A–E, and 20 progeny) and the liver from a dead female were provided by the Sado Japanese Crested Ibis Conservation Center (Niigata, Japan). Sample-collection protocols were based on a conservation project of the Ministry of the Environment of Japan and approved by the Animal Research Committee of Niigata University. Genomic DNA samples were prepared from whole blood and liver using the Wizard Genomic DNA Purification kit (Promega) according to the manufacturer's instructions.

Primers

Primers and annealing temperatures used for polymerase chain reaction (PCR) analysis are shown in Table S1. To determine the partial sequence of the Japanese crested ibis MHC class IIB gene, degenerate primers BRMHC05 and AIEx3R (Table S1) (Miller and Lambert, 2004; Alcaida et al., 2007) were used for amplification and an amplified 1,057 bp fragment containing exon 2 was cloned and then sequenced. The resulting sequence represented a part of the *DAB2* locus (Figure 1). Primers 2F_pen1 and intron2-01R (Table S1) were designed on the basis of a previous report (Tsuda et al., 2001) and the sequence of this PCR product. Other primers were designed on the basis of sequences determined in this study.

Polymorphisms and PCR-restriction fragment length polymorphism (RFLP) of MHC-IIB Exon 2

A 279 bp fragment of MHC-IIB exon 2 was amplified using primers 2F_pen1 and intron2-01R (Table S1) from two founders (D and E) and 20 progeny (At the time, three founder samples, A–C were not available to us). In order to analyze polymorphism, PCR products were cloned using the TOPO TA Cloning Kit (Invitrogen) and at least 8 positive clones were sequenced per individual. Based on the sequence information of these clones, we identified possible restriction sites for further PCR-RFLP analysis and chose the restriction enzymes *Pst*I, *Rsa*I, or *Sal*I. Digested fragments were separated on a 3% agarose gel and revealed individual banding patterns for each of the different MHC-IIB exon 2 sequences. MHC class II genotypes of five founders were examined by PCR-RFLP.

Construction of a genomic library and Screening

To determine the genomic structure of the *N. nippon* MHC class II region, two genomic libraries were constructed from a dead female and founder E. Genomic DNA was partially digested with *Sau*3AI and separated on a 0.5% agarose gel. Digested fragments (15–23 kb) were gel purified using the Wizard SV Gel and PCR Clean-Up system (Promega), according to the manufacturer's instructions, and ligated into the Lambda DASH II vector (Stratagene). The ligated DNA mixture was then packaged using Gigapack III Gold or XL packaging extract (Stratagene). Screening was performed using a PCR-based method (Figure S1). Five lambda phage clones (16-D8, 6-F3, 5-3-1, 19-A1 and 1-C1) were isolated by the PCR-screening with primers

2F_pen1 and intron2-01R (Figure 1 and Table S1). On the basis of the sequences of both ends in two lambda phage clones (16-D8 and 1-C1), primers 16-D8Rev_S1, 16-D8Rev_A1, 16-D8Uni_S1, 16-D8Uni_A1, 1-C1BRD2_S1 and 1-C1BRD2_A1 (Table S1) were designed and used for genome walking. Two lambda phage clones (2-L10 and 6-S9) were isolated by the PCR-screening with 16-D8Uni_S1/A1 primer pair, as were a clone 6-F3 with 16-D8Rev_S1/A1 primer pair and two clones (16-C4 and 13-F9) with 1-C1BRD2_S1/A1 primer pair (Figure 1). These positive clones were isolated from the genomic library constructed from a dead female. In addition, two positive clones (XL2-6-8 and XL10-T2) were isolated from the founder E genomic library by the PCR-screening with primers 2F_pen1 and intron2-01R (Figure 1 and Table S1). .

Analysis of isolated lambda phage clones

Phage DNAs of positive clones were purified, digested with *Bam*HI and/or *Eco*RI, and subcloned into pBluescript II vector (Stratagene). In some lambda phage clones, additional restriction enzymes (*Kpn*I, *Sac*I, *Sal*I and/or *Xho*I) were also used (Figure 1). Sequences of both ends of some subclones were determined and analyzed through homology searches using BLAST (http://blast.ncbi.nlm.nih.gov/blast). Moreover, subclones containing MHC-IIB exon 2 were identified by colony-PCR with primers 2F_pen1 and intron2-01R (Table S1) and analyzed by PCR-RFLP for exon 2 sequence. By combining with these results, restriction maps of each lambda phage clones were constructed.

For complete sequencing of MHC class II genes, insert DNA of subclones containing MHC class II genes was completely or partially digested with *Sac*I or *Sau3*AI and the

resulting fragments were re-cloned into pBluescript II vector (Stratagene). Positive clones were sequenced using M13 forward and reverse primers. The remaining gaps were filled by sequencing with specific primers. Every nucleotide position was sequenced at least twice using both strands of a single subclone or the same strand of several subclones. Sequencing was carried out by Greiner Japan Co., Ltd. (Tokyo, Japan). Sequences were manually assembled using GENETYX version 11 (Software Development). The sequences of MHC class II gene regions from haplotypes 1, 2, and 3 (approximately 8, 13, and 18 kb, respectively) were deposited in the DNA Data Bank of Japan (DDBJ) under accession numbers AB872442, AB872443, and AB872444, respectively. The sequence of the *BRD2* gene in haplotype 1 was deposited in the DDBJ under accession number AB890383. The partial sequence of *COL11A2*-like gene is shown in Figure S2.

Gene Identification and Sequence Analysis

The assembled sequences were analyzed for coding regions using the GENSCAN program (http://genes.mit.edu/GENSCAN.html) with vertebrate parameters and through homology searches using BLAST (http://blast.ncbi.nlm.nih.gov/blast). Sequence alignments and dot-matrix analyses were performed using GENETYX with the default parameters.

Construction of Phylogenetic Tree

Phylogenetic relationships of MHC-IIB genes within Pelecaniformes as well as among avian orders were analyzed using exon 2 or partial exon 3 (first 128bp) sequences, since only 662 bp sequences containing full exon 2, intron 2 and partial exon3 are available in

Pelecaniformes except for the Japanese crested ibis (Li et al., 2011). Eleven Pelecaniformes and two species from each of seven avian orders used for the analyses were shown in Table S3. The best-fitting nucleotide substitution model for each codon position was evaluated using Find Best DNA/Protein Models (ML) in MEGA version 5.2 (Tamura et al., 2011) according to the Akaike information criterion. A maximum-likelihood tree with exon 2 sequences was constructed by using a Kimura 2-parameter model with gamma distribution in MEGA. A maximum-likelihood tree with partial exon 3 sequences was constructed by using a Tamura 3-parameter model with gamma distribution in MEGA.

Southern blotting

In order to confirm haplotype identity for all founder individuals and to examine a presence of additional MHC loci, we performed Southern blotting. Genomic DNA (10 μg) from the five founders was digested with *Bam*HI and *Eco*RI and separated on a 0.7% agarose gel. These restriction enzymes were chosen on the basis of restriction maps and sequences of three haplotypes and preliminary experiments. The gel was blotted onto Hybond-N+ nylon membrane (GE Healthcare) and immobilized by UV cross-linking. Three fragments (238 bp of MHC-IIA exon 3, 279 bp of MHC-IIB exon 2, and 307 bp of MHC-IIB exon 3) were used as probes for the detection of MHC-IIA/IIB fragments. The probes were labeled using a PCR DIG Probe Synthesis Kit (Roche) with primers IIAex3-F and IIAex3-R for MHC-IIA exon 3, 2F_pen1 and intron2-01R for MHC-IIB exon 2, and IIBex3-F and IIBex3-R for MHC-IIB exon 3 (Table S1). Hybridization and detection were performed according to the manufacturer's instructions.

Results

Polymorphism of MHC-IIB Exon 2

To determine the polymorphism of the Japanese crested ibis MHC class IIB gene, 279 bp of exon 2 sequences were amplified from two founders (D and E) and 20 progeny. Sequencing of the PCR products revealed four types of exon 2 sequences (temporally named types I–IV for the 279 bp exon 2 sequences), suggesting that the *N. nippon* population contained at least four alleles of the MHC class IIB gene (see below). Each individual possessed one to three exon 2 sequences (Table S2). We developed a PCR-RFLP method to distinguish the type (I–IV) of MHC-IIB exon 2. The four sequences of exon 2 could be assigned to types by the combinations of profiles digested with restriction enzymes of *Pst*I, *Rsa*I, or *Sal*I (Table 1).

Table 1. Digestion profiles of 279 bp of MHC-IIB exon 2 sequence

Exon 2 sequence	Restriction enzyme		
	PstI	RsaI	SalI
type I	279bp	124bp, 54bp, 101bp	279bp
type II	279bp	279 bp	279bp
type III	279bp	124bp, 54bp, 101bp	201bp, 78bp
type IV	168bp, 111bp	279 bp	279bp

Genomic Structure of MHC Class II Region

To determine the genomic structure of the *N. nippon* MHC class II region, we screened a genomic library constructed from the liver of a dead female and assembled two contigs (Figure 1). Restriction mapping and partial sequencing suggested that the contigs represented two different homologous chromosomes. One contig (approximately 35 kb) contained a partial copy of *COL11A2*, one copy of an MHC IIA/IIB pair (*DAA1*01*/*DAB1*01*), and *BRD2*; this contig was designated haplotype 1

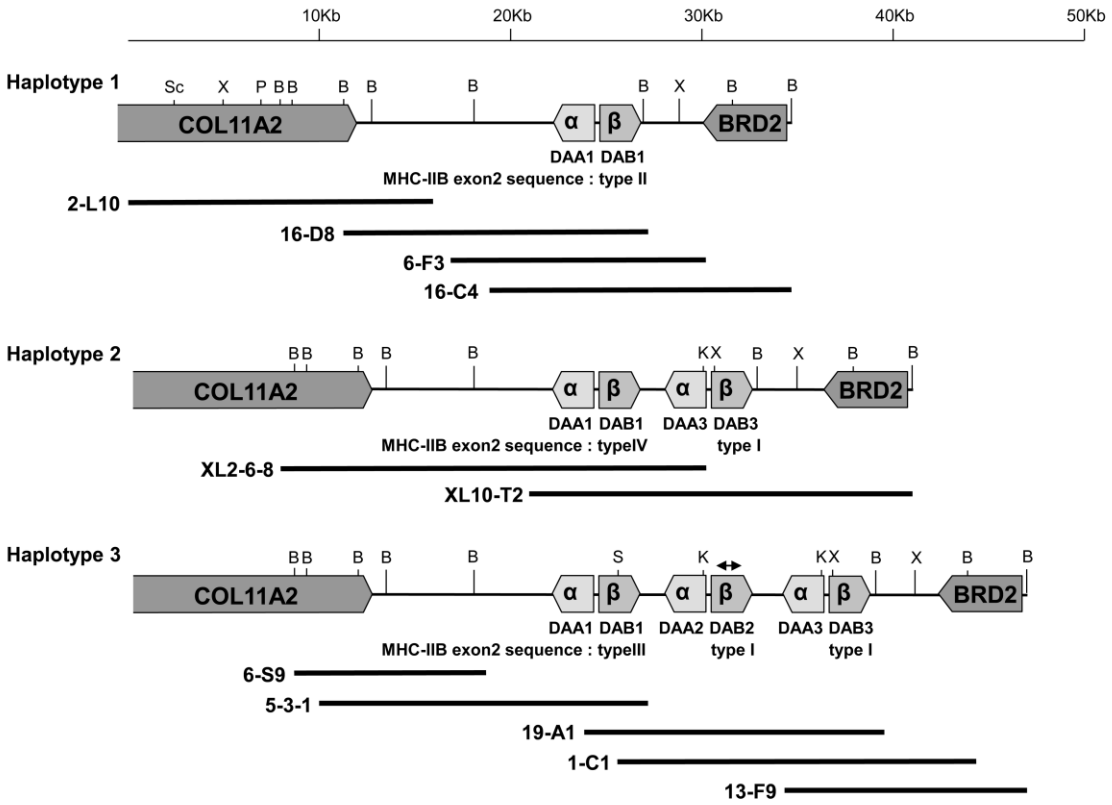

Figure 1. Genomic organization of the Japanese crested ibis MHC class II region. Three contigs representing different haplotypes were constructed. *Collagen-type XI α-2* (*COL11A2*), MHC-IIA (*α*), MHC-IIB (*β*) and *bromodomain-containing 2* (*BRD2*) genes and their orientations are indicated. Locus names are indicated below the MHC-IIA and -IIB genes. The types of MHC-IIB exon 2 sequences (Table 1) are shown below the locus names. B, K, P, S, Sc, and X represent restriction sites used for subcloning of *Bam*HI, *Kpn*I, *Pst*I, *Sal*I, *Sac*I, and *Xho*I, respectively. Solid bars below the map represent locations of isolated lambda phage clones. The bidirectional arrow above *DAB2* in haplotype 3 indicates the first amplified MHC-IIB fragment.

(HP1). (Note: the *DAA1/DAB1*-nomenclature describes an allele of the full gene.) PCR-RFLP analysis of the isolated clone revealed that the exon 2 sequence of the *DAB1*01* allele in HP1 was type II. Another contig (approximately 40 kb) contained a partial copy of *COL11A2*, three copies of MHC IIA/IIB pairs (*DAA1*03*/*DAB1*03*, *DAA2*03*/*DAB2*03*, and *DAA3*03*/*DAB3*03*), and *BRD2*; this contig was designated haplotype 3 (HP3). The exon 2 sequences of the *DAB1*03*, *DAB2*03*, and *DAB3*03*

alleles in HP3 were types III, I, and I, respectively.

To isolate a haplotype containing type IV of MHC-IIB exon 2, an additional genomic library was constructed from founder E, in which type IV was revealed by the sequencing of exon 2 PCR products (Table S2). Two positive clones were isolated and a contig (approximately 35 kb) was assembled and designated as haplotype 2 (HP2). HP2 contained a partial copy of *COL11A2*, two copies of MHC IIA/IIB pairs (*DAA1*02*/*DAB1*02* and *DAA3*02*/*DAB3*02*), and *BRD2* (Figure 1). The exon 2 sequences of *DAB1*02* and *DAB3*02* alleles in HP2 were types IV and I, respectively. The names of the three MHC IIA/IIB pairs in HP3 were assigned according to the proposal for naming vertebrate MHC genes suggested by Klein et al. (1990). Names of MHC-IIA and -IIB genes in HP1 and HP2 were assigned based on homology with genes in HP3. MHC-IIA and -IIB genes were arranged head-to-head in all pairs.

