Journal of Heredity 2006:97(2):133–142
doi:10.1093/jhered/esj013
Advance Access publication February 17, 2006

Patterns of Variation in MHC Class II β Loci of the Little Greenbul (*Andropadus virens*) with Comments on MHC Evolution in Birds

ANDRES AGUILAR, SCOTT V. EDWARDS, THOMAS B. SMITH, AND ROBERT K. WAYNE

From the Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095 (Aguilar, Smith, and Wayne); the Center for Tropical Research, Institute of the Environment, 1609 Hershey Hall, University of California, Los Angeles, CA 90095 (Aguilar, Smith, and Wayne); and the Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138 (Edwards). Andres Aguilar is now at the Department of Ocean Sciences and Southwest Fisheries Science Center, 110 Shaffer Road, University of California, Santa Cruz, CA 95060.

Address correspondence to A. Aguilar at the address above, or e-mail: andres.aguilar@noaa.gov.

Abstract

We have isolated major histocompatibility complex (MHC) class II β loci from the little greenbul (*Andropadus virens*), an African songbird. We utilized preexisting information about conserved regions of the avian MHC to design primers to amplify a pool of sequences representing multiple loci. From this pool, a unique locus spanning 1109 bp that we designate as *Anvi*-DAB1 was cloned and sequenced. We designed locus-specific primers based on this sequence information and amplified six alleles from seven individuals. Compared to other *A. virens* MHC sequences obtained from genomic DNA or cDNA, the variability of sequences from *Anvi*-DAB1 was low and the ratio of nonsynonymous to synonymous substitution was much less than one, suggesting that *Anvi*-DAB1 may either be a pseudogene or a nonclassical MHC locus. Phylogenetic analysis revealed that the *Anvi*-DAB1 locus was highly divergent when compared with other passerine or *A. virens* genomic or transcribed MHC sequences. The use of conserved MHC primers followed by analysis of cloned sequences allows rapid isolation of MHC loci from exotic species and avoids laborious large-scale cloning and sequencing.

The major histocompatibility complex (MHC) is a 0.1- to 4-Mb segment of DNA in vertebrates and can contain over 80 genes arranged in three functional classes (Kaufman et al. 1999; Klein 1986; Trowsdale 1995). Class I genes present endogenous antigens, whereas class II MHC genes are responsible for the presentation of exogenous antigens to the immune system. This presentation of antigens to the immune system is integral to cell-mediated immune response (Klein 1986). The MHC has received considerable attention from evolutionary biologists as natural selection is widely thought to operate on MHC genes (Edwards and Hedrick 1998; Hedrick 1994). Evidence suggesting that these genes are under selection include (1) high allelic diversity (Robinson et al. 2000), (2) levels of heterozygosity that often exceed those predicted by neutrality (Black and Salzano 1981), (3) relatively long coalescent times compared to neutral expectations (Klein et al. 1993; Takahata 1990), and (4) increased nonsynonymous to synonymous substitution ratios for active portions of genes (Hughes and Nei 1988, 1989).

A growing body of research has revealed that concerted evolution is common in the avian MHC (Edwards et al. 1995b; Miller and Lambert 2004; Wittzell et al. 1999) and that genomic organization differs among avian taxa (Hess et al. 2000; Kaufman et al. 1999; Shiina et al. 1999). The entire MHC of the chicken is less than 100 kb and has a relatively simple organization with small introns and few pseudogenes (Kaufman et al. 1999). This organization has been designated the minimal essential MHC (Kaufman et al. 1995) and refers to a gene family that is highly reduced in physical size, gene copy number, and expression and contains few nonfunctional loci. Observations on the chicken MHC have led to suggestions that the avian MHC differs dramatically from the mammalian MHC as the latter is often large (4 Mb), can vary in gene copy number, and possesses many pseudogenes (MHC Sequencing Consortium 1999; Milner and Campbell 1992; Trowsdale 1995). However, limited research on songbirds indicates that the minimal essential MHC may not apply to this group as many transcribed and genomic

copies of MHC genes have been discovered within single genomes (Edwards et al. 1998, 2000; Freeman-Gallant et al. 2002; Hess et al. 2000; Westerdahl et al. 2000; Wittzell et al. 1999).

Most research on the vertebrate MHC has been done on mammals (Trowsdale 1995). Studies on the avian MHC are restricted to relatively few taxa, primarily gallinaceous birds (chicken, quail, and pheasant) (Kaufman et al. 1999; Shiina et al. 1999; Wittzell et al. 1999; Zoorob et al. 1993) and several passerines (Edwards et al. 1995a, 1995b, 1998, 2000; Gasper et al. 2001; Hess et al. 2000). Surveys of MHC variation in natural populations are often facilitated by the close phylogenetic relationship of the target taxa to a model species for which there is detailed knowledge of the MHC. For example, studies on wild canids (Hedrick et al. 2000, 2002) have been assisted by previous work on the domestic dog that allowed the design of locus-specific polymerase chain reaction (PCR) primers (Wagner 2003). In contrast, the investigation of MHC variation in exotic species without well-studied close relatives can be laborious and time consuming and often involves the development, screening, and sequencing of cDNA or genomic libraries (see Edwards et al. 1998, 2000 or Hess et al. 2000).

