

Characterization, Polymorphism and Selection of Major Histocompatibility Complex (MHC) DAB Genes in Vulnerable Chinese Egret (*Egretta eulophotes*)

Zeng Wang¹, Xiaoping Zhou^{1,2}, Qingxian Lin^{1,2}, Wenzhen Fang^{1,2*}, Xiaolin Chen^{1,2*}

1 Key Laboratory of Ministry of Education for Coast and Wetland Ecosystems, School of Life Sciences, Xiamen University, Xiamen, Fujian, People's Republic of China, **2** Key Laboratory of Ministry of Education for Coast and Wetland Ecosystems, College of the Environment and Ecology, Xiamen University, Xiamen, Fujian, People's Republic of China

Abstract

The major histocompatibility complex (MHC) is an excellent molecular marker for the studies of evolutionary ecology and conservation genetics because it is a family of highly polymorphic genes that play a key role in vertebrate immune response. In this study, the functional genes of MHC Class II B (DAB) were isolated for the first time in a vulnerable species, the Chinese egret (*Egretta eulophotes*). Using a full length DNA and cDNA produced by PCR and RACE methods, four potential MHC DAB loci were characterized in the genome of this egret and all four were expressed in liver and blood. At least four copies of the MHC gene complex were similar to two copies of the minimal essential MHC complex of chicken, but are less complex than the multiple copies expressed in passerine species. In MHC polymorphism, 19 alleles of exon 2 were isolated from 48 individuals using PCR. No stop codons or frameshift mutations were found in any of the coding regions. The signatures of positive selection detected in potential peptide-binding regions by Bayesian analysis, suggesting that all of these genes were functional. These data will provide the fundamental basis for further studies to elucidate the mechanisms and significance of MHC molecular adaptation in vulnerable Chinese egret and other ardeids.

Citation: Wang Z, Zhou X, Lin Q, Fang W, Chen X (2013) Characterization, Polymorphism and Selection of Major Histocompatibility Complex (MHC) DAB Genes in Vulnerable Chinese Egret (*Egretta eulophotes*). PLoS ONE 8(9): e74185. doi:10.1371/journal.pone.0074185

Editor: Christian Schönbach, Kyushu Institute of Technology, Japan

Received: April 18, 2013; **Accepted:** July 28, 2013; **Published:** September 3, 2013

Copyright: © 2013 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by the National Natural Science Foundation of China (<http://www.nsf.gov.cn>) (Grant # 30970380 to XC, 40876077 to WF, and 31000963 to XZ), the Fujian Natural Science Foundation of China (<http://xmgl.fjkjt.gov.cn/>) (Grant # 2010Y2007 and 2012J01149 to XC, 2009J01195 to WF, and 2010J05084 to XZ), and the Fundamental Research Funds for the Xiamen Universities (<http://www.xmu.edu.cn/>) (Grant # 2011121038 to XZ). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: wzfang@xmu.edu.cn (WF); xlchen@xmu.edu.cn (XC)

☉ These authors contributed equally to this work.

Introduction

The major histocompatibility complex (MHC) is a multi-gene family, representing an important component in the acquired immune system of the vertebrate species. Traditionally, MHC genes are classified into two major classes, designated class I and class II. The MHC class II genes consists of two genes, named A and B, in which the B genes are responsible for the majority of the polymorphism. MHC is a favored molecular marker for evolutionary ecology and conservation genetics because of their diverse functions and characteristics relevant to evolutionary and adaptive processes [1–6].

Studies on the functional genes of MHC II B (DAB) complex in birds have indicated that multiple copies of these genes are highly variable among different species [5,7–18]. The most complex of the MHC structures were found in passerine birds [19–21] while the “minimal essential” structures were found in a

non-passerine bird, the red jungle fowl (Galliformes, Phasianidae, *Gallus gallus*) [22,23]. Other non-passerine birds, including raptors and seabirds, paleognath birds, and penguins have been reported to have more complex MHC structures than the “minimal essential” genes [24].

Efficient and reliable genotyping of MHC genes is a difficult task, yet remains a basic prerequisite to understand the MHC complex in these species [13,16,25]. Target loci are commonly expressed as multiple copies, and allele sequences, even at a single locus, may be very divergent. This makes for a more challenging problem to identify all alleles carried by an individual and reconstruction of multilocus genotype in an individual. The presence of both expressed loci and pseudogenes further complicates the problem in identifying functional variants [14,20,26]. Moreover, the copy number of MHC genes in birds is highly variable. Therefore, the precise genotyping and full understanding of the molecular architecture

of the multilocus MHC complex remains an important and challenging problem [12,27].

The Chinese egret (*Ciconiiformes*, *Ardeidae*, *Egretta eulophotes*) is listed as a vulnerable species with an estimated global population of 2600–3400 individuals [28,29]. This egret is a migratory colonial waterbird, wintering in the south of Asia and breeding on offshore islands in Russia, North Korea, South Korea and China. This breeding pattern may facilitate pathogen transmission and advance MHC polymorphisms [24]. Several molecular studies have been carried out on this species [30–32], including reports of MHC and two DNA sequences of the DAB gene from exon 2 to exon 3 but these studies did not confirm their expression [33]. The aims of the present study were to: (1) isolate the complete MHC DAB genes in the Chinese egret and determine their architecture; (2) confirm the expression of all the loci; (3) estimate variability in each locus and test to determine if positive selection was acting on the MHC DAB genes; and (4) to investigate the usefulness of noncoding sequence variation for the design of locus-specific amplification strategies. Successful completion of these objectives will provide the essential fundamentals for further studies to determine the mechanism and significance of MHC molecular adaptation in vulnerable Chinese egret and other ardeids.