To characterize the MHC-IIA and -IIB genes in detail, the complete sequences of *Bam*HI fragments (approximately 8, 13, and 18 kb from HP1, HP2, and HP3, respectively) were determined. The coding sequences in MHC-IIA and -IIB genes were predicted using GENSCAN and deduced amino acid sequences were then aligned (Figures 2 and 3). Frameshift mutations or premature stop codons were not detected in predicted MHC-IIA and -IIB genes, suggesting that every MHC class II gene could encode functional proteins. In the MHC-IIA genes, the *DAA1*01*, *DAA1*02*, and *DAA1*03* alleles encoded identical proteins. Another three alleles (*DAA3*02*, *DAA2*03*, and *DAA3*03*) also encoded identical proteins. No single-nucleotide polymorphism (SNP) sites were observed in exons within either of the two gene groups whereas two SNP sites were detected in intron 1 of the *DAA1* locus and one SNP site was found in the promoter region of the *DAA3-DAB3* locus. In the MHC-IIB genes, the *DAB3*02*

and *DAB3*03* alleles encoded identical proteins. Although the *DAB2*03* allele was very similar to *DAB3* genes, one amino acid substitution was detected in exon 1 and seven were detected in exon 3. The *DAB1* loci (*DAB1*01–*03*) were highly conserved; 12 polymorphic amino acid residues were detected only within exon 2.

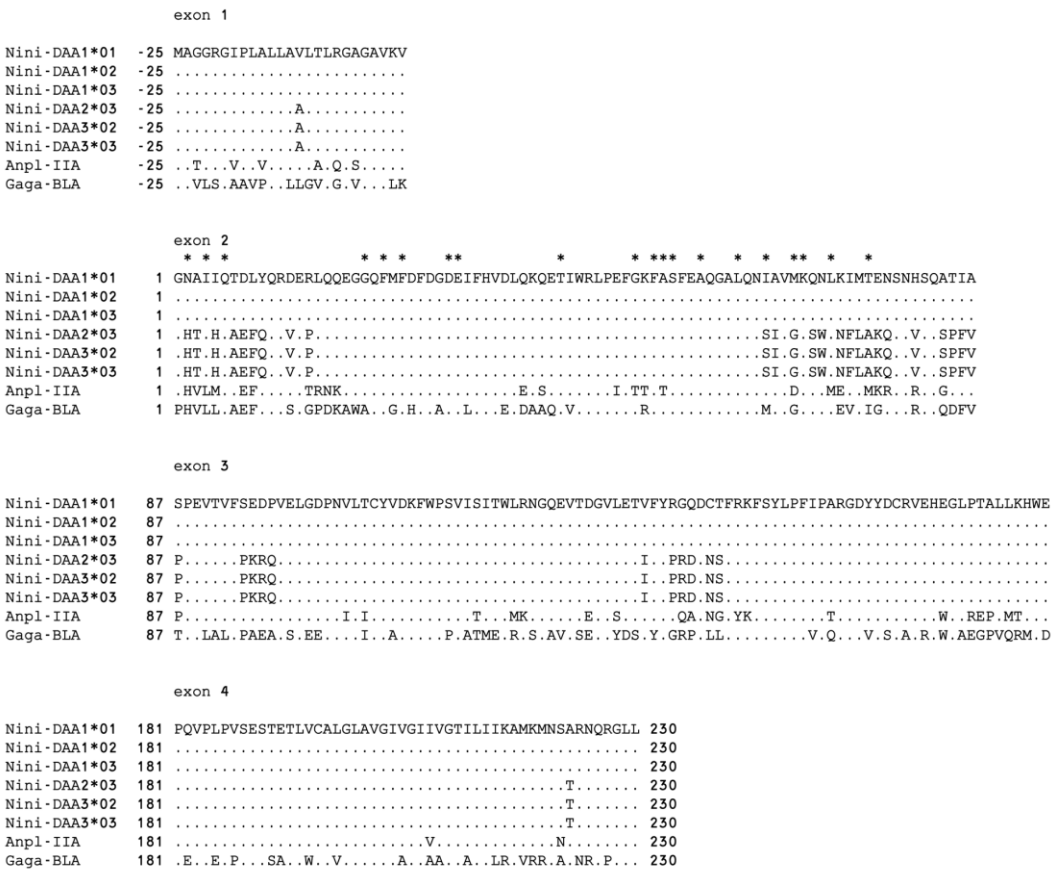

Figure 2. Alignment of predicted amino acid sequences of six MHC class IIA alleles from three haplotypes. The duck (Anpl-IIa) and chicken (Gaga-BLA) amino acid sequences are provided for reference. The first amino acid of the α1 domain was designated as position 1. Identity with the *Nini-DAA1*01* sequence is indicated with a dot. Gaps are indicated by dashes. Asterisks above the sequence of the α1 domain indicate peptide-binding residues (Stern et al., 1994).

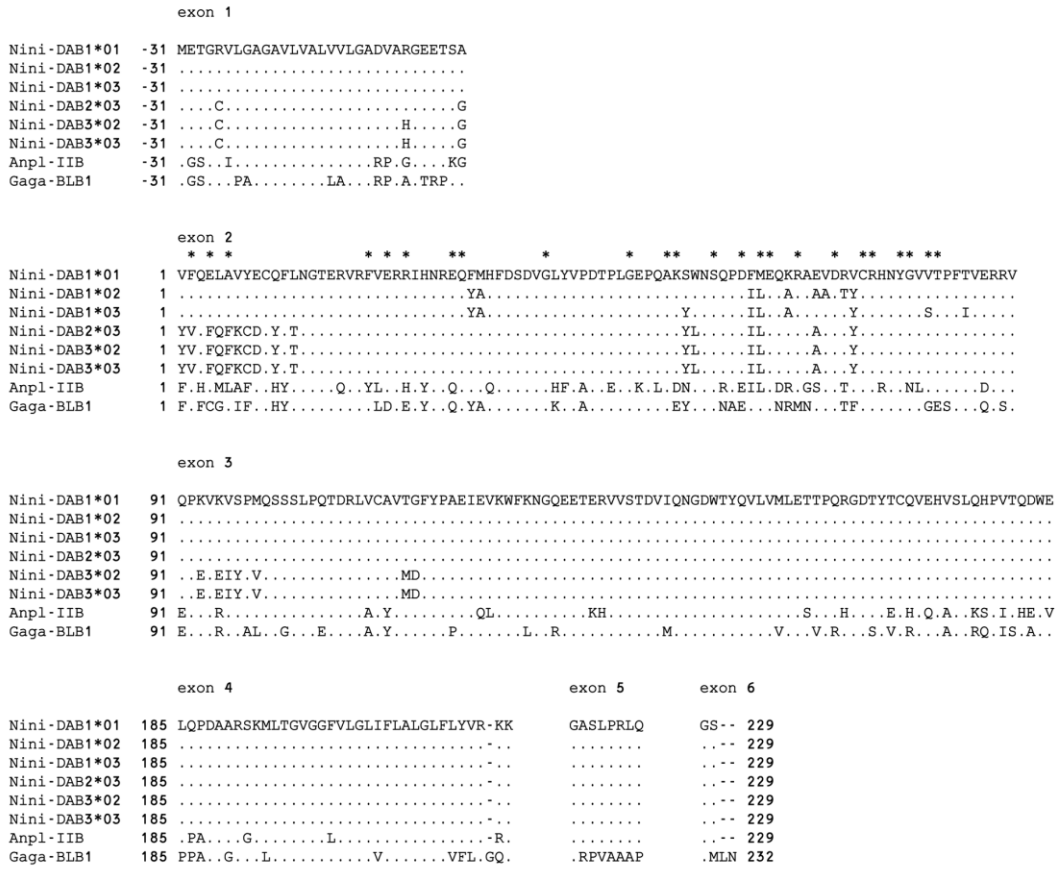

Figure 3. Alignment of predicted amino acid sequences of six MHC class IIB alleles from three haplotypes. Duck (Anpl-IIb) and chicken (Gaga-BLB1) amino acid sequences are provided for reference. The first amino acid of the β1 domain was designated as position 1. Identity with the *Nini-DAB1*01* sequence is indicated with a dot. Gaps are indicated by dashes. Asterisks above the sequence of the β1 domain indicate peptide-binding residues (Brown et al., 1993).

A dot-matrix analysis of exons and introns revealed that the *DAA2-DAB2* locus in HP3 was more similar to the *DAA3-DAB3* locus than the *DAA1-DAB1* locus (Figure 4). A remarkable feature was observed in the *DAB1*01* allele in HP1: its 3′-terminal region was identical to that of the *DAB3* locus, whereas its 5′-terminal region was identical to that of the *DAB1* locus.

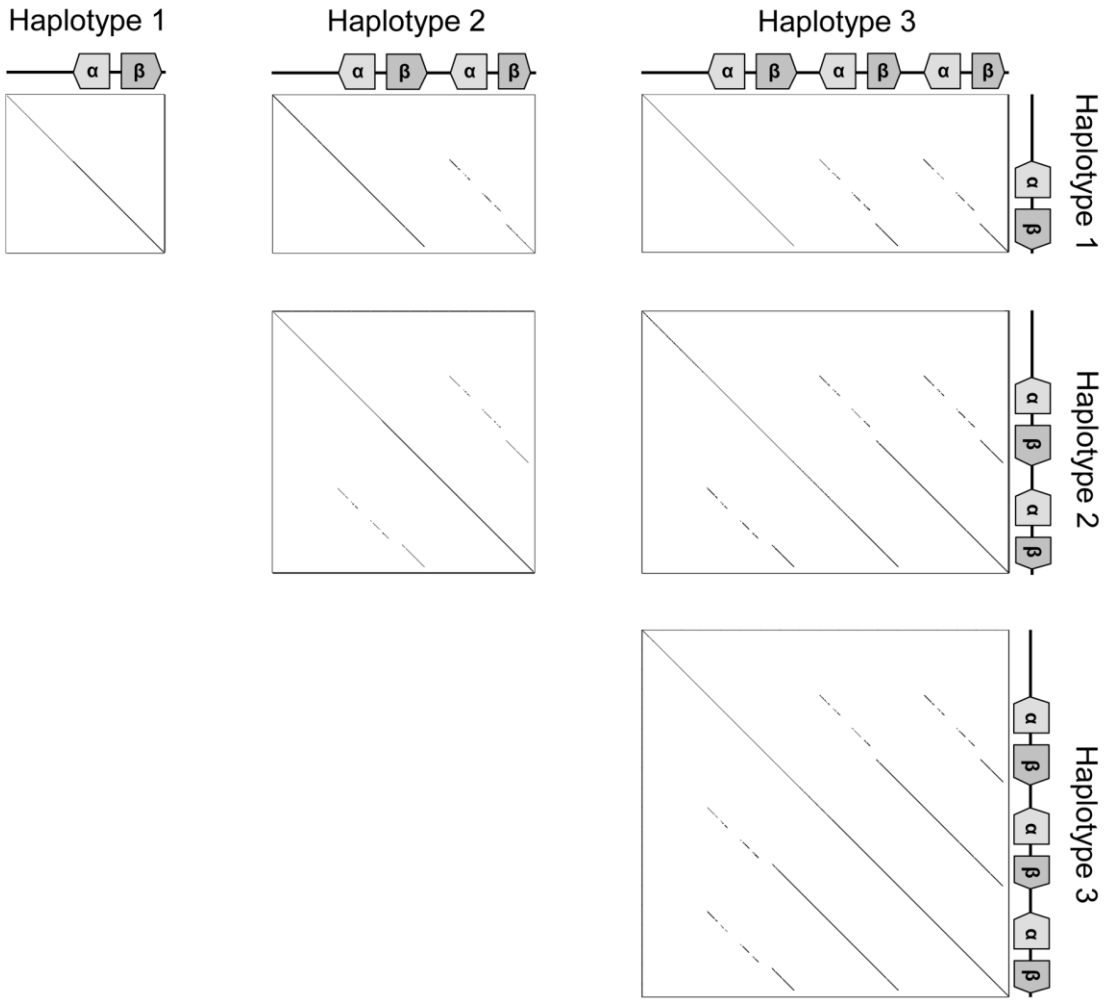

Figure 4. Dot-matrix analysis between three haplotypes of MHC class II region. Diagonal lines indicate regions where contiguous sequences align.