Here, we use knowledge of conserved regions within the avian MHC to design PCR primers to amplify and sequence a pool of MHC loci from genomic DNA as well as transcripts from class II β genes. Putative single-locus amplifications are confirmed by analysis of single-strand conformational polymorphisms (SSCP; Orita et al. 1989). We use this rapid approach to isolate MHC class II β chain loci in the little greenbul (*Andropadus virens*), an African nonmigratory rainforest passerine (Keith et al. 1992). This species has been the subject of studies in speciation and geographic variation (Smith et al. 1997, 2000, 2001) and pathogen incidence (Kirkpatrick and Smith 1988; Sehgal et al. 2001) and hence is an interesting species in which to characterize MHC variation. We analyzed MHC sequence variation in the little greenbul and used phylogenetic techniques to elucidate functional significance and the evolutionary modes of MHC diversification in this and other avian taxa.

Methods

Gene Isolation

Previous research on passerine MHC (Edwards et al. 1995a,b) has revealed conserved regions of MHC class II β genes. We used knowledge of these conserved regions to obtain sequences from the MHC class II β genes of *A. virens* (see Potts 1996). We focused on the regions surrounding the peptide-binding region because this functional domain is involved in antigen presentation and is highly polymorphic (Hughes and Nei 1988, 1989; Klein 1986). Using conserved primers, we amplified two separate fragments of the class II β gene from a single individual *A. virens*. The first fragment amplified with primers 1a (ATG GGA CCC CAA AAG TGA TT) and 2a (CCG AGG GGA CAC GCT CT) contained a segment of intron 1 as well as exon 2 (fragment A; Figure 1). A second

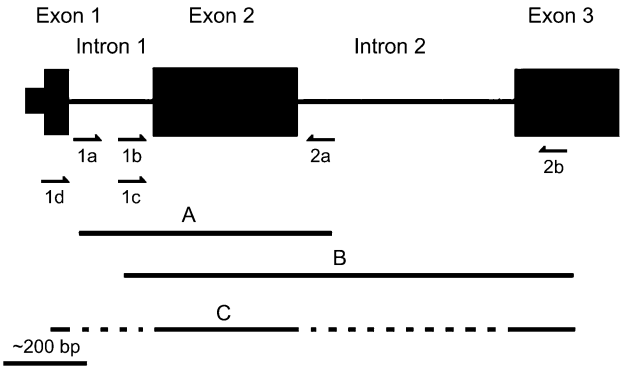

Figure 1. PCR-based strategy used to obtain exon and intron sequences for MHC class II β loci in songbirds. Primer names are as follows: 1a (Int1f.7), 1b (AnviEX2f.1), 1c (AnviEX2f.2), 1d (Songex1F.2), 2a (Int2r.1), and 2b (SongEX3R.1). Primer 1a was from Edwards SV (unpublished data) and 2a was from Edwards et al. (1998). Primers 1b and 1c (*Anvi*-DAB1 specific) were designed from conserved regions of *Andropadus virens* intron 1 sequences. Primers 1d and 2b were designed from conserved regions found in published exon 1 and exon 3 sequences of passerines. The two regions amplified from genomic DNA are shown as fragments A and B. Fragment C is the cDNA sequence of MHC transcripts amplified with primers 1d and 2b.

fragment spanning the entire region from exon 2 through exon 3 (fragment B; Figure 1) was amplified using forward primer 1b designed from fragment A sequences and primer 2b based on previously described exon 3 sequences.

Analysis of cloned PCR sequences from these two fragments allowed us to identify a unique locus, *Anvi*-DAB1, and design locus-specific primers that could be used to characterize variation in a sample of seven little greenbuls (see Results).

Samples of *A. virens* blood used for screening by PCR were collected by T. B. Smith in Cameroon and stored in lysis buffer (Seutin et al. 1991). Blood was digested overnight with Proteinase K and DNA extracted with standard phenol-chloroform protocols (Sambrook et al. 1989). The resulting DNA was suspended in Tris-EDTA. All PCR conditions contained ~20 ng genomic DNA, 20 pmol of forward and reverse primer, 1× Sigma PCR buffer, 1 mM deoxynucleoside triphosphates (dNTPs), and 2.5 mM $MgCl_2$. Fragment A had a PCR temperature profile of 3-min initial denaturing at 94°C and 30 cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C with a final extension of 5 min at 72°C. Fragment B had a PCR temperature profile of 3-min initial denaturing at 94°C and 30 cycles of 1 min at 94°C, 1 min at 60°C, and 1.5 min at 72°C with a final extension of 5 min at 72°C. PCR products were run on 1% agarose gels to confirm amplification, and then products corresponding to the correct size were gel purified using the UltraClean Kit (Mobio Laboratories Inc., Carlsbad, CA). DNA was ligated into the TA or TOPO-TA cloning vector (Invitrogen Corp., Carlsbad, CA) and transformed

into TOP10 chemically competent cells (Invitrogen Corp, Carlsbad, CA). Colonies containing inserts of the correct size were sequenced in one of two ways. In the first method, a colony was selected and