Materials and Methods

Ethics Statement

This research and all procedures involving collection of animal tissue in the wild were approved by the Administration Center for Wildlife Conservation in Fujian Province (FJWCA-1208). The scientific license for access to the study site was issued by the Administration Department of Xiamen Egret Natural Reserve (XMENR-1005). Blood samples (~0.5 mL) were collected from 32 nestlings (aged around 15 days) of *E. eulophotes* by puncturing the wing vein and using a syringe to draw up the blood. The nestlings were immediately returned to the nest after stanching the blood with cotton. Collection of the blood sample was conducted during the morning, and visits to the breeding colony were restricted to a maximum of two hours per day.

Methods for DNA Extraction, RNA Extraction and Reverse Transcription

Blood samples were collected from 48 nestlings (aged around 15 days) of the Chinese egret on the Riyu Islet (27°01'N, 120°25'E) in Fujian Province of China as described above, and frozen at -80°C until DNA extraction. Total genomic DNA was isolated using the Universal Genomic DNA Extraction Kit Ver. 3.0 (TaKaRa), and total RNA was extracted from the blood and from the liver of one individual Chinese egret that died at this Islet by Trizol (Invitrogen, Switzerland), respectively.

Isolation of MHC DAB Genes

A fragment from exon 1 to exon 3 of the MHC DAB gene complex in the Chinese egret was first amplified with

degenerate primers 34F and Rap3R [11]. PCR reactions were carried out on a Biometra T gradient thermocycler in a final volume of 20 µL containing 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂), 0.2 mM of each dNTP, 0.4 µM of each primer, 0.7 U of Taq polymerase (TaKaRa) and 100 ng of genomic DNA. The conditions for PCR amplification were a denaturing step at 94 °C for 3 min, followed by 35 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 2 min, and a final extension at 72 °C for 10 min.

The 3' and 5' rapid amplification strategy of cDNA ends (RACE) [34] was used to isolate the entire gene sequence and confirm the gene transcription by using 3'-full RACE core set ver. 2.0 (TaKaRa) and 5'-full RACE Kit (TaKaRa). Based on the DNA sequence obtained from the fragment, gene specific outer/inner primers were designed in conserved exon 3 regions for 3' RACE (23sp1: ACAGGCTGGTTTGCTACGTGACG 23sp2: CCGCGGAGAT TGAGGTGAAGTGG) and 5' RACE (25sp1: CCACCTGGCACATGTAG GTGTCC 25sp2: TGGTTTCCAGCATCACCAGCACCTG) respectively, using Oligo 6.0 (Molecular Biology Insights). PCR reactions were carried out following the protocols and applications as described in the kit. A 580 bp segment and a 650 bp fragment were obtained from the 3' RACE and 5' RACE, respectively. After removing the overlapping region, a 1060 bp long cDNA segment of the MHC DAB genes from the 5' UTR to the 3' UTR was obtained.

According to the complete cDNA sequence, two primers 25utrF CAGAACTCTGCCCGGAGACGG and 23utrR CTGCAGAGCAGCGACAG CGAA were designed to amplify the complete DNA sequence of the MHC DAB gene. PCR reactions were carried out on a Biometra T gradient thermocycler in a final volume of 50 µL containing 1 × LA PCR buffer II, 2.5 mM MgCl₂, 0.4 mM dNTP, 0.5 µM of each primer, and 2 U TaKaRa, LA Taq. PCR conditions included an initial denaturation step at 95°C for 3 min, 35 cycles of denaturation at 95°C for 30 s, annealing at 62°C for 30 s, primer extension at 72°C for 3 min, and a final extension at 72 °C for 10 min.

To verify differences among the four MHC DAB loci in the intron 1 and locate the conserved regions to design locus-specific primers, two primers 25utrF1: AGCTCATTTTCGGGAGGGGGTC and 23exr: CATCACCAGCAC CTGGTAGGTCC were used to amplify a fragment from the 5' UTR to exon 3 for the 32 nestlings randomly selected. PCR reactions were carried out in a final volume of 20 µL containing 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂), 0.2 mM of each dNTP, 0.4 µM of each primer, 0.7 U of Taq polymerase (TaKaRa) and 100 ng of genomic DNA. The condition for PCR amplification included a denaturing step at 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 62 °C for 30 s, 72 °C for 2 min, and a final extension at 72 °C for 10 min. All PCR products were purified using the Agarose Gel DNA Purification Kit Ver. 2.0 (TaKaRa), then ligated into pMD 18-T vector (TaKaRa) and transformed into *Escherichia coli* DH5α. Ten positive colonies of each band were selected to sequence bidirectionally on an automatic sequencer (ABI PRISM 3730; Invitrogen Biotechnology Co. Ltd.) using universal M13 sequencing primers and BigDye version 3 (Applied Biosystems).