Phylogenetic Analysis of the Japanese Crested Ibis MHC Class IIB Gene

Phylogenetic relationships of MHC-IIB genes within 11 Pelecaniformes as well as among eight avian orders were analyzed using exon 2 or partial exon 3 sequences. The maximum-likelihood tree constructed using MHC-IIB exon 2 sequences showed that all Pelecaniformes formed one cluster separated from other avian orders (Figure 5a). The Japanese crested ibis belongs to family Threskiornithidae; the other 10 Pelecaniformes belong to family Ardeidae. In the Ardeidae, two MHC-IIB loci, *DAB1* and *DAB2*, were

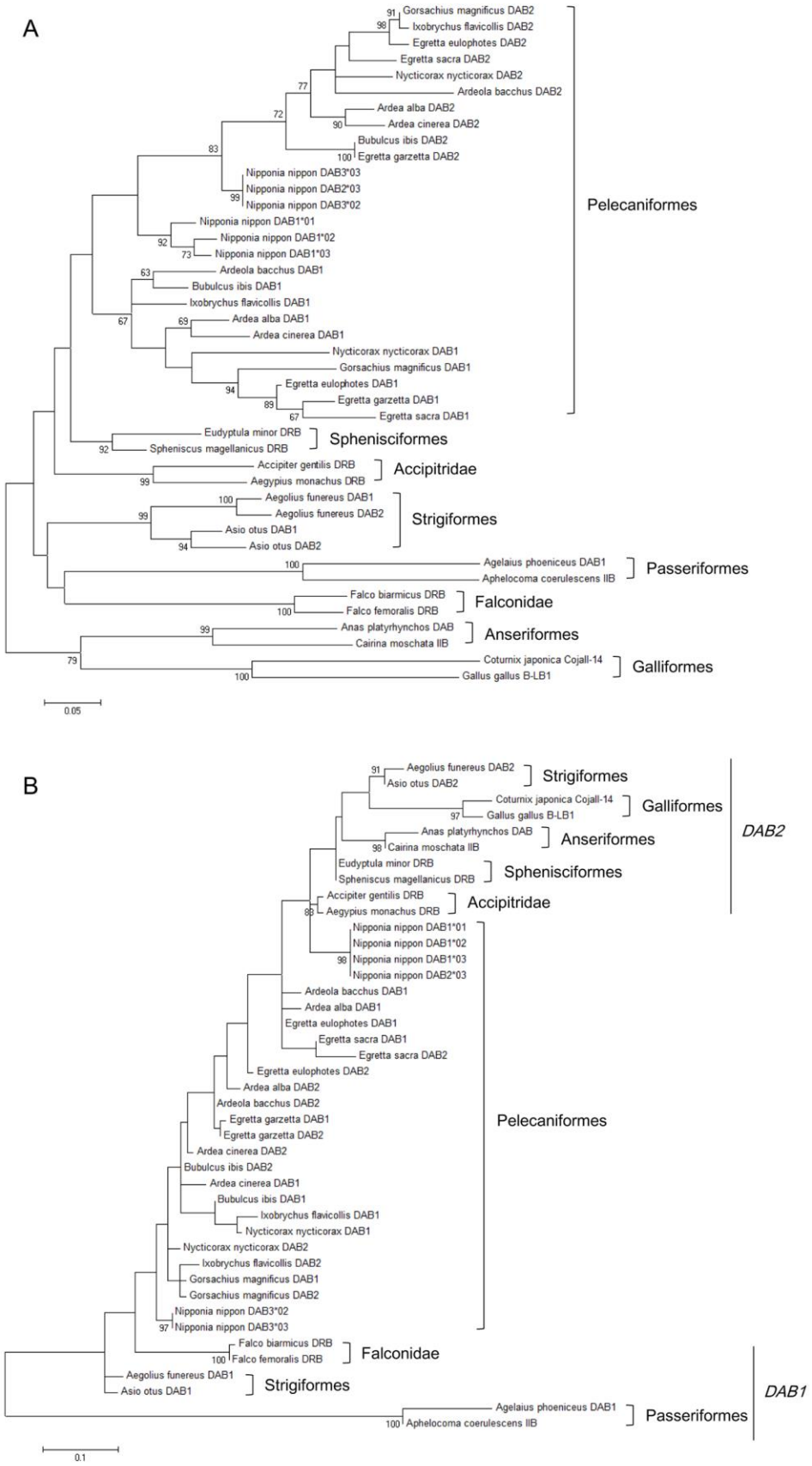
A
Gorsachius magnificus DAB2
Ixobrychus flavicollis DAB2
Egretta eulophotes DAB2
Egretta sacra DAB2
Nycticorax nycticorax DAB2
Ardeola bacchus DAB2
Ardea alba DAB2
Ardea cinerea DAB2
Bubulcus ibis DAB2
Egretta garzetta DAB2
Nipponia nippon DAB3*03
Nipponia nippon DAB2*03
Nipponia nippon DAB3*02
Nipponia nippon DAB1*01
Nipponia nippon DAB1*02
Nipponia nippon DAB1*03
Ardeola bacchus DAB1
Bubulcus ibis DAB1
Ixobrychus flavicollis DAB1
Ardea alba DAB1
Ardea cinerea DAB1
Nycticorax nycticorax DAB1
Gorsachius magnificus DAB1
Egretta eulophotes DAB1
Egretta garzetta DAB1
Egretta sacra DAB1
Pelecaniformes
Eudyptula minor DRB
Spheniscus magellanicus DRB
Sphenisciformes
Accipiter gentilis DRB
Aegypius monachus DRB
Accipitridae
Aegolius funereus DAB1
Aegolius funereus DAB2
Asio otus DAB1
Asio otus DAB2
Strigiformes
Agelaius phoeniceus DAB1
Aphelocoma coerulescens IIB
Passeriformes
Falco biarmicus DRB
Falco femoralis DRB
Falconidae
Anas platyrhynchos DAB
Cairina moschata IIB
Anseriformes
Coturnix japonica Cojall-14
Gallus gallus B-LB1
Galliformes
0.05
B
Aegolius funereus DAB2
Asio otus DAB2
Strigiformes
Coturnix japonica Cojall-14
Gallus gallus B-LB1
Galliformes
Anas platyrhynchos DAB
Cairina moschata IIB
Anseriformes
Eudyptula minor DRB
Spheniscus magellanicus DRB
Sphenisciformes
Accipiter gentilis DRB
Aegypius monachus DRB
Accipitridae
DAB2
Nipponia nippon DAB1*01
Nipponia nippon DAB1*02
Nipponia nippon DAB1*03
Nipponia nippon DAB2*03
Ardeola bacchus DAB1
Ardea alba DAB1
Egretta eulophotes DAB1
Egretta sacra DAB1
Egretta sacra DAB2
Egretta eulophotes DAB2
Ardea alba DAB2
Ardeola bacchus DAB2
Egretta garzetta DAB1
Egretta garzetta DAB2
Ardea cinerea DAB2
Bubulcus ibis DAB2
Ardea cinerea DAB1
Bubulcus ibis DAB1
Ixobrychus flavicollis DAB1
Nycticorax nycticorax DAB1
Nycticorax nycticorax DAB2
Ixobrychus flavicollis DAB2
Gorsachius magnificus DAB1
Gorsachius magnificus DAB2
Nipponia nippon DAB3*02
Nipponia nippon DAB3*03
Pelecaniformes
Falco biarmicus DRB
Falco femoralis DRB
Falconidae
Aegolius funereus DAB1
Asio otus DAB1
Strigiformes
Agelaius phoeniceus DAB1
Aphelocoma coerulescens IIB
Passeriformes
DAB1
0.1

Figure 5. Maximum-likelihood tree with MHC-IIB exon 2 or partial exon 3 sequences from *Nipponia nippon* and other bird species. The best-fitting nucleotide substitution model for each codon position was evaluated using Find Best DNA/Protein Models (ML) in MEGA version 5.2 (Tamura et al., 2011) according to the Akaike information criterion. (A) The tree of exon 2 was constructed by using a Kimura 2-parameter model with gamma distribution in MEGA. (B) The tree of partial exon 3 was constructed by using a Tamura 3-parameter model with gamma distribution in MEGA. Bird species used for the analyses were shown in Table S3. In both analyses, bootstrap values were evaluated with 1000 replications. Bootstrap values >60 are shown in this tree. Branch lengths represent the number of changes per site.

identified on the basis of differences in the length of intron 1 (Li et al., 2011). Within the Pelecaniformes cluster, three alleles (*DAB3*02*, *DAB2*03*, and *DAB3*03*) of the Japanese crested ibis were included in the *DAB2* subclade. Although the other three *DAB* alleles (*DAB1*01–*03*) were grouped into independent subclade, this branching was ambiguous because of a low bootstrap value (< 60). The maximum-likelihood tree with partial exon 3 sequences was largely different from that with exon 2 sequences (Figure 5b). Among seven avian orders except for Pelecaniformes, two MHC-IIB lineages (*DAB1* and *DAB2*) were observed (Burri et al., 2010). Within Pelecaniformes Ardeidae, partial exon 3 sequences tended to cluster together within species and did not show gene-specific cluster (Li et al., 2011). In contrast, four alleles (*DAB1*01–*03* and *DAB2*03*) and two alleles (*DAB3*02* and *DAB3*03*) of the Japanese crested ibis were grouped into two independent subclades separated from other Pelecaniformes. Phylogenetic relationships with MHC-IIB exon 2 and partial exon 3 sequences suggested that evolution of the MHC-IIB gene in *N. nippon* might be different from other Pelecaniformes.

Genetic Diversity of MHC Class II Region among 5 Founders

The current *N. nippon* population in Japan originates from only five founders. To investigate the genetic diversity of the MHC class II region among these founders, their MHC genotypes were examined by PCR-RFLP (Figure 6). Non-digested fragments with *Rsa*I, *Pst*I-digested fragments, and *Sal*I-digested fragments represented type II (*DAB1*01* in HP1), type IV (*DAB1*02* in HP2), and type III (*DAB1*03* in HP3), respectively (Table 1 and Figure 1). The genotypes of founders A, B, C, D, and E were estimated to be *hp1/2*, *hp1/3*, *hp1/3*, *hp1/1*, and *hp2/3*, respectively.

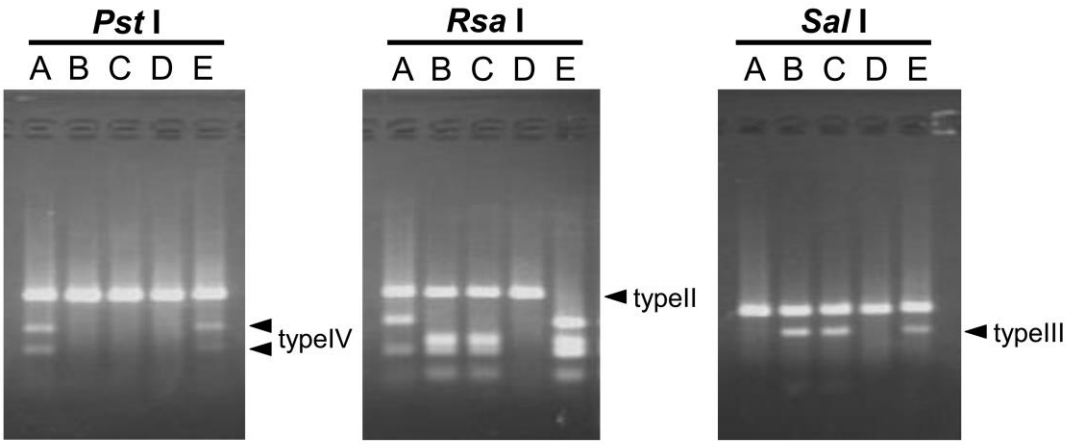

Figure 6. Genotyping of MHC class II gene regions by PCR-RFLP. A 279-bp fragment of MHC IIB exon 2 was amplified from five founder genomes (A–E). PCR products were digested with *Pst*I, *Rsa*I, or *Sal*I and digested fragments were analyzed by 3% agarose gel. Non-digested fragments with *Rsa*I, *Pst*I-digested fragments, and *Sal*I-digested fragments represented type II (*DAB1*01* in HP1), type IV (*DAB1*02* in HP2), and type III (*DAB1*03* in HP3), respectively (Table 1 and Figure 1).

The MHC genotypes of the founders were also analyzed by Southern blotting (Figure 7). Three bands of approximately 8, 13, and 18 kb were detected that corresponded to HP1, HP2, and HP3, respectively. The genotypes indicated by Southern blotting were the same as those determined by PCR-RFLP.

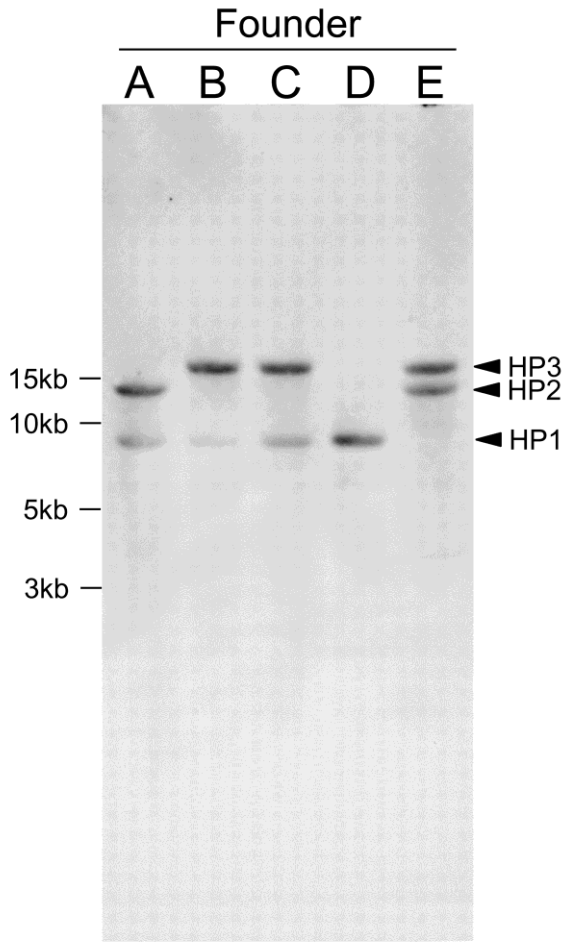

Figure 7. Detection of MHC class II gene regions by Southern blotting. Genomic DNA from five founders (A-E) was digested with *Bam*HI and *Eco*RI and hybridized with a mixture of three probes (MHC IIA exon 3, MHC IIB exon 2, and MHC IIB exon 3). Three bands of approximately 8, 13, and 18 kb represented HP1, HP2, and HP3, respectively.

In four founder genomes (A–D), both of the alleles at the *DAB1* locus were individually amplified. Primers DAA-F1 and DAB*01-R1 were used for *DAB*01*, and primers DAA-F1 and DAB*02,03-R1 were used for *DAB1*02* and *DAB1*03*. The *DAB1*02* and *DAB1*03* alleles in founder E were amplified from isolated lambda phage clones. The resulting PCR products were directly sequenced and sequences (approximately 1 kb) from exon 2 to exon 4 were compared. Four *DAB1*01* sequences

from founders A–D were completely identical, as were two *DAB1*02* sequences from founders A and E and three *DAB1*03* sequences from founders B, C, and E. These results strongly suggested that the founder population possessed only three MHC class II haplotypes. Moreover, no additional hybridizing bands besides the three haplotype bands were detected in the Southern blotting, suggesting that the MHC class II region was a single locus in the Japanese crested ibis genome.