Polymorphism of the MHC DAB Gene Complex

To detect polymorphism within the MHC DAB genes in wild populations of the Chinese egret, 48 nestlings on the Riyu Islet were genotyped by semi-nested PCR combined with single strand conformation polymorphism (SSCP). First, four forward locus-specific primers in intron 1 were designed and designated as DAB01F-04F (DAB01F: GAGGTGCTG GGTGTGGGTGG, DAB02F: GAGGTGCTGGGTGTGGGAATGG, DAB03F: CACTACCAGGACA CGGCTTGTGG and DAB04F: CTACAATTCCAGTTTAA GTGCGACTG). These were combined with the reverse primer DAB2exR: CCACGT GCTCACCTCTCCTATCC to amplify the four loci, respectively. PCR was carried out in a 20 μ L reaction mixture containing 1 μ L genomic DNA, 0.7 U of Taq polymerase (TaKaRa), 1.5 mM $MgCl_2$, 200 μ M of each dNTP and 0.4 μ M of each primer, for 25 cycles at 94°C for 30 s, 58°C for 30 s and 72°C for 1 min. Second, to obtain suitable length fragments for SSCP genotyping and scanning of the variation in exon 2 of each locus, a second round PCR as carried out using the primer DAB2exF: CTGCACAAACAGGG GTTTCCAG. DAB2exR was used to amplify the entire exon 2 in each locus. The reaction conditions for the second round PCR were identical with those described for the first round. The PCR products from first round were diluted 100-fold and used as the template. The amplicons were separated by SSCP analyses as described by Fain [35]. Five μ L of each PCR product were mixed with an equal volume of 95% formamide, denatured for 5 min at 99°C and immediately chilled on ice water for 10 min. The reaction mixture was loaded on an 8% nondenaturing polyacrylamide gel (PAGE) (37.5:1). Electrophoresis was carried out at 10°C for 5 h at 8 W per gel followed by a sensitive silver-staining procedure, as described by Budowle [36]. Samples with equal banding patterns were rearranged and electrophoresed a second time on a nondenaturing PAGE on adjacent lanes to ensure the genotyping, and samples with unique genotypes, were typed using two independent PCR reactions. All SSCP-bands were cut from at least two individuals of each genotype and incubated in 80 μ L water for at least 3 h at 37°C. One μ L was used as template in a second PCR under the same conditions as described above. The products were screened in an additional SSCP analysis and every allele was directly sequenced in both directions from at least two individuals or two independent PCRs from one individual.

Exclusion of PCR Errors and Definition of Alleles

For all PCRs, the amplification, cloning or sequencing was carried out twice and only sequences found in both experiments were included in the analysis to avoid the inclusion of PCR artifacts. Since recombination of cloned PCR products can introduce additional artifacts [37,38], direct sequencing of uncloned PCR products was used for agreement with the polymorphic sites. The sequence obtained from SSCP was directly sequenced in both directions. Only completely identical sequences found in two independent PCR events were defined as alleles. These alleles were referred to as “verified” because the same sequence could not arise twice from independent amplification errors. Throughout this study,

the word “allele” is used to describe a 270 bp exon 2 sequence derived from SSCP genotyping.

Data Analyses

The complete DNA and cDNA sequences for MHC DAB, derived from the Chinese egret, were aligned using the DNAMAN software package (Lynnon Biosoft, Quebec, Canada). Exons and introns were distinguished using the GenScan Program [39]. DNA fragments obtained from the 32 nestlings were edited in the BioEdit Sequence Alignment Editor [40]. Alignment gaps were treated as missing data in all analyses. Sequence variability statistics were calculated using the Mega 4.0 Program [41]. The proportion of synonymous (d_s) and non-synonymous (d_n) substitutions were calculated for both the 24 possible peptide-binding regions (PBRs) as defined by Brown [42], and the remaining non-antigen binding sites (non-PBRs) for each of the four loci. The Z-Test in the Mega 4.0 was used to test for directed selection. Standard errors were obtained with 1,000 bootstrap replicates, including average rates of synonymous (d_s) and nonsynonymous (d_n) substitutions per site according to the Nei–Gojobori method with Jukes–Cantor correction for multiple substitutions [43]. Population allele frequencies and tests of deviation from Hardy–Weinberg equilibrium were calculated using GENEPOP 4.0 [44]. The programme CODEML in the PAML package, version 3.15 [45] was used to test for the presence of codon sites affected by positive selection and to identify those sites in exon 2 sequences of the Chinese egret, and the neutral model M7 and the positive selection model M8 were compared using maximum likelihood ratio tests (LRT).

Results

Characterization of MHC DAB Gene Complex

In this study, DNA sequences of four different MHC DAB genes were isolated from one individual Chinese egret, which was highly similar over large stretches but distinguished by the divergent intron 1. According to Klein [46], these four genes were designated as Egeu-DAB1, Egeu-DAB2, Egeu-DAB3 and Egeu-DAB4 (GenBank Accession KC282841–KC282844). The result of RACE analyses indicated that each of the four genes comprised a total of 780 bp encoding for 260 amino acid residues sequences in exons. This encoded for six of the expected exons, including the signal peptide, the β 1, β 2, transmembrane proteins and the cytoplasmic domains (Figure 1).

When expression analyses were carried out by RT-PCR in six individuals, a total of nine alleles were obtained after sequencing of the cloned RT-PCR products. Four of these alleles were derived from Egeu-DAB1, two from Egeu-DAB2, one from Egeu-DAB3 and two from Egeu-DAB4 (GenBank Accession KC282845–KC282853). The cDNA sequences contained features expected of functional class II molecules, and BLAST searches produced significant hits with known MHC DAB protein sequences. CD-Search identified a β 1 domain in exon 2 and an IGc1 domain (β 2 domain) in exon 3. SMART confirmed these domains, and identified a signal peptide and a trans-membrane domain. The manual survey

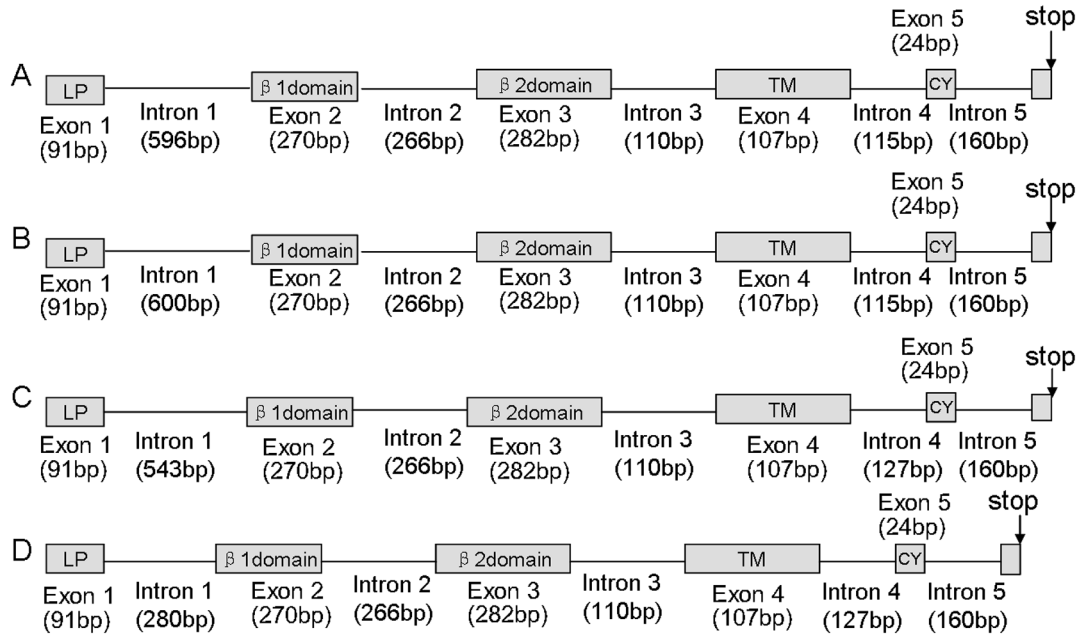


Figure 1. Schematic illustration of MHC II B genes in the Chinese egret. Egeu-DAB1, Egeu-DAB2, Egeu-DAB3 and Egeu-DAB4 are labeled as A, B, C and D respectively. Exons are represented in boxes. Functional domains are indicated in light grey. The highly divergent regions encompassing intron 1. LP, leader peptide; TM, trans-membrane domain; CY, cytoplasmic domain.

doi: 10.1371/journal.pone.0074185.g001

confirmed the conserved intra-domain cysteine salt bridges within $\beta 1$ and $\beta 2$ domains (Figure 2). Additionally, the successful amplification from liver and blood shows that all four genes were expressed in these tissues.

Locus Identification and Polymorphism

Using the primers 25utrF1 and 23exr, 32 of 48 nestlings were amplified, and 22 sequences were obtained from independently replicated individuals. According to the sequence differences in intron 1, 9 of the 22 sequences were derived from Egeu-DAB1, 5 from Egeu-DAB2, 3 from Egeu-DAB3 and the remainder from Egeu-DAB4 (GenBank Accession KC282854-KC282875). This grouping was also confirmed by the neighbor joining analysis (Figure 3). Based on the clustering results, four forward locus-specific primers were designed on intron 1 for each locus, respectively. The SSCP results showed one or two alleles per individual on each locus, certifying the locus separation.

Polymorphism in the MHC DAB Gene Complex

Complete four loci genotypes were obtained for 48 nestlings using a combination of SSCP and direct sequencing for each allele. Nineteen alleles in total were obtained from four loci, including 9 from Egeu-DAB1, 6 from Egeu-DAB2, 1 from Egeu-DAB3 and 3 from Egeu-DAB4 (GenBank Accession KF314166-KF314184). Among the four genes, Egeu-DAB1 had the greatest number of alleles and showed the most variability in nucleotide and amino acid sequences, whereas Egeu-DAB3 was monomorphic on exon 2 (Table 1 and Table S1).

Frequency of Allele and Hardy-Weinberg Equilibrium

Our allele frequency results (Table 2) showed that alleles of each locus had an uneven distribution. On each locus there were one or two alleles had higher frequency than others, such as Egeu-DAB1*05, Egeu-DAB2*02, Egeu-DAB2*05 and Egeu-DAB4*01. The departure from Hardy-Weinberg equilibrium within each locus was statistically significant ($P < 0.001$).

Tests of Selection

When the ratio of non-synonymous to synonymous substitution (d_N/d_S) was analyzed in predicted peptide-binding regions sites (PBR) and remain sites (non-PBR) for each gene separately [42], significant signs of positive selection was only found in the PBR of Egeu-DAB1 alleles, with a d_N/d_S ratio of 2.93 (Z test, $P < 0.05$) (Table 1). Furthermore, it was tested if positive selection was acting on the exon 2 in the Chinese egret. The LRT statistic comparing M7 and M8 models indicated that M8 fitted the data significantly better than M7 ($P < 0.001$). The estimates from M8 suggested that 17 amino acid sites of the exon 2 were under strong positive selection. As shown in Table 3, Egeu-DAB1 had the greatest number of residues with posterior probabilities of positive selection ($P \geq 0.95$). The peptide binding regions (PBR) were slight different from humans [42], but highly consistent with some birds, such as chicken, thin-billed prion and raptor [11, 16].