Discussion

The Japanese crested ibis is a critically threatened and internationally conserved species belonging to order Pelecaniformes, an avian lineage that is highly divergent from order Galliformes. We isolated genomic clones encompassing MHC-IIB genes from the *N. nippon* genomic libraries and constructed three contigs covering the MHC class II regions by genome walking (Figure 1). These contigs represented 3 different haplotypes of MHC class II regions. Our sequencing data revealed that the MHC class II genomic structure in *N. nippon* was largely different from that of chicken. The MHC-IIB gene was flanked by the MHC-IIA gene and the MHC IIA/IIB gene pair was located between *COL11A2* and *BRD2* in *N. nippon*, whereas the chicken MHC-B contained no MHC-IIA gene (*BLA*) and two MHC-IIB genes (*BLB1* and *BLB2*) located on both sides of the *Tapasin* gene (Kaufman et al., 1999). We did not detect a *Tapasin*-like gene in cloned fragments. The gene order *COL11A2*-MHC-IIA, MHC-IIB-*BRD2* in *N. nippon* was more similar to that in humans than to that in chicken (Kaufman et al., 1999; Horton et al., 2004). In the duck MHC, a single MHC-IIA gene was located next to five MHC-IIB genes, whereas the genes on both sides of the MHC IIA/IIB cluster are unknown (Ren et al., 2011). In contrast with the duck MHC, in

which only the MHC-IIB gene was duplicated, gene duplications occurred as a unit with MHC-IIA/IIB gene pairs in *N. nippon*.

Our results for polymorphism of MHC-IIB exon 2 (Table 1), cloning (Figure 1), and Southern blotting (Figure 7) suggested that the MHC class II region was a single locus in the Japanese crested ibis genome, although we cannot exclude the possible existence of other MHC class II loci with low sequence homology to those presented here. These results are similar to findings in duck (Ren et al., 2011) but different from those in chicken, in which additional MHC class II genes were found in the MHC-Y region (Miller et al., 1994; Delany et al., 2009). These combined results suggest that there is large variability in MHC genomic organization among avian species at the order level.

The three haplotypes of MHC class II regions in *N. nippon* had different copy numbers of MHC IIA/IIB gene pairs (Figure 1). In MHC-IIA genes, three alleles (*DAA1*01–*03*) at the *DAA1* locus encoded identical proteins, as did another three alleles (*DAA3*02*, *DAA2*03*, and *DAA3*03*) (Figure 2). However, three MHC-IIB alleles (*DAB1*01–03*) contained 12 polymorphic amino acid residues within exon 2 (Figure 3). MHC-IIB genes were apparently more polymorphic than MHC-IIA genes. In contrast with the *DAB1* locus, the alleles *DAB3*02*, *DAB2*03,* and *DAB3*03* had an identical exon 2. Haplotypes 2 and 3 might have been produced by duplication and/or gene conversion during a relatively recent period in evolutionary time. Moreover, the characteristic of *DAB1*01*, in which its 5′- and 3′-terminal regions were identical to those of *DAB1* and *DAB3* loci, respectively (Figure 4), might indicate that *DAB1*01* was produced by recombination between the *DAB1* and *DAB3* loci. Copy number variations of MHC-IIB gene has been reported in certain bird species (Hosomichi et al., 2006; Eimes et al., 2011; Strandh et al., 2012; Alcaide et al., 2014). These results might

indicate that gene duplication, gene loss and/or gene conversion are operating at relatively high rate within a single bird species

Our phylogenetic analyses showed that Pelecaniformes formed one cluster separated from other seven avian orders (Figure 5a), suggesting that MHC-IIB genes of *N. nippon* were closely related to those of the other Pelecaniformes from family Ardeidae that were examined in this study. Two MHC-IIB loci, *DAB1* and *DAB2*, with longer and shorter intron 1 lengths, respectively, were detected in Ardeidae (Li et al., 2011). The three *DAB1* alleles in *N. nippon* possessed a longer intron 1 (661 bp) than did *DAB3*02*, *DAB2*03*, and *DAB3*03* (285 bp). The latter three alleles were grouped into the *DAB2* subclade, but *DAB1*01–*03* were not in the *DAB1* subclade in the maximum-likelihood tree with exon 2 sequences , suggesting that evolution of the MHC-IIB gene might differ among families within the same order. Moreover, the relationship between two MHC-IIB loci (*DAB1* and *DAB2*) in Pelecaniformes Ardeidae was obviously different from that in Strigiformes (Figure 5). Although the maximum-likelihood tree with partial exon 3 sequences suggested that the *DAB3* and *DAB1* loci in *N. nippon* might represent two MHC-IIB lineages (*DAB1* and *DAB2*) respectively (Figure 5b), the genome structures emerged that three *DAB* loci might be a cluster in the same lineage (Figure 1). To elucidate the long-term evolutionary history of the avian MHC, more data on genome structure of MHC from a wide diversity of bird species are apparently needed.

As the current Japanese population of *N. nippon* originates from a small number of founders, polymorphism among these founders should largely limit the genetic diversity of progeny. We identified only three kinds of MHC class II haplotypes among the five founders. In the Chinese population of *N. nippon*, five alleles of MHC-IIB exon 2 were detected (Zhang et al., 2006). Three out of five alleles were not observed

among the founder birds in Japan. Microsatellite markers in the Chinese population contained two to five alleles, and only two haplotypes were detected in mitochondrial DNA control regions (Ji et al., 2004; He et al., 2006; He et al., 2013). These findings suggested that the Chinese population experienced a severe evolutionary bottleneck. There were two to three alleles of microsatellite markers and haplotype numbers of 202-bp sequences containing multiple SNP sites in the five founders of the Japanese population (Urano et al., 2013; Chapter 1). These results strongly suggested that the number of alleles/haplotypes in the current Japanese population are three or fewer in most genomic loci and that genetic diversity is extremely low. The finding of lower genetic diversity in the Japanese population compared with the Chinese population is reasonable because the founders of the Japanese population originated from China.

To our knowledge, this is the first report of MHC class II genomic organization in Pelecaniformes. Our results revealed that the MHC class II genomic structure of the Japanese crested ibis was largely different from that of Galliformes or Anseriformes. Moreover, the fact that five founders possessed only three kinds of MHC class II haplotypes strongly suggested that the genetic diversity of the MHC region in the Japanese population is extremely low. The recovery of a large population from a small number of founders with low genetic diversity will be a significant challenge. If successful, the process might provide a good model for investigating the expansion of genetic diversity in a closed population through mutations and recombination events. The structure of the MHC class II region presented here will provide valuable insights for future studies on the evolution of the avian MHC and for conservation of the Japanese crested ibis.

Supporting information

Table S1. List of primers used for polymerase chain reaction analysis

Table S2. Polymorphism of MHC-IIB exon 2 in two founders and 20 progeny of Japanese crested ibis

Table S3. Bird species used in phylogenetic analysis

Figure S1. Schematic of polymerase chain reaction (PCR) screening for the lambda phage library

Figure S2. Partial sequence of *collagen-type XI α-2-like* gene in the *Nipponia nippon* MHC class II region

Table S1.List of primers used for PCR

Primer name	Sequence 5'-3'	Annealing temperature (°C)	Product size (bp)	Usage	Reference
BRMHC05	CGTRCTGGTGGCACTGGTGGYGCT	68	1,057	cloning	Miller & Lambert (2004)
AIEx3R	CACCAGCASCTGGTASGTCCAGTC				Alcaide et al. (2007)
2F_pen1	AACGGCACCGAGCGGGTGAGGT	68	279	amplification of exon 2, screening, probe	Tsuda et al. (2001)
intron2-01R	GGGCTTGGCTTGTGCCCATCC				
IIAex3-F	TCACCCCAGTGCTTCAGGAG	58	238	probe	
IIAex3-R	ACCCCAACGTCCTGACCTGC				
IIBex3-F	TCGCTCTCCCCCAGTTCAGC	58	307	probe	
IIBex3-R	CAGGGCCTTACCCCAGTCCT				
DAA-F1	CCATGCAGACCCTGGGACCTTGA	58		specific amplification for DAB1 locus	
DAB*01-R1	GGAGGGATCCCCAAGCACGGAAC		2,999	specific amplification for DAB1*01 allele	
DAB*02,03-R1	ACAACCAGGACCGGAGGTTACCG		3,109	specific amplification for DAB1*02, 03 alleles	
16-D8Rev_S1	CCTAATAAACACGCGTCAAG	52	436	genome walking	
16-D8Rev_A1	TGGGTTTGACTCCGTTTC				
16-D8Uni_S1	GTGCTCTTTCCCAGCATTCC	58	432	genome walking	
16-D8Uni_A1	GTCAGGCTGGACCAGACTAC				
1-C1BRD2_S1	GAGTGCAGCGACTTGAGGA	58	355	genome walking	
1-C1BRD2_A1	CAGATGCCACCAGAAGAGCA				

Table S2. Polymorphism of MHC-IIB exon 2 in two founders and 20 progeny of the Japanese Crested Ibis

ID of individual	Detected exon2 sequences
51, 55, 61, 75	type II, IV
67, 81, 130, 116, founder E	type I, III, IV
69, 113, 118,131, 133	type I, III
83, 99, 102, 121	type I, II, III
119, 122, founder D	type II
143	type I, II, IV

Table S3 Bird species used in phylogenetic analysis				
Order/Family	Species	Accession Number	Gene Nomenclature	Reference
Pelecaniformes/Threskiornithidae	*Nipponia Nippon*	AB872442-AB872444	*Nini*-DAB1*01-*03, DAB2*03, DAB3*02-*03	
Pelecaniformes/Ardeidae	*Ardea cinerea*	HM991018, HM991041	*Arci*-DAB1, -DAB2	Li et al. (2011)
	Egretta garzetta	HM991035, HM991056	*Egga*-DAB1, -DAB2	
	Egretta eulophotes	HM991028, HM991052	*Egeu*-DAB1, -DAB2	
	Egretta sacra	HM991037, HM991058	*Aral*-DAB1, -DAB2	
	Bubulcus ibis	HM991033, HM991054	*Buib*-DAB1, -DAB2	
	Ardeola bacchus	HM991020, HM991044	*Arba*-DAB1, -DAB2	
	Ardea alba	HM991022, HM991046	*Aral*-DAB1, -DAB2	
	Nycticorax nycticorax	HM991039, HM991060	*Nyny*-DAB1, -DAB2	
	Gorsachius magnificus	HM991024, HM991048	*Goma*-DAB1, -DAB2	
	Ixobrychus flavicollis	HM991026, HM991050	*Ixfl*-DAB1, -DAB2	
Galliformes	*Gallus gallus*	AL023516	B-LB1	Kaufman et al. (1999)
	Coturnix japonica	AB265805	CojaII-14	Shiina et al. (2004)*
Anseriformes	*Anas platyrhynchos*	AF390589	—	Chan SWS, et al. (unpublished)
	Cairina moschata	DQ490138	—	Mannes NK and Schulz U (unpublished)
Spheniscifformes	*Eudyptula minor*	AB302187	—	Tsuda et al. (2001)
	Spheniscus magellanicus	AB325529	—	Kikkawa et al. (2009)**
Accipitriformes	*Accipiter gentilis*	EF370953	*Acge*-DRB	Alcaide et al. (2007)
	Aegypius monachus	EF370954	*Aemo*-DRB	
Strigiformes	*Aegolius funereus*	EF641252, EF641253	*Aefu*-DAB1, -DAB2	Burri et al. (2008)
	Asio flammeus	EF641250, EF641251	*Asfl*-DAB1, -DAB2	
Falconiformes	*Falco biarmicus*	EF370989	*Fabi*-DRB	Alcaide et al. (2007)
	Falco femoralis	EF370988	*Fafe*-DRB	
Passeriformes	*Agelaius phoeniceus*	AF030997	*Agph*-DAB1	Edwards et al. (1998)***
	Aphelocoma coerulescens	U23958	—	Edwards et al. (1995)****

* Shiina T, Shimizu S, Hosomichi K, Kohara S, Watanabe S, et al. (2004) Comparative genomic analysis of two avian (quail and chicken) MHC regions. J Immunol 172: 6751-6763.

** Kikkawa EF, Tsuda TT, Sumiyama D, Naruse TK, Fukuda M, et al. (2009) Trans-species polymorphism of the Mhc class II DRB-like gene in banded penguins (genus Spheniscus). Immunogenetics 61: 341-352.

*** Edwards SV, Gasper J, March M (1998) Genomics and polymorphism of Agph-DAB1, an Mhc class II B gene in red-winged blackbirds (Agelaius phoeniceus). Mol Biol Evol 15: 236-250.

**** Edwards SV, Wakeland EK, Potts WK (1995) Contrasting histories of avian and mammalian Mhc genes revealed by class II B sequences from songbirds. Proc Natl Acad Sci U S A 92: 12200-12204.