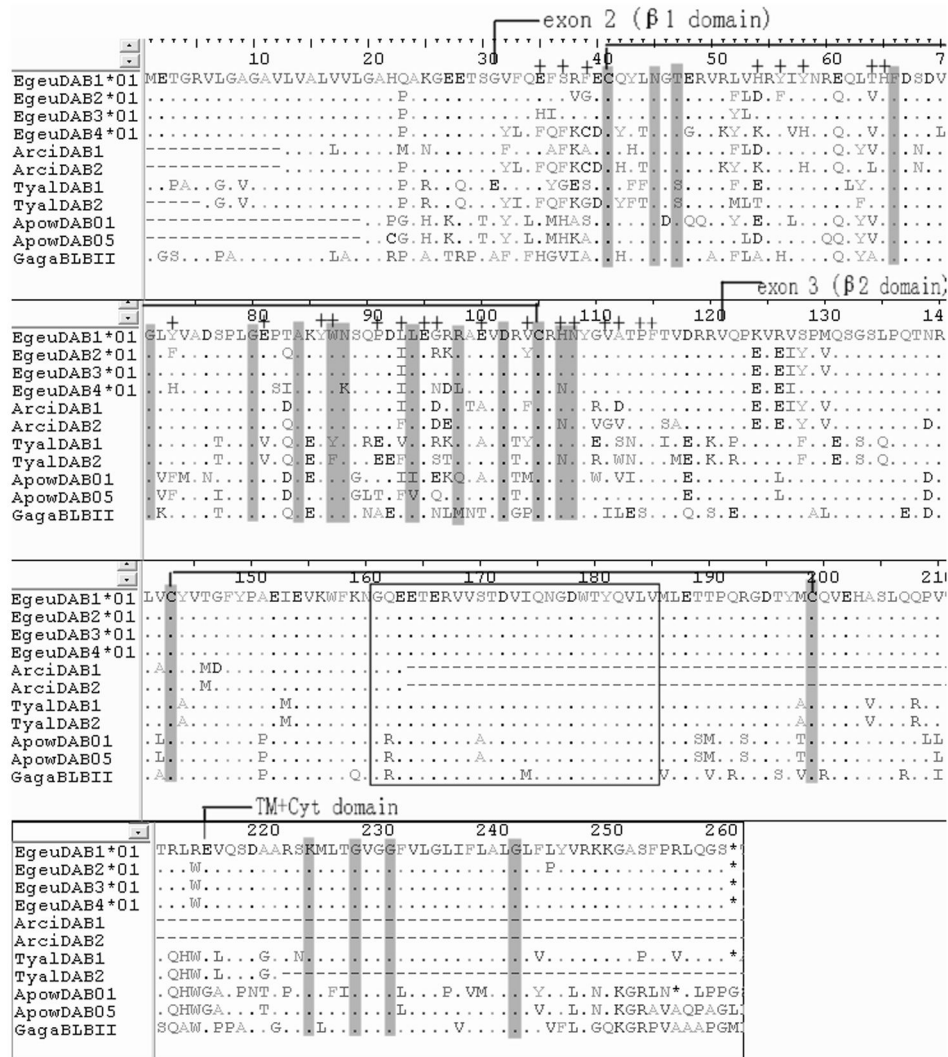


Figure 2. Amino acid alignment of transcribed Chinese egret MHC class II β chain sequences compared with those of other bird species. Dots indicate identity with the Egeu-DAB1*01 sequence, dashes indicate gaps. Conserved residues characteristic of classical class II β chain molecules [41] are shaded. Amino acids that contact the peptide-binding region in the human DR β molecule [42] are indicated with a plus sign (+) on alignment top. Cysteine bridges in the β 1 and β 2 domains are indicated by brackets. Residues in the β 2 domain that are implicated in CD4 binding in humans are boxed. Species and accession numbers for other bird sequences are as follows: ArciDAB1 Grey Heron, HM991016; ArciDAB2 Grey Heron, HM991017; TyalDAB1 Barn Owl, EU442606; TyalDAB2 Barn Owl, EU442607; ApowDAB01 kiwi, HQ639683; ApowDAB05 kiwi, HQ639685; GagaBLBII chicken, NM_001044679.

doi: 10.1371/journal.pone.0074185.g002

Discussion

Using a PCR-based approach, four MHC DAB genes in the Chinese egret were isolated and characterized. Compared with the results of Li's study [33], two loci, Egeu-DAB2 and Egeu-DAB3, were newly discovered in this egret. These four MHC DAB genes were all functional, based on the following findings: (1) they were expressed in several tissues; (2) some sites on PBR were under strong positive selection; (3) no frame shift mutations and stop codons were found in any coding region.

These findings are thus in accord with the postulate that MHC organization is relatively simple, and contains rare pseudogenes in the nonpasserine birds [11,22,47,48].

The dominantly expressed 'major' chicken MHC II β -chain (B-LB) gene showing high sequence variation and high expression levels across tissue types, was previously described [22,49]. Among the four MHC DAB genes in the Chinese egret, genetic diversity in Egeu-DAB1 locus was higher than other three loci, and the number of alleles obtained from this locus by RT-PCR was more than that from the other three loci. Additionally, the