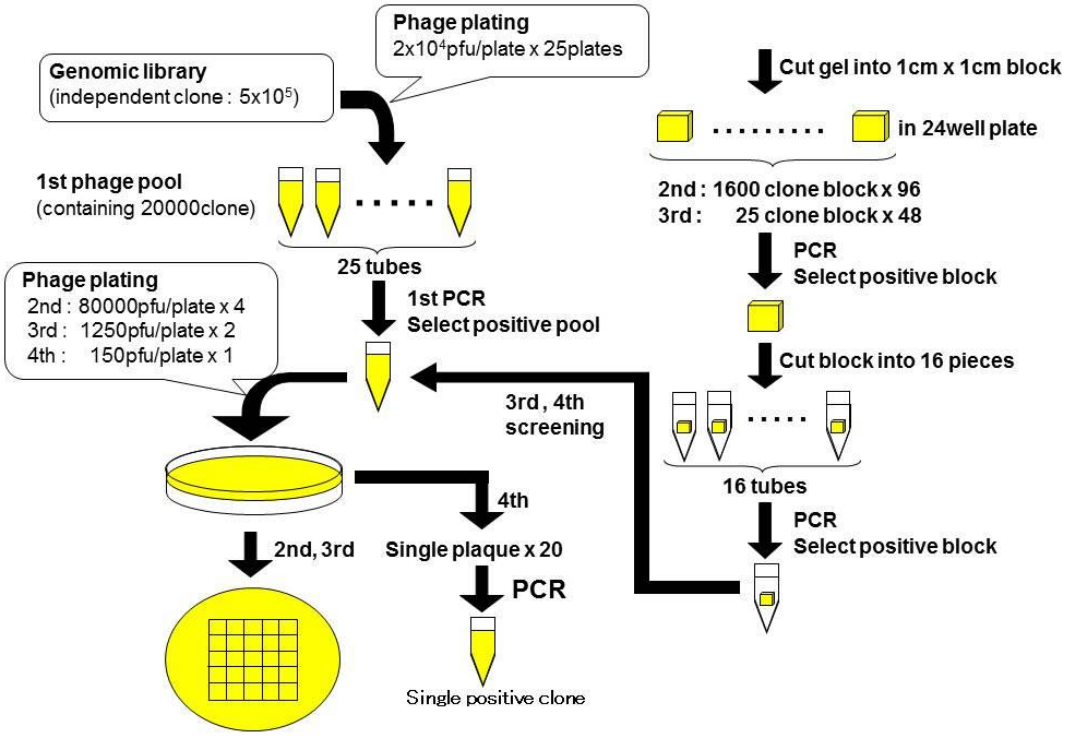

Figure S1. Schematic of polymerase chain reaction (PCR) screening for the lambda phage library. In the first screening, lambda phages from the primary library were plated on 90-mm plates at 20,000 pfu/plate. Twenty-five plates were used for the screening of 500,000 independent clones. After clear plaques appeared, the plates were overlaid with 3 mL of SM buffer and stored at 4 ℃ overnight. Next, phage solutions were individually collected into 25 tubes (first phage pool containing 20,000 clones). Phage solutions were directly used as template DNA for PCR. The positive first phage pool was selected by PCR using KOD-FX Neo DNA polymerase (Toyobo). In the second screening, phages from the positive pool were plated on four plates at 80,000 pfu/plate. After clear plaques appeared, NZY agarose gel with plaques was cut into 1- cm^2 blocks (containing approximately 1600 clones/block). Each block was transferred into a 24-well plate with 400 μL of SM buffer. The second screening used 96 blocks. Positive blocks were selected by PCR, further divided into 16 sections (0.0625 cm^2 containing approximately 100 clones each), and transferred into 1.5-mL tubes with 100 μL of SM buffer. Positive 0.0625 cm^2 blocks were selected by PCR. In the third screening, phages were plated at 1,250 pfu/plate (n = 2 plates). Positive 0.0625 cm^2 blocks (containing several clones) were selected in the same manner as for the second screening. In the fourth screening, phages from the positive pool were plated 150 pfu/plate (n = 1 plate). After clear plaques appeared, 20 single plaques were transferred into 1.5-mL tubes with 100 μL of SM buffer. A single positive plaque was selected by PCR.

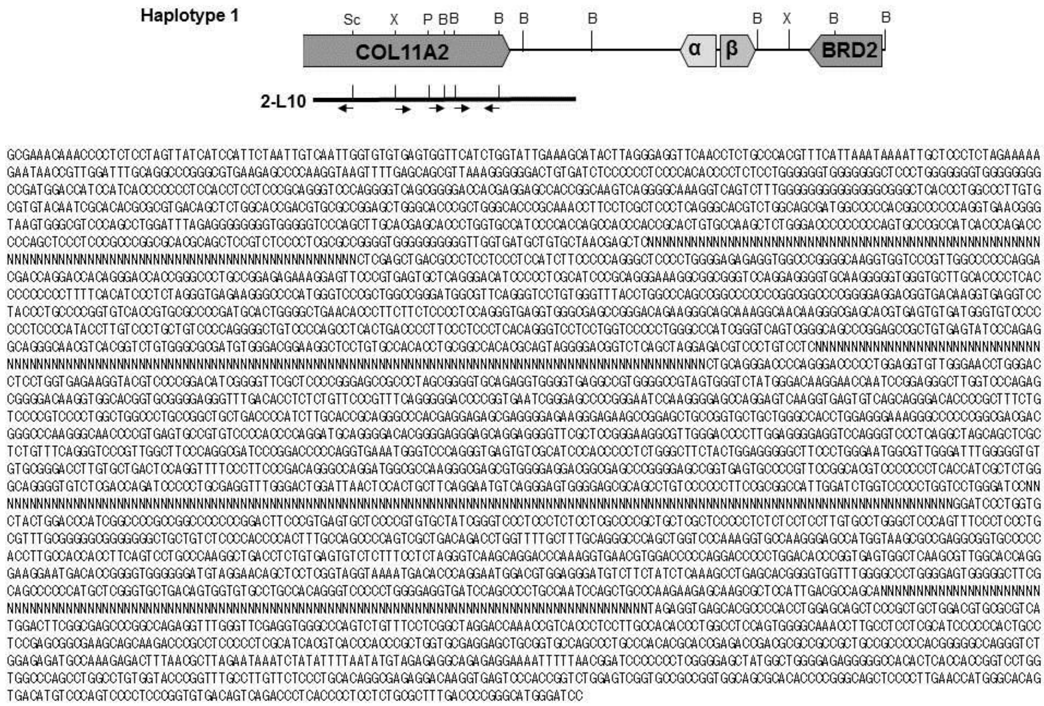

Figure S2. Partial sequence of *collagen-type XI α-2-like* gene in the *Nipponia nippon* MHC class II region. B, P, Sc, and X represent restriction sites used for subcloning of *Bam*HI, *Pst*I, *Sac*I, and *Xho*I, respectively. Solid bars and arrows below the map indicate the location of isolated lambda phage clone and sequenced regions, respectively. Gap regions are shown as "N." A BLAST search against the human genome + transcripts database revealed that this partial sequence was homologous to *collagen-type XI α-2 transcript*.

CHAPTER 3

Structure and Polymorphism of the Major Histocompatibility Complex in the Oriental Stork, *Ciconia boyciana*

Introduction

The major histocompatibility complex (MHC) is the most important region in the vertebrate genome regarding infectious disease defense and is crucial in adaptive and innate immunity. It is notable for the extremely high polymorphism of class I and II genes that are responsible for presenting pathogen-derived peptides and triggering the adaptive immune response (Klein, 1986). Polymorphisms at MHC class I and class II genes facilitate binding of a diversity of pathogens, and these evolutionary selection pressures possibly contribute to the high genetic variation in MHC loci (Zinkernagel and Doherty, 1979).

Polymorphism in the MHC is not restricted to allelic variation. The molecular evolution of the MHC involves frequent gene duplication and gene loss resulting in vast rearrangements and pronounced variation in gene number and genomic organization among organisms (Kulski et al., 2002; Kelley et al., 2005). Since the discovery of the mouse MHC in 1936, MHCs of many mammalian species have been characterized. However, avian MHCs are not well characterized. The genomic organizations of the MHC spanning class I and class II regions have been reported in only 6 Galliformes, domestic chicken (*Gallus gallus*) (Kaufman et al., 1999), Japanese quail (*Coturnix coturnix japonica*) (Hosomichi et al., 2006), domestic turkey (*Meleagris gallopavo*) (Chaves et al., 2009), golden pheasant (*Chrysolophus pictus*) (Ye et al., 2012), black grouse (*Tetrao tetrix*) (Wang et al., 2012), and greater prairie-chicken (*Tympanuchus*

cupido) (Eimes et al., 2013), and a Pelecaniformes, the Japanese crested ibis (*Nipponia nippon*) (Chen et al., 2015).

The chicken MHC (also known as MHC-B or B-complex) has a remarkable structure referred to as a "minimal essential MHC" and consists of only 19 genes spanning 92 kb (Kaufman et al., 1999). The overall MHC-B structures of 5 other galliform species are largely similar to that of chicken, whereas gene number, order, and orientation in these structures vary among the species. In contrast, the MHC class II genomic structure in *N. nippon* was largely different from that of chicken (Chen et al., 2015). For example, the MHC-IIB gene was flanked by the MHC-IIA gene, and the MHC IIA/IIB gene pair was located between *COL11A2* and *BRD2* in *N. nippon*; whereas, the chicken MHC-B contained no MHC-IIA gene (*BLA*), and 2 MHC-IIB genes (*BLB1* and *BLB2*) were located on both sides of the *Tapasin* gene (Kaufman et al., 1999). These results suggest that there is large variability in MHC genomic organization among avian species at the order level. Because taxonomic and genomic sampling of avian MHC regions is limited, it is unclear whether the minimal essential MHC represents the ancestral condition or whether it is a highly derived condition unique to Galliformes. To determine the long-term evolutionary history of the avian MHC, more data of genomic structures from other avian species are required.

The Oriental stork (*C. boyciana*, order Ciconiiformes) is a large migrant and listed as "endangered" in the 2015 International Union for Conservation of Nature Red List of Threatened Species (http://www.iucnredlist.org). Historically, the Oriental stork was found in Japan, China, Korea, and Russia, but habitat loss and overhunting caused a drastic decline in its numbers resulting in its extinction in 1986 in Japan. In 1985, captive-breeding programs were conducted at the Hyogo Park of the Oriental White

Stork, by using 6 wild birds as founders provided by Russia. Thereafter, 19 Oriental storks from Russia and some zoos in Japan such as Tama Zoological Park and Tennoji Zoo were introduced into the Hyogo Park of the Oriental White Stork and used as founders. The current size of the captive-breeding population in Japan is approximately 96 birds. A project to release *C. boyciana* into the wild was launched in 2005; and in May 2007, one chick hatched in the wild, the first of this species born in the wild in 40 years (www.stork.u-hyogo.ac.jp).

Molecular ecology studies have shown that, in addition to adaptive immune responses, the MHC genotype influences patterns of mate choice, local adaptation, and expression of sexually selected ornaments (Von Schantz et al., 1997; Ekblom et al., 2007; Hale ML et al., 2009; Roberts, 2009). Therefore, the diversity of MHCs is fundamental to the conservation of endangered species.

The Oriental White stork belongs to a different avian order from that of the chicken or Japanese crested ibis. Characterizing the MHC region in *C. boyciana* may provide valuable information about primordial avian MHCs. In addition, insight into the genetic diversity of this genome region could be vital to conservation of the *C. boyciana* population. In the present study, we investigated the genomic organization of the *C. boyciana* MHC region and polymorphisms among the 6 dominant founders of the Japanese population.

Materials and Methods

Samples

Blood samples from *C. boyciana* (6 founders, A–F) were provided by the Hyogo Park of the Oriental White Stork (Toyooka, Japan). Six founders were selected on the basis of

a larger contribution to the progeny population. Genomic DNA samples were prepared from whole blood using the Wizard Genomic DNA Purification kit (Promega) according to the manufacturer's instructions.

Primers

Primers and annealing temperatures used for the polymerase chain reaction (PCR) are shown in Table S1. Degenerate primer pairs MHC IIBexon4-F1/R1, TAP1exon7-F1/R1, and TNXB-S1/A1 for amplification of MHC-IIB, TAP1, and TNXB genes, respectively, were designed on the basis of chicken, quail, and crested ibis sequences. Other primers were designed on the basis of sequences determined in this study.

Construction of a Genomic Library and Screening

To determine the genomic structure of the *C. boyciana* MHC region, a genomic library was constructed from founder B. Genomic DNA was partially digested with *Sau*3AI and separated on a 0.5% agarose gel. Digested fragments (15–23 kb) were gel purified using the Wizard SV Gel and PCR Clean-Up system (Promega), according to the manufacturer's instructions, and ligated into the Lambda DASH II vector (Stratagene). The ligated DNA mixture was then packaged using Gigapack III Gold packaging extract (Stratagene). Screening was performed using a PCR-based method as described in chapter 2. Once a positive phage clone was isolated, sequences of both its ends were determined and then used to design primers to obtain the extended segments in the next round of screening. In some PCR-screenings, positive- and negative-selection primers were used to effectively isolate target phage clones. All isolated lambda phage clones and primer pairs used for screening of each phage clone are shown in Figure S1 and

Table S2.

Three lambda phage clones (C3D, J15A, and A11B) were isolated by the PCR-screening with primer pair MHC IIBexon4-F1/R1 (Figure 1, Figure S1, and Table S2). A lambda phage clone Q6A was isolated by genome walking to upstream of C3D. A lambda phage clone Q17G was isolated by genome walking to downstream of A11B.

A lambda phage clone Q13C was isolated by the PCR-screening with primer pair TAP1exon7-F1/R1. Two lambda phage clones (K8A and K6J) were isolated by genome walking to upstream and downstream of Q13C, respectively. T69H and subsequent V66E were obtained by 2 rounds of genome-walking to downstream of K6J. Five phage clones (O19R, O62E, P32B, S14A, and X65K) were obtained as V66E F/R primer pair positive clone.

A lambda phage clone Z34D was isolated with primer pair MHC-I F/R that were designed on the basis of the MHC-I gene sequence within the clone Q13C. A lambda phage clone N49F was isolated by genome walking to downstream of Z34D.

Two lambda phage clones (L57C and G57G) were isolated by the PCR-screening with primer pair TNXB-S1/A1.

Analysis of Isolated Lambda Phage Clones

Phage DNAs of positive clones were purified, digested with *Eco*RI and *Xho*I, and then subcloned into pBluescript II vector (Stratagene). In positive phage clones C3D, J15A, A11B, and Q17G, *Eco*RI, and *Bam*HI were used (Figure 1). Sequences of both ends of some subclones were determined by using M13 forward and reverse primers and analyzed through homology searches using BLAST (http://blast.ncbi.nlm.nih.gov/blast). Sequencing was carried out by Greiner Japan Co., Ltd. (Tokyo, Japan). By combining

these results, restriction maps of each lambda phage clone were constructed. Complete sequencing of the whole MHC region is ongoing (see results).

Genetic Diversity of MHC Class II Region among 6 Founders

To investigate the genetic diversity of the MHC class II region, genomic fragments containing both *DAB1* and *DAB2* loci were amplified from 6 founders (A–F) with primers DAB1int1-F1 and BRD2-R2 (Table S1). PCR products were cloned into pBluescriptII (Stratagene) and 2 positive clones representing different alleles were selected from every founder. Exon 2 sequences of both *DAB1* and *DAB2* genes in positive clones were determined with primers DAB1int1-F1 and DAB2int1-F1 (Table S1), respectively. The haplotypes of the MHC class II were decided by the combination of *DAB1* and *DAB2* exon 2 sequences within a single clone.

Results and Discussion

Genomic Structure of MHC Region

The Oriental stork is a critically threatened avian species belonging to the order Ciconiiformes, an avian lineage that is highly divergent from order Galliformes, but close to the Japanese crested ibis (order Pelecaniformes). To determine the genomic structure of the Oriental stork MHC region, we screened a genomic library and assembled one contig (approximately 160 kb) (Figure 1, all isolated phage clones are shown in Figure S1). Partial sequences of subclones suggested that this contig contained a partial copy of *COL11A2*, 2 copies of MHC IIA/IIB pairs (*DAA1*/*DAB1* and *DAA2*/*DAB2*), *BRD2*, *DMA*, *DMB1*, *DMB2*, MHC I (*UAA*), *TAP1*, *TAP2*, 2 copies of MHC I (*UBA1* and *UBA2*), and a partial copy of *TNXB*. MHC-IIA and -IIB genes were

arranged head-to-head. Partial sequences of the DNA fragment containing *UBA1* were almost identical to those of *UBA2*, suggesting that 2 regions were very long tandem repeats. Although the number of tandem repeats containing the MHC-I gene was estimated to be 2 from comparison with short sequences (540 bp) from 9 phage clones (O19R, O62E, P32B, Z34D, S14A, X65K, V66E, Z57G, and L57C) (Figure S2), we cannot exclude the possibility that the number of tandem repeats are 3 or more. To resolve this, further analysis is necessary.

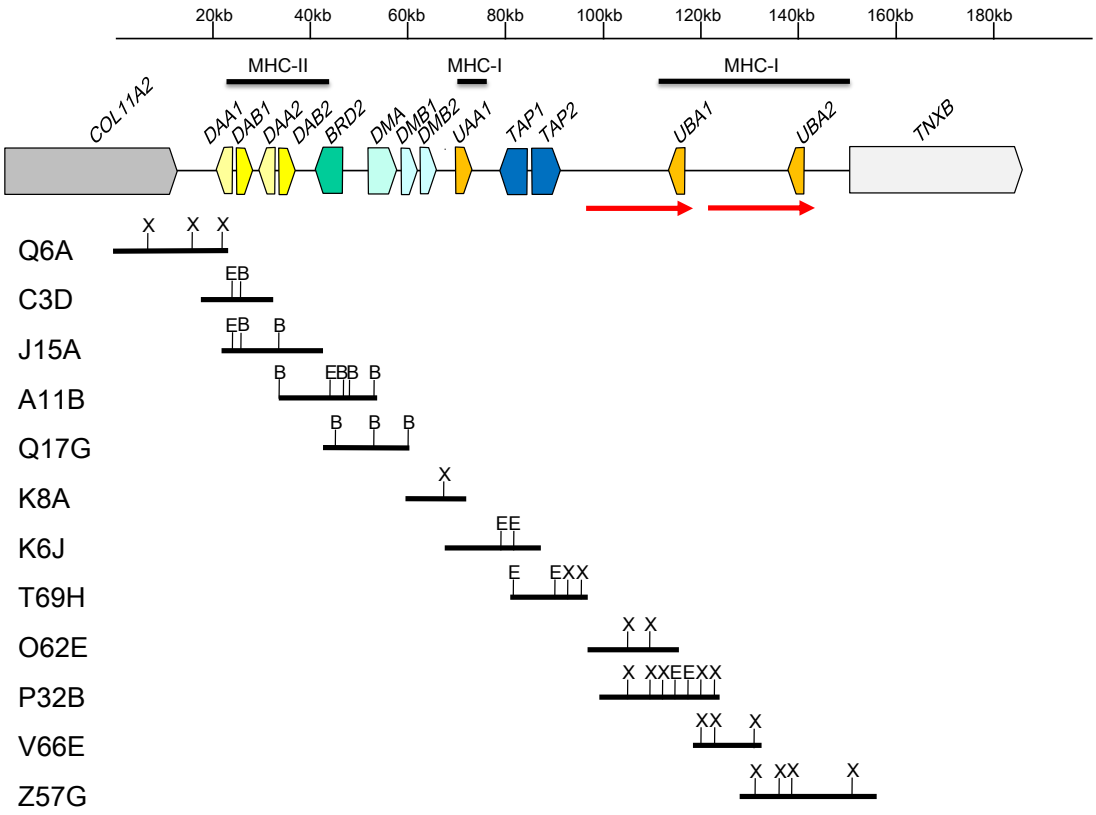

Figure 1. Genomic organization of the Oriental stork MHC. A single contig was assembled. *Collagen-type XI α-2* (*COL11A2*), 2 copies of MHC IIA/IIB pairs (*DAA1*/*DAB1* and *DAA2*/*DAB2*), *BRD2*, *DMA*, *DMB1*, *DMB2*, MHC I (*UAA*), *TAP1*, *TAP2*, 2 copies of MHC I (*UBA1* and *UBA2*), and *TNXB* genes and their orientations are indicated. Locus names are indicated below the MHC-I and -II genes. Solid bars below the map represent locations of isolated lambda phage clones consisting of minimum tailing path. B, E, and X represent restriction sites used for subcloning of *Bam*HI, *Eco*RI, and *Xho*I, respectively. Red arrows below the map represent very long repeat sequences containing the MHC-I gene.

For complete sequencing of the MHC region, insert DNAs of subclones from phage clones Q6A, A11B, Q17G, K8A, K6J T69H, V66E, and Z34D were purified, mixed, and sequenced with the next-generation sequencer HiSeq2500 (Illumina). Sequence data of approximately 1.2 Gb (12 million reads) was obtained and reads were assembled using the "Velvet" program. However, since the assembled sequence contained many gaps and many ambiguous nucleotides, especially in repeat sequence regions, phage clones consisting of minimal tailing paths were selected (Figure 1) for ongoing sequencing with a Sanger sequencer.

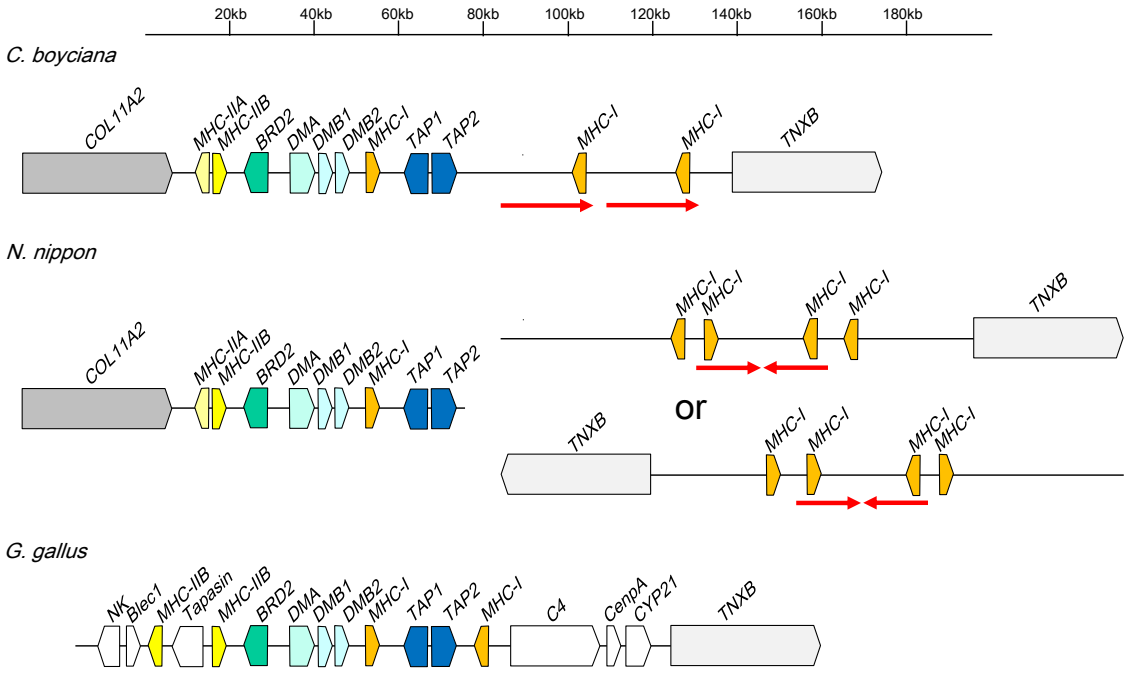

Figure 2. Comparison of the genomic organizations of the Oriental stork, Japanese crested ibis, and chicken MHC. Since the Japanese crested ibis MHC contains a gap downstream of *TAP2* gene, the orientation of the fragment containing the MHC-I cluster and *TNXB* remains unknown (Chen et al., 2015).

The genomic structure of the Oriental stork MHC region was compared with those of the Japanese crested ibis, and chicken (Figure 2). The MHC class II regions (*COL11A2* to *BRD2*) of the Oriental stork and the Japanese crested ibis showed complete synteny, but they were largely different from that of chicken. The MHC-IIB gene was flanked by the MHC-IIA gene and the MHC IIA/IIB gene pair was located between *COL11A2* and *BRD2* in the Oriental stork and the Japanese crested ibis, whereas the chicken MHC-B contained no MHC-IIA gene (*BLA*) and 2 MHC-IIB genes (*BLB1* and *BLB2*) located on both sides of the *Tapasin* gene (Kaufman et al., 1999). In contrast to the MHC class II region, the internal region (*BRD2* to *TAP2*) was highly conserved with perfect syntenic gene order between the 3 avian species. The genomic structures of MHC class I region (*TAP2* to *TNXB*) were different in each species. The chicken MHC class I region contained *C4*, *CenpA*, and *CYP21* genes, but these genes were not detected in the MHC class I regions of the Oriental stork or Japanese crested ibis. Whereas the orientation of the fragment containing the MHC-I cluster and *TNXB* remains unknown because of a gap downstream of *TAP2* in the Japanese crested ibis (Chen et al., 2015), the number and orientation of the MHC-I genes were different between the Oriental stork and Japanese crested ibis. These results suggested that evolution of the MHC region might differ largely among avian species. To elucidate the long-term evolutionary history of avian MHCs, more data on genome structure of MHCs from a wide variety of bird species are apparently needed.

Genetic Diversity of MHC Class II Region among 6 Founders

To investigate the genetic diversity of the MHC class II region among the 6 founders, genomic fragments containing both *DAB1* and *DAB2* loci were amplified with primers

DAB1int1-F1 and BRD2-R2 that were located in *DAB1* intron 1 and within *BRD2* gene, respectively. PCR products of approximately 11 kb were detected in founders A, B, C, and F (Figure 3). In addition to the 11 kb fragment, PCR products of approximately 5 kb were detected in founders D and E, suggesting that shorter fragments represented one copy of a MHC-IIA/IIB pair. Copy number variations (CNV) of MHC-IIB gene has been reported in *N. nippon* (Chapter 2) and certain bird species (Hosomichi et al., 2006; Eimes et al., 2011; Strandh et al., 2012; Alcaide et al., 2014). These results might indicate that CNVs by gene duplication, gene loss, and/or gene conversion occur at a relatively high rate within a single bird species.

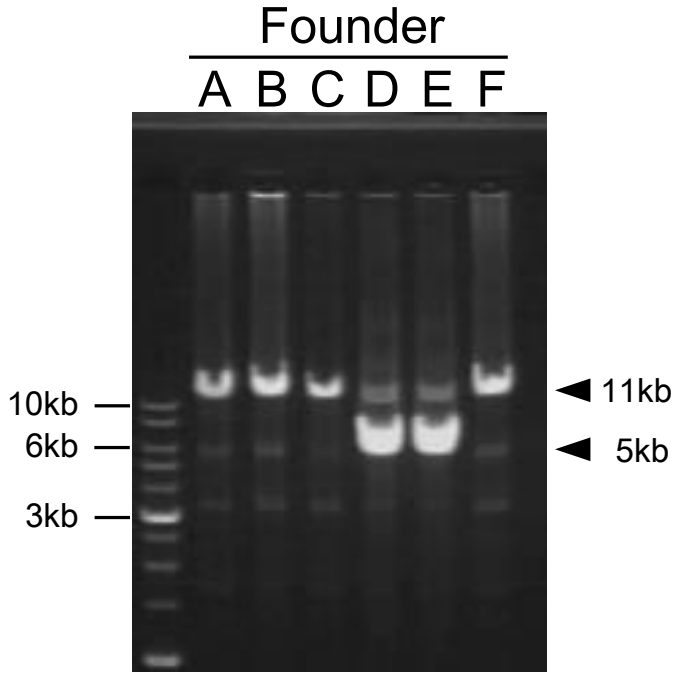

Figure 3. Amplification of the MHC-IIB gene locus from 6 founders. Genomic fragments containing both *DAB1* and *DAB2* loci were amplified from 6 founder genomes (A–F) and analyzed by 0.7% agarose gel. In contrast, a single PCR product (approximately 11 kb) was detected in founders A, B, C, and F, and 2 PCR products (approximately 5 and 11 kb) were amplified in founders D and E.

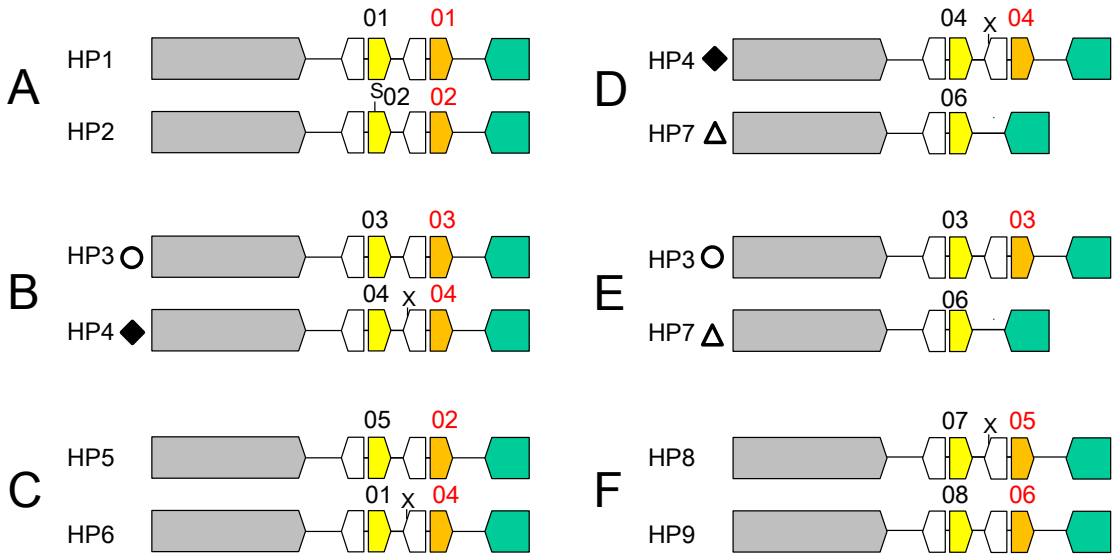

Figure 4. Haplotypes of the MHC class II region in 6 founders. PCR products containing both *DAB1* and *DAB2* loci from 6 founders (A–F) were cloned and 2 positive clones representing different alleles were selected in every founder. Exon 2 sequences of both *DAB1* and *DAB2* genes in positive clones were determined. Eight and 6 exon 2 alleles were detected in *DAB1* and *DAB2* genes, respectively. The haplotypes of the MHC class II were decided by the combination of *DAB1* and *DAB2* exon 2 sequences within a single clone.