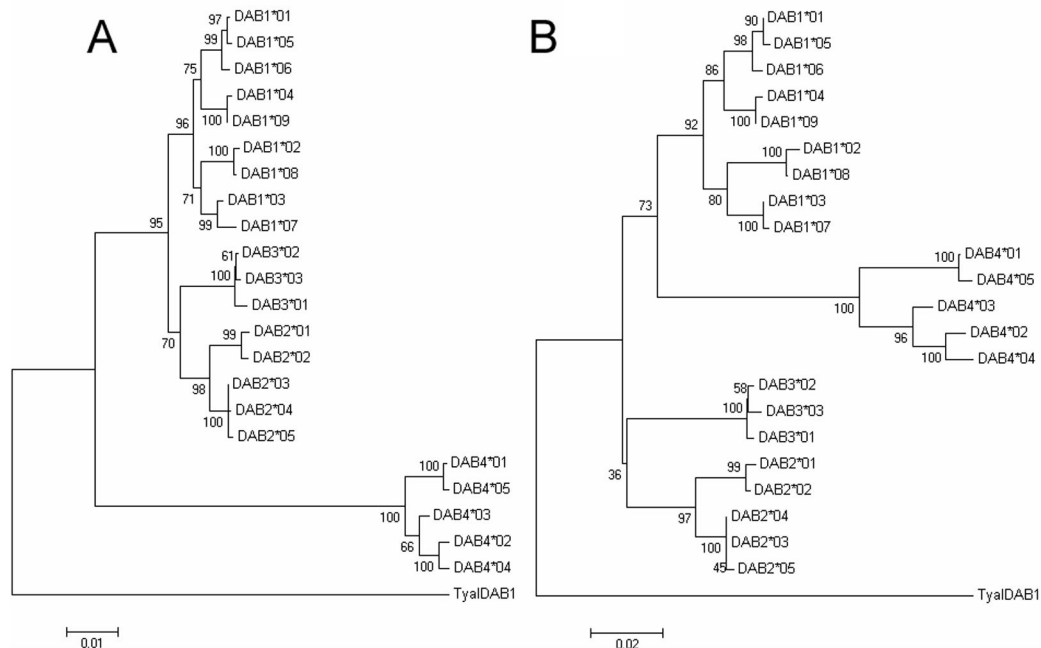


Figure 3. Neighbor joining tree of the 22 alleles obtained from 32 nestlings of the Chinese egret. The two trees are constructed with sequences from 5' UTR to exon 3 (A) and only exons (B), respectively by using the MEGA 4.0 software program. A Barn Owl MHC class II sequence, TyalDAB1 (accession numbers EU442606), served as the outgroup.

doi: 10.1371/journal.pone.0074185.g003

Table 3. Log-likelihood values and parameter estimates of MHC DAB exon 2 alleles in the Chinese egret.

Locus	Model	In L	Parameter estimates	Positively selected sites	2 Δ L(LRT)
DAB01	M7	-829.2402	p = 0.00822, q = 0.01089	Not allowed	
	M8	-802.7880	p0=0.84765, p=0.00500, q=0.00753, (p1=0.15244), ω = 10.80638	5** , 7* , 8** , 9** , 33** , 34** , 53** , 66** , 67** , 70* , 86**	52.9044(P<0.001)
DAB02	M7	-561.4464	p = 0.00745, q = 0.00500	Not allowed	
	M8	-554.4012	p0=0.88891, p=0.03131, q=0.02570, (p1=0.11109), ω = 15.72271	5** , 34** , 57** , 63* , 82**	14.0904(P<0.001)
DAB04	M7	-489.1666	p = 0.00500, q = 0.01175	Not allowed	
	M8	-481.6974	p0=0.79860, p=0.00500, q = 20.92618, (p1=0.20140), ω = 12.41712	27* , 42* , 43* , 53** , 66* , 82**	14.9384(P<0.001)

Sites inferred to be under positive selection at the 95% () and 99% () confidence interval level are indicated. InL, log-likelihood value; ω , selection parameter; p_n , proportion of sites that falls into ω_n site class. The sites consistent with human PBR sites are indicated in *bold*.

d_N/d_S values (all site or PBR site) suggested that Egeu-DAB1 locus was under positive selection. Therefore, these results suggest that Egeu-DAB1 may represent a major class II gene in the Chinese egret. In contrast to the greater sequence variation of Egeu-DAB1, Egeu-DAB3 showed extremely limited variation and was monomorphic in the PBR-site, suggesting this locus may represent a minor DAB gene [49]. Moreover, CODEML analyses revealed 17 residues with posterior probabilities of positive selection ($P \geq 0.95$) in MHC DAB exon 2, indicating that positive selection played a part in the adaptive evolution of this egret.

Many studies on characterization of the MHC gene complex have, as a major goal, identification of single MHC gene for population genetic studies. However, it is difficult to identify individual genes because the molecular mechanisms that

underlie MHC evolution, including selection and concerted evolution, can obscure the genealogical relationships among alleles [50]. Segregation analyses using individuals with well-defined pedigrees may aid in clarifying the number of loci and assigning sequences to loci. However, such pedigree information is not available for wild birds such as the Chinese egret. Therefore, using noncoding and conserved regions as an indicator of which sequences representing individual loci, is an acceptable option, as it is less variable and not subject to balancing selection. Efforts to sequence noncoding regions of MHC loci can be rewarded by the ability to characterize individual loci, and can enhance the utility of these genes as molecular markers in evolutionary ecology research. This experimental strategy also contributes greatly to the study of MHC gene family evolution across the breadth of avian

Table 1. Summary of sequence variation of MHC DAB in the Chinese egret.

	S _{nt}	S _{AA}	π	d _N	d _S	d _N /d _S	p	z
PBR sites								
Egeu-DAB1	27	14	0.198±0.036	0.237±0.063	0.081±0.042	2.93	0.019	2.033
Egeu-DAB2	17	8	0.148±0.033	0.153±0.043	0.142±0.058	1.08	0.693	0.279
Egeu-DAB3	-	-	-	-	-	-	-	-
Egeu-DAB4	13	7	0.129±0.037	0.122±0.052	0.159±0.070	0.77	0.477	-0.634
Non-PBR sites								
Egeu-DAB1	32	18	0.067±0.012	0.068±0.018	0.057±0.022	1.19	0.650	0.455
Egeu-DAB2	16	12	0.042±0.011	0.043±0.012	0.042±0.022	1.38	0.981	0.024
Egeu-DAB3	-	-	-	-	-	-	-	-
Egeu-DAB4	18	10	0.065±0.015	0.053±0.018	0.076±0.037	0.70	0.598	-0.529

Total size 270 bp (90 residues) for all sites, 72 bp (24 residues) for PBR sites and 198 bp (66 residues) for non-PBR sites; S number of variable nucleotide (nt) or amino acid (AA) sites; π nucleotide diversity ± SE; d average rate of nonsynonymous (d_N) or synonymous (d_S) substitutions per site ± SE; z the Z test of positive selection.

Table 2. Allele frequency of MHC DAB in the Chinese egret.

Allele\ Locus	Egeu-DAB1	Egeu-DAB2	Egeu-DAB3	Egeu-DAB4
01	0.0729	0.0833	1.0000	0.6250
02	0.0208	0.2917		0.2083
03	0.0208	0.0833		0.1667
04	0.1042	0.0833		
05	0.4167	0.2708		
06	0.1354	0.1875		
07	0.1042			
08	0.0417			
09	0.0833			

diversity. In this study, intron 1 was used to show that this intron was more divergent than the other regions of the MHC

References

- Reichard M, Spence R, Bryjová A, Bryja J, Smith C (2012) Female rose bitterling prefer MHC-dissimilar males: experimental evidence. *PLOS ONE* 7: e40780. doi:10.1371/journal.pone.0040780. PubMed: 22815816.
- Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front Zool* 2: 16. doi:10.1186/1742-9994-2-16. PubMed: 16242022.
- Ujvari B, Belov K (2011) Major histocompatibility complex (MHC) markers in conservation biology. *Int J Mol Sci* 12: 5168-5186. doi:10.3390/ijms12085168. PubMed: 21954351.
- Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012) Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat Communications* 3: 621. doi:10.1038/ncomms1632. PubMed: 22233631.

DAB genes, suggesting it may be the best candidate for locus specific primers in single locus typing of MHC class II in the Chinese egret. The low number of loci in this egret will allow us to completely genotype each locus of the MHC II B in individual egrets. Moreover, the primers designed for this study are targeting highly conserved regions across class II genes, and therefore, similar fragments in other Ardeid species are likely to be cross-amplified successfully.

Supporting Information

Table S1. Genotyping data collected from 4 MHC DAB loci in the Chinese egret. The null alleles are indicated by dots. (DOC)

Acknowledgements

We thank Professor Frederic A. Troy II of the University of California (Davis) School of Medicine, for help in reviewing the manuscript. We also thank Peng Cai, Yufei Dai, Wei Lei and Jing Lin for help in collecting some samples and Bo Wang and Kaijian Zhu for assistance in field work of this study.

Author Contributions

Conceived and designed the experiments: XC ZW WF. Performed the experiments: ZW QL. Analyzed the data: ZW XZ XC WF. Contributed reagents/materials/analysis tools: XC WF XZ. Wrote the manuscript: ZW XZ XC WF. Edited manuscript: XC WF. Graduate Supervisor: XC WF. Funding Source: XC WF XZ.