Sequencing of PCR products revealed 8 exon 2 alleles in the *DAB1* gene and 6 exon 2 alleles in the *DAB2* gene among the 6 founders (Figure 3). The haplotypes of the MHC class II region were determined by the combination of the *DAB1* exon 2 and *DAB2* exon 2 alleles within a single PCR product. As a result, 9 haplotypes were detected in 6 founders (Figure 4). The MHC class II genotypes of founders A, B, C, D, E, and F were estimated to be *hp1/2*, *hp3/4*, *hp5/6*, *hp4/7*, *hp3/7*, and *hp8/9*, respectively.

Nine MHC class II haplotypes were detected in 6 founders. The current *C. boyciana* population in Japan originates from not only 6 founders used here but also a few other founders. For the Japanese crested ibis, the current *N. nippon* population in Japan originates from only 5 founders, from which only 3 MHC class II haplotypes were

detected. These results indicated that the genetic diversity of the MHC region in the Oriental stork population is to some extent larger than that in the Japanese crested ibis population.

Conclusion

We screened a genomic library and assembled one contig (approximately 160 kb). Partial sequences of subclones suggested that this contig contained a partial copy of *COL11A2*, 2 copies of MHC IIA/IIB pairs (*DAA1*/*DAB1* and *DAA2*/*DAB2*), *BRD2*, *DMA*, *DMB1*, *DMB2*, MHC I (*UAA*), *TAP1*, *TAP2*, 2 copies of MHC I (*UBA1* and *UBA2*) and a partial copy of *TNXB*. This is the first report of MHC genomic organization in Ciconiiformes. Our results revealed that the MHC genomic structure of the Oriental stork was close to that of the Japanese crested ibis, but largely different from that of Galliformes. Nine MHC class II haplotypes were detected in 6 founders, suggesting that the genetic diversity of the MHC region in the *C. boyciana* population in Japan is to some extent larger than that in the *N. nippon* population. The structure and polymorphism of the MHC region presented here will provide valuable insights for future studies on the evolution of the avian MHC and for the conservation of the Oriental stork.

Supporting information

Table S1. List of primers used for polymerase chain reaction analysis

Table S2. Primer pairs used for screening of each positive clone

Figure S1. Genomic organization of the Oriental stork MHC

Figure S2. Alignment of nucleotide sequences of 540 bp from the *Xho*I site among 9 phage clones

Table S1.List of primers used for PCR

Primer name	Sequence 5'-3'	Annealing temperature (°C)	Product size (bp)	Usage	sequence used for primer design
MHC IIBexon4-F1	AGACGGAGCGCGTGGTGTCCAC	63	356		MHC-IIB gene sequence from chicken, quail and crested ibis
MHC IIBexon4-R1	GAAGAGCCCCAGCGCCAGGAAG				
TAP1exon7-F1	GCCCAGGGTAGGAGAACCA	60	176	cloning	TAP1 gene sequence from chicken and quail
TAP1exon7-R1	CTGTCCCCAGGTCCTGCTC				
TNXB-S1	GTGCACCTCCACCCGGTACA	63	288	cloning	TNXB gene sequence from crested ibis
TNXB-A1	GTGGCCCGTGCCATCCTCTC				
C3D-S	AAGTGCTCTGATTCACACCAG	53	290	genome walking	5' terminal sequence of clone C3D
C3D-A	CCAGACAAGGAGCAAATCCC				
A11B-F	CACCATCCACCCAACTCCT	58	386	genome walking	3' terminal sequence of clone A11B
A11B-R	ACCAGTGACGAGTGGTGTC				
A11B BRD2-F	GCCGGAGGTTTCCAACCCCAAG	63	300	genome walking	BRD2 gene sequence within clone A11B
A11B BRD2-R	ACATCCGGGGTCAGGCAGCA				
Q13C-S	ATGTGGGACCCCAGGGGTAGT	62	293	genome walking	5' terminal sequence of clone Q13C
Q13C-A	TGAGCCCCTCAGACCGTCCTT				
MHC I-F	GGGAGGGACTTCATCGCCTT	60	304	cloning, genome walking	MHC-I gene sequence within clone Q13C
MHC I-R	GCCTCCTTCCCCGACACT				
Q13C-F	TCCCCTCCAAGTGCCAGTGA	60	217	genome walking	3' terminal sequence of clone Q13C
Q13C-R	GGCACAAAGTGGGGAGGACAA				
K6J-F	CGTCTACCAAAGGAAACTGCC	57	195	genome walking	3' terminal sequence of clone K6J
K6J-R	TTTTCTTACCGGTTGTGCTCC				
T69H-F	TACCAGCTGGAGCTCTGCC	60	281	genome walking	3' terminal sequence of clone T69H
T69H-R	AGGACAAGGCACCCCAGAAG				
V66E-F	TGTGACTAGTCCTTGCAGCGT	58	201	cloning, genome walking	3' terminal sequence of clone V66E
V66E-R	AGGGCCTTATACTTCCATGGGT				
Z34D-F	GGCAACTCTGCAGACAGCAAG	59	374	genome walking	3' terminal sequence of clone Z34D
Z34D-R	TGGGCTCACTTTCACACCCAA				
DAB1int1-F1	GATGAATTCCTCAGGATCGCACCGCCAGTGCTATCAG	72		amplification of fragments containing both DAB1 and DAB2 loci	
BRD2-R2	CATGAATTCAAGCACAAAGGCCGAGGAGGTGACGA				
DAB1int1-F1	ATCAGAGCGGTGTGGCTATG			sequencicg of DAB1 exon 2	
DAB2int1-F1	GGAGGAGATGAGTGCAAAAG			sequencicg of DAB2 exon 2	

Table S2. Primer pairs used for screening of each positive clones		
clone name	positive selection primer pair	negative selection primer pair
Q6A	C3D-S/A	IIBexon4-F1/R1
C3D	IIBexon4-F1/R1	
J15A	IIBexon4-F1/R1	
A11B	IIBexon4-F1/R1	
Q17G	A11B-F/R	A11B BRD2-F/R
K8A	Q13C-S/A	Q13C-F/R
Q13C	TAP1 exon7-F/R	
K6J	Q13C-F/R	Q13C-S/A
T69H	K6J-F/R	MHC I-F/R
O19R	V66E-F/R	
S14A	V66E-F/R	
X65K	V66E-F/R	
O62E	V66E-F/R	
P32B	V66E-F/R	
Z34D	MHC I-F/R	
N49F	Z34D-F/R, MHC I-F/R	
V66E	T69H-F/R	K6J F/R
Z57D	TNXB-S1/A1	
L57C	TNXB-S1/A1	

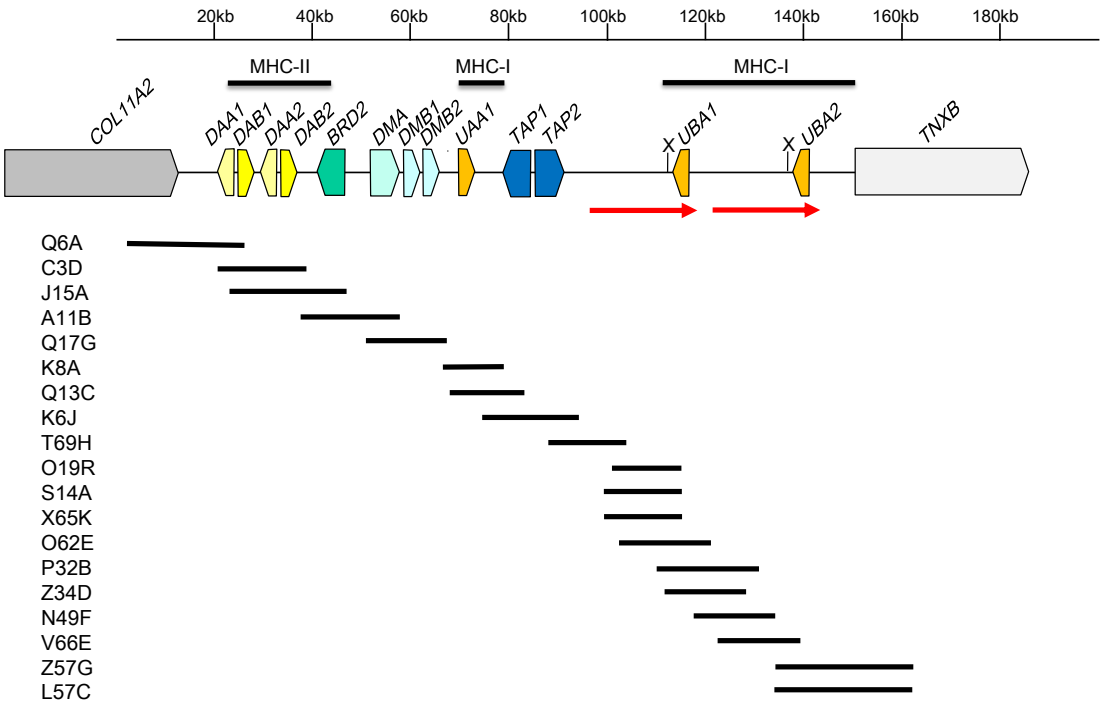

Figure S1. Genomic organization of the Oriental stork MHC. *Collagen-type XI α-2* (*COL11A2*), 2 copies of MHC IIA/IIB pairs (*DAA1*/*DAB1* and *DAA2*/*DAB2*), *BRD2*, *DMA*, *DMB1*, *DMB2*, MHC I (*UAA*), *TAP1*, *TAP2*, 2 copies of MHC I (*UBA1* and *UBA2*), and *TNXB* genes and their orientations are indicated. All isolated lambda phage clones are depicted below the map. X on the map indicates the *Xho*I site used for sequence comparison between 9 phage clones O19R, S14A, X65K, O62E, P32B, Z34D, V66E, Z57G, and L57C (Figure S2).

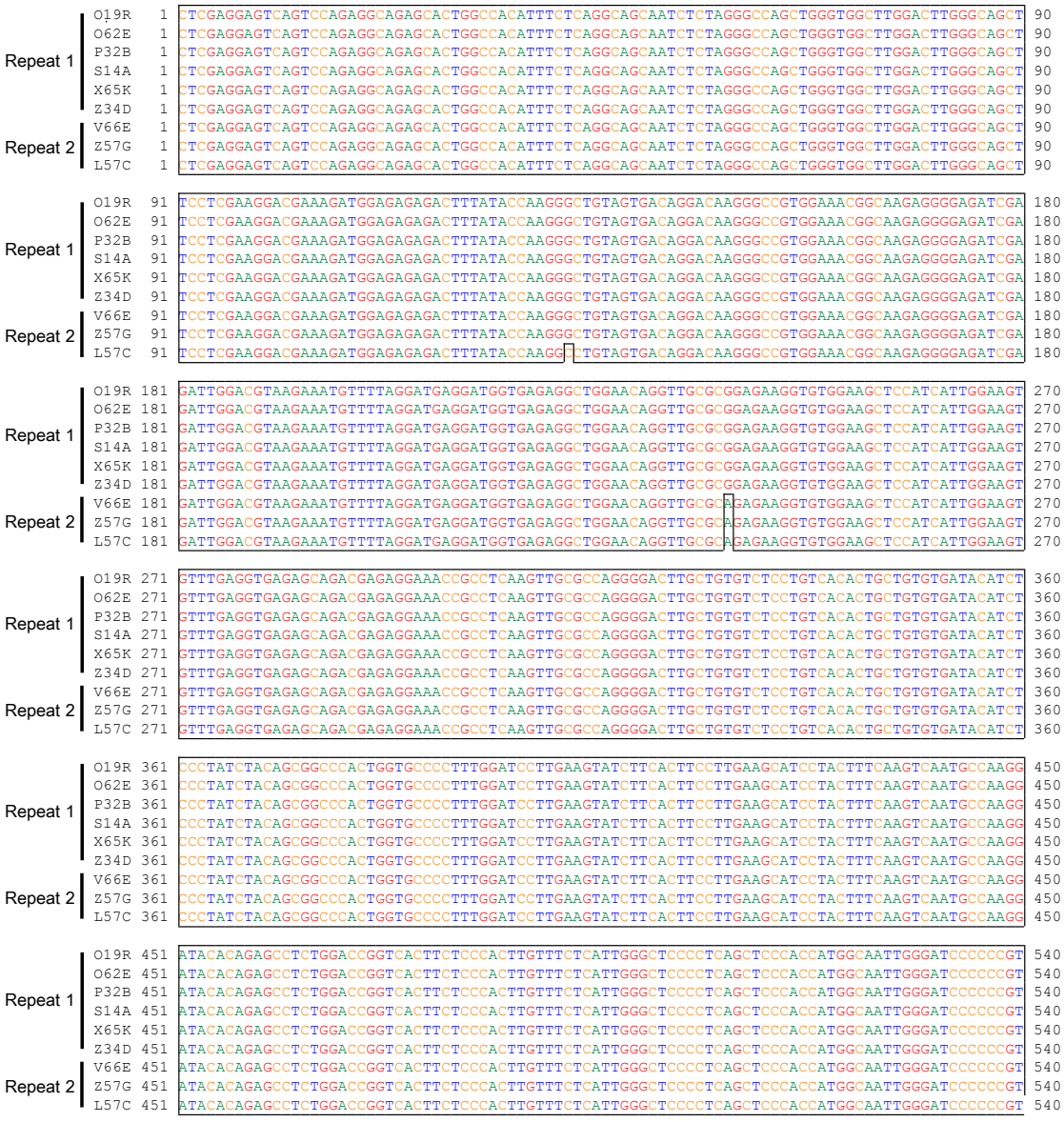

Figure S2. Alignment of nucleotide sequences of 540 bp from the *Xho*I site among 9 phage clones. 540 bp sequences from 9 phage clones O19R, S14A, X65K, O62E, P32B, Z34D, V66E, Z57G, and L57C were aligned. Nucleotide substitutions were detected in 2 positions. Polymorphism at nt239 suggested the different loci (repeat 1: O19R, S14A, X65K, O62E, P32B, and Z34D, repeat 2: V66E, Z57G, and L57C), whereas polymorphism at nt132 suggested 2 alleles within repeat 2.