5. Hess CM, Edwards SV (2002) The evolution of the major histocompatibility complex in birds. *BioScience* 52: 423-431. doi: 10.1641/0006-3568(2002)052[0423:TEOTMH]2.0.CO;2.
6. Babik W (2010) Methods for MHC genotyping in non-model vertebrates. *Mol Ecol Resour* 10: 237-251. doi:10.1111/j.1755-0998.2009.02788.x. PubMed: 21565019.
7. Aguilar A, Edwards SV, Smith TB, Wayne RK (2006) Patterns of variation in MHC class II β loci of the little greenbul (*Andropadus virens*) with comments on MHC evolution in birds. *J Hered* 97: 133-142. doi: 10.1093/jhered/esj013. PubMed: 16489149.
8. Ekblom R, Grahn M, Höglund J (2003) Patterns of polymorphism in the MHC class II of a non-passerine bird, the great snipe (*Gallinago media*). *Immunogenetics* 54: 734-741. PubMed: 12557060.
9. Gasper JS, Shiina T, Inoko H, Edwards SV (2001) Songbird Genomics: Analysis of 45 kb Upstream of a Polymorphic MHC Class II Gene in Red-Winged Blackbirds (*Agelaius phoeniceus*). *Genomics* 75: 26-34. doi:10.1006/geno.2001.6596. PubMed: 11472064.
10. Edwards SV, Grahn M, Potts WK (2008) Dynamics of Mhc evolution in birds and crocodylians: amplification of class II genes with degenerate primers. *Mol Ecol* 4: 719-730. PubMed: 8564010.
11. Alcaide 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. doi:10.1007/s00239-007-9033-9. PubMed: 17925996.
12. Westerdahl H, Wittzell H, Von Schantz T, Bensch S (2004) MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGE. *Heredity* 92: 534-542. doi:10.1038/sj.hdy.6800450. PubMed: 15162116.
13. Alcaide M, Edwards SV, Cadahía L, Negro JJ (2009) MHC class I genes of birds of prey: isolation, polymorphism and diversifying selection. *Conserv Genet* 10: 1349-1355. doi:10.1007/s10592-008-9653-7.
14. Bonneaud C, Sorci G, Morin V, Westerdahl H, Zoorob R et al. (2004) Diversity of Mhc class I and IIB genes in house sparrows (*Passer domesticus*). *Immunogenetics* 55: 855-865. doi:10.1007/s00251-004-0648-3. PubMed: 14963619.
15. Mardon J, Bonadonna F (2009) Atypical homing or self-odour avoidance? Blue petrels (*Halobaena caerulea*) are attracted to their mate's odour but avoid their own. *Behav Ecol Sociobiol* 63: 537-542. doi:10.1007/s00265-008-0688-z.
16. Silva MC, Edwards SV (2009) Structure and evolution of a new avian MHC class II B gene in a sub-Antarctic seabird, the thin-billed prion (*Procellariiformes: Pachyptila belcheri*). *J Mol Evol* 68: 279-291. doi: 10.1007/s00239-009-9200-2. PubMed: 19209378.
17. Ekblom R, Saether SA, Fiske P, Kålås JA, Höglund J (2010) Balancing selection, sexual selection and geographic structure in MHC genes of Great Snipe. *Genetica* 138: 453-461. doi:10.1007/s10709-008-9335-x. PubMed: 19052880.
18. Strandh M, Lannefors M, Bonadonna F, Westerdahl H (2011) Characterization of MHC class I and II genes in a subantarctic seabird, the blue petrel, *Halobaena caerulea* (Procellariiformes). *Immunogenetics* 63: 653-666. doi:10.1007/s00251-011-0534-8. PubMed: 21607694.
19. Balakrishnan CN, Ekblom R, Völker M, Westerdahl H, Godinez R et al. (2010) Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biol* 8: 29. doi: 10.1186/1741-7007-8-29. PubMed: 20359332.
20. Bollmer JL, Dunn PO, Whittingham LA, Wimpee C (2010) Extensive MHC class II B gene duplication in a passerine, the common yellowthroat (*Geothlypis trichas*). *J Hered* 101: 448-460. doi:10.1093/jhered/esq018. PubMed: 20200139.
21. Westerdahl H, Wittzell H, von Schantz T (2000) Mhc diversity in two passerine birds: no evidence for a minimal essential Mhc. *Immunogenetics* 52: 92-100. doi:10.1007/s002510000256. PubMed: 11132162.
22. 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. doi:10.1038/44856. PubMed: 10553909.
23. Kaufman J, Volk H, Wallyn HJ (2006) A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunol Rev* 143: 63-88.
24. 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. PubMed: 15153492.
25. Westerdahl H (2007) Passerine MHC: genetic variation and disease resistance in the wild. *J of Ornithol* 148: 469-477. doi:10.1007/s10336-007-0230-5.
26. Castillo S, Srithayakumar V, Meunier V, Kyle CJ (2010) Characterization of major histocompatibility complex (MHC) DRB exon 2 and DRA exon 3 fragments in a primary terrestrial rabies vector (*Procyon lotor*). *PLOS ONE* 5: e12066. doi:10.1371/journal.pone.0012066. PubMed: 20706587.
27. Cloutier A, Mills JA, Baker AJ (2011) Characterization and locus-specific typing of MHC class I genes in the red-billed gull (*Larus scopulinus*) provides evidence for major, minor, and nonclassical loci. *Immunogenetics* 63: 377-394. doi:10.1007/s00251-011-0516-x. PubMed: 21327606.
28. BirdLife International (2011) Species factsheet: *Egretta eulophotes*. Available: <http://www.birdlife.org>. Accessed: 2013, August 5
29. IUCN (2012) IUCN Red List of Threatened Species, version 2011.1. Available: www.iucnredlist.org. Accessed: 2013, August 5.
30. Zhou X, Wang Y, Chen X, Lin Q, Fang W et al. (2008) A set of primer pairs for amplifying the complete mitochondrial DNA of endangered Chinese egret (Aves, Ardeidae, *Egretta eulophotes*). *Mol Ecol Resour* 8: 412-414. doi:10.1111/j.1471-8286.2007.01975.x. PubMed: 21585806.
31. Huang X, Zhou X, Chen M, Fang W, Chen X (2010) Isolation and characterization of microsatellite loci in vulnerable Chinese egret (*Egretta eulophotes*: Aves). *Conservation Genetics* 11: 1211-1214.
32. Wang Z, Zhou X, Lin Q, Fang W, Chen X (2011) New primers for sex identification in the Chinese Egret and other ardeid species. *Mol Ecol Resour* 11: 176-179. doi:10.1111/j.1755-0998.2010.02879.x. PubMed: 21429119.
33. 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. doi:10.1007/s00239-011-9446-3. PubMed: 21590337.
34. Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85: 8998-9002. doi:10.1073/pnas.85.23.8998. PubMed: 2461560.
35. Fain MA, Zhao T, Kindt TJ (2001) Improved typing procedure for the polymorphic single-copy RLA-DQA gene of the rabbit reveals a new allele. *Tissue Antigens* 57: 332-338. doi:10.1034/j.1399-0039.2001.057004332.x. PubMed: 11380942.
36. Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC (1991) Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am J Hum Genet* 48: 137-144. PubMed: 1670750.
37. Bradley RD, Hillis DM (1997) Recombinant DNA sequences generated by PCR amplification. *Mol Biol Evol* 14: 592-593. doi:10.1093/oxfordjournals.molbev.a025797. PubMed: 9159937.
38. Meyerhans A, Vartanian JP, Wain-Hobson S (1990) DNA recombination during PCR. *Nucleic Acids Res* 18: 1687-1691. doi: 10.1093/nar/18.7.1687. PubMed: 2186361.
39. Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268: 78-94. doi:10.1006/jmbi.1997.0951. PubMed: 9149143.
40. Hall TA (1997) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp S No. 41*: 95-98.
41. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599. doi:10.1093/molbev/msm092. PubMed: 17488738.
42. 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: 33-33. doi:10.1038/364033a0. PubMed: 8316295.
43. Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418-426. PubMed: 3444411.
44. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8: 103-106. doi:10.1111/j.1471-8286.2007.01931.x. PubMed: 21585727.
45. Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 17: 32-43. doi:10.1093/oxfordjournals.molbev.a026236. PubMed: 10666704.
46. Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllenstein UB et al. (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics* 31: 217-219. PubMed: 2329006.
47. Zoorob R, Béhar G, Kroemer G, Auffray C (1990) Organization of a functional chicken class II B gene. *Immunogenetics* 31: 179-187. PubMed: 1969383.
48. 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. doi:10.1093/molbev/msn065. PubMed: 18359775.

49. Jacob JP, Milne S, Beck S, Kaufman J (2000) The major and a minor class II beta-chain (B-LB) gene flank the Tapasin gene in the BF/B-L region of the chicken major histocompatibility complex. *Immunogenetics* 51: 138–147. doi:10.1007/s002510050022. PubMed: 10663576.
50. Miller HC, Andrews-Cookson M, Daugherty CH (2007) Two patterns of variation among MHC Class I loci in Tuatara (*Sphenodon punctatus*). *J Hered* 98: 666–677. doi:10.1093/jhered/esm095. PubMed: 18032462.