SUMMARY

The Japanese crested ibis *N. nippon* and the Oriental stork *C. boyciana* are endangered avian species and their conservation projects are intensively promoted in Japan. It is important for these conservation projects to obtain information on genetic relatedness by using molecular tools such as single nucleotide polymorphism (SNP) and short tandem repeat (STR) markers and maintain the genetic diversity among populations.

The current Japanese population of *N. nippon* originates from only 5 founders donated by the Chinese government. To discover genome-wide SNPs and STRs while obtaining genotype data of these polymorphic markers in each founder, reduced representation libraries were independently prepared from each of the founder genomes and sequenced on an Illumina HiSeq2000. This yielded 316 million 101-bp reads. Consensus sequences were created by clustering sequence reads, and then sequence reads from each founder were mapped to the consensus sequences, resulting in the detection of 52,512 putative SNPs and 162 putative STRs. Moreover, genotype data of 32,157 putative SNPs were estimated. The haplotypes numbers and STR alleles and the investigation of genetic similarities suggested that the total genetic diversity between the founders was low, although we could not identify a pair with closely related genome sequences.

In addition to SNP and STR makers, polymorphisms of functional genes could provide valuable information for the conservation of endangered species. The major histocompatibility complex (MHC) is a highly polymorphic genomic region that plays a central role in the immune system. Because of its high polymorphism and importance for disease resistance and other functions, the MHC has been an important focus in the conservation of endangered species.

We investigated the structure and polymorphism of the *N. nippon* MHC class II region and the *C. boyciana* MHC region. In *N. nippon*, 3 contigs representing different haplotypes of MHC class II regions were assembled. A pair of MHC-IIA and -IIB genes was arranged head-to-head between the *COL11A2* and *BRD2* genes. The 3 haplotypes contained 1–3 copies of MHC-IIA/IIB gene pairs. Genotyping of the MHC class II region detected only 3 haplotypes among the 5 founders, suggesting that the genetic diversity of the current *N. nippon* population is extremely low. In *C. boyciana*, one contig (approximately 160 kb) was assembled. Partial sequences of subclones suggested that this contig contained a partial copy of *COL11A2*, 2 copies of MHC IIA/IIB pairs (*DAA1*/*DAB1* and *DAA2*/*DAB2*), *BRD2, DMA*, *DMB1, DMB2*, MHC I (*UAA*), *TAP1*, *TAP2*, 2 copies of MHC I (*UBA1* and *UBA2*), and a partial copy of *TNXB*. These results revealed that the MHC genomic structure of *N. nippon* and *C. boyciana* were conserved, but they were largely different from that of Galliformes. Nine MHC class II haplotypes were detected in 6 founders of *C. boyciana*, suggesting that the genetic diversity of the MHC region in the *N. nippon* population is lower than that in the *C. boyciana* population in Japan.

The present study provided important insight into protocols for genetic management of the captive breeding population of *N. nippon* and *C. boyciana* in Japan, toward the national project for the reintroduction of captive-bred individuals into the wild and for future studies on the evolution of the avian MHC. In addition, we proposed a simple, efficient, and cost-effective approach for the simultaneous detection of genome-wide polymorphic markers and their genotypes for species currently lacking a reference genome sequence.

ACKNOWLEDGMENT

The author wishes to express his most sincere appreciation and respect to Professor, Dr. H. Iwaisaki for his valuable advice, assistance and encouragement throughout the course of this study.

The author wishes to express his deep gratitude Dr. Y. Sasaki, Dr. T. Yamada, Dr. K. Moriya and Dr. H. Matsuda for their help in conducting this study.

The author thanks Dr. Y. Kaneko and staffs at the Sado Japanese Crested Ibis Conservation Center, and Dr. K. Naito and staffs at the Hyogo Park of the Oriental White Stork for providing rare tissue samples of the Japanese Crested Ibis and the Oriental Stork, respectively.

The author is also grateful to colleagues in the laboratory on Animal Breeding and Genetics.

Finally, the author wishes to express his deepest respect and gratitude to the late Dr. Y. Kitagawa at Nagoya University for assistance and encouragement at the beginning of his research career.

REFERENCES

Alcaida M, Edwards SV, Negro JJ (2007) Characterization, polymorphism, and evolution of MHC class II B genes in birds of prey. J Mol Evol 65: 541-554.

Alcaide M, Muñoz J, Martínez-de la Puente J, Soriguer R, Figuerola J (2014) Extraordinary MHC class II B diversity in a non-passerine, wild bird: the Eurasian Coot Fulica atra (Aves: Rallidae). Ecol Evol 4: 688-698

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3: e3376.

Balakrishnan C, Ekblom R, Volker M, Westerdahl H, Godinez R, et al. (2010) Gene duplication and fragmentation in the zebra finch major histocompatibility complex. BMC Biol 8: 29.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, et al. (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364: 33-39.

Burri R, Hirzel HN, Salamin N, Roulin A, Fumagalli L (2008) Evolutionary patterns of MHC class II B in owls and their implications for the understanding of avian MHC evolution. Mol Biol Evol 25: 1180-1191.

Burri R, Salamin N, Studer RA, Roulin A, Fumagalli L. (2010) Adaptive divergence of ancient gene duplicates in the avian MHC class II beta. Mol Biol Evol 27: 2360-2374.

Chaves LD, Krueth SB, Reed KM (2009) Defining the turkey MHC: sequence and genes of the B locus. J Immunol 183: 6530-6537.

Chen LC, Lan H, Sun L, Deng YL, Tang KY, Wan QH (2015) Genomic organization of

the crested ibis MHC provides new insight into ancestral avian MHC structure. Sci Rep 5:7963.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, et al. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12: 499-510.

Delany ME, Robinson CM, Goto RM, Miller MM (2009) Architecture and organization of chicken microchromosome 16: order of the NOR, MHC-Y, and MHC-B subregions. J Hered 100: 507-514.

Durrell L, Mallinson J (1987) Reintroduction as a political and educational tool for conservation. J Jersey Wildlife Preserv Trust 24: 6-19.

Eimes JA, Bollmer JL, Whittingham LA, Johnson JA, VAN Oosterhout C, et al. (2011) Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. J Evol Biol 24: 1847-1856.

Eimes JA, Reed KM, Mendoza KM, Bollmer JL, Whittingham LA, et al. (2013) Greater prairie chickens have a compact MHC-B with a single class IA locus. Immunogenetics 65: 133-144.

Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, et al. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). Mol Ecol 16: 1439-1451.

Ekblom R, Stapley J, Ball AD, Birkhead T, Burke T, et al. (2011) Genetic mapping of the major histocompatibility complex in the zebra finch (*Taeniopygia guttata*). Immunogenetics 63: 523-530.

Gregory TR, Nicol JA, Tamm H, Kullman B, Kullman K, et al. (2007) Eukaryotic genome size databases. Nucleic Acids Res 35: D332-D338.

Hale ML, Verduijn MH, Moller AP, Wolff K, Petrie M (2009) Is the peacock's train an honest signal of genetic quality at the major histocompatibility complex? J Evol Bio 22: 1284-1294.

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 Gen 7: 157-160.

He X-L, Ding C-Q, Han J-L (2013) Lack of structural variation but extensive length polymorphisms and heteroplasmic length variations in the mitochondrial DNA control region of highly inbred Crested Ibis, *Nipponia nippon*. PLoS ONE 8(6): e66324.

Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, et al. (2004) Gene map of the extended human MHC. Nat Rev Genet 5: 889-899.

Hosomichi K, Shiina T, Suzuki S, Tanaka M, Shimizu S, et al. (2006) The major histocompatibility complex (MHC) class IIB region has greater genomic structural flexibility and diversity in the quail than the chicken. BMC Genomics 7: 322.

Ji YJ, Liu YD, Ding CQ, Zhang DX (2004) Eight polymorphic microsatellite loci for the critically endangered Crested Ibis, *Nipponia nippon* (Ciconiiformes: Threskiornithidae). Mol Ecol Notes 4: 615-617.

Kaufman J, Milne S, Göbel TW, Walker BA, Jacob JP, et al. (1999) The chicken B locus is a minimal essential major histocompatibility complex. Nature 401: 923-925.

Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. Immunogenetics 56: 683-695.

Kerstens H, Crooijmans R, Veenendaal A, Dibbits B, Chin-A-Woeng T, et al. (2009)

Large scale single nucleotide polymorphism discovery in unsequenced genomes using second generation high throughput sequencing technology: applied to turkey. BMC Genomics 10: 479.

Klein J (1986) Antigen-major histocompatibility complex-T cell receptors: inquiries into the immunological ménage à trois. Immunol Res 5: 173-190.

Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, et al. (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. Immunogenetics 31: 217-219.

Kraus RH, Kerstens HH, Van Hooft P, Crooijmans RP, Van Der Poel JJ, et al. (2011) Genome wide SNP discovery, analysis and evaluation in mallard (Anas platyrhynchos). BMC Genomics 12: 150.

Kulski JK, Shiina T, Anzai T, Kohara S, Inoko H (2002) Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. Immunol Rev. 190: 95-122.

Lande R (1988) Genetics and demography in biological conservation. Science 241: 1455-1460.

Li L, Zhou X, Chen X (2011) Characterization and evolution of MHC class II B genes in Ardeid birds. J Mol Evol 72: 474-483.

Li R, Fan W, Tian G, Zhu H, He L, et al. (2010) The sequence and de novo assembly of the giant panda genome. Nature 463: 311-317.

Liu Y (1981) Recovery of Japanese crested ibis in Qin-Ling range. Acta Zoology Sinica 27: 273.

Miller HC, Lambert DM (2004) Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). Immunogenetics. 56: 178-191.

Miller MM, Goto R, Bernot A, Zoorob R, Auffray C, et al. (1994) Two MHC class I and two MHC class II genes map to the chicken Rfp-Y system outside the B complex. Proc Natl Acad Sci USA 91: 4397-4401.

News of the week, Sado Island, Japan Back From the Brink (2012) Around the World. Science 336: 524-525.

Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, et al. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS One 4: e6524.

Ren L, Yang Z, Wang T, Sun Y, Guo Y, et al. (2011) Characterization of the MHC class II α-chain gene in ducks. Immunogenetics 63: 667-678.

Roberts SC (2009) Complexity and context of MHC-correlated mating preferences in wild populations. Mol Ecol 18: 3121-3123.

Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, et al. (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature 368: 215-221.

Strandh M, Westerdahl H, Pontarp M, Canbäck B, Dubois MP, et al. (2012) Major histocompatibility complex class II compatibility, but not class I, predicts mate choice in a bird with highly developed olfaction. Proc Biol Sci 279: 4457-4463.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.

Tsuda TT, Tsuda M, Naruse T, Kawata H, Ando A, et al. (2001) Phylogenetic analysis of penguin (Spheniscidae) species based on sequence variation in MHC class II genes. Immunogenetics 53: 712-6.

Urano K, Tsubono K, Taniguchi Y, Matsuda H, Yamada T, et al. (2013) Genetic diversity and structure in the Sado captive population of the Japanese crested ibis. Zoolog Sci. 30(6): 432-438.

van Bers NE, van Oers K, Kerstens HH, Dibbits BW, Crooijmans RP, et al. (2010) Genome-wide SNP detection in the great tit Parus major using high throughput sequencing. Mol Ecol 19 Suppl 1: 89-99.

van Orsouw NJ, Hogers RCJ, Janssen A, Yalcin F, Snoeijers S, et al. (2007) Complexity reduction of polymorphic sequences (CRoPS): a novel approach for large-scale polymorphism discovery in complex genomes. PLoS One 2: e1172.

Van Tassell CP, Smith TP, Matukumalli LK, Taylor JF, Schnabel RD, et al. (2008) SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat Methods 5: 247-252.

Von Schantz T, Wittzell H, Goransson G, Grahn M (1997) Mate choice, male condition-dependent ornamentation and MHC in the pheasant. Hereditas 127: 133-140.

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, et al. (2009) Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326: 865-867.

Wang B, Ekblom R, Strand TM, Portela-Bens S, Höglund J (2012) Sequencing of the core MHC region of black grouse (*Tetrao tetrix*) and comparative genomics of the galliform MHC. BMC Genomics 13: 553.

Wang J, Wang W, Li R, Li Y, Tian G, et al. (2008) The diploid genome sequence of an Asian individual. Nature 456: 60-65.

Wiedmann RT, Smith TP, Nonneman DJ (2008) SNP discovery in swine by reduced

representation and high throughput pyrosequencing. BMC Genet 9: 81.

Ye Q, He K, Wu SY, Wan QH (2012) Isolation of a 97-kb minimal essential MHC B locus from a new reverse-4D BAC library of the golden pheasant. PLoS One 7(3): e32154.

Zhang B, Fang S-G, Xi Y-M (2004) Low genetic diversity in the endangered crested ibis Nipponia Nippon and implications for conservation. Bird Conserv Int 14: 183-190.

Zinkernagel RM, Doherty PC (1979) MHC-restricted cytotoxic T cells: studies on the biological role of polymorphic major transplantation antigens determining T-cell restriction-specificity, function and responsiveness. Adv Immunol 27: 51-177.

Zhang B, Fang S-G, Xi Y-M (2006) Major histocompatibility complex variation in the endangered crested ibis Nipponia nippon and implications for reintroduction. Biochem Genet. 44: 113-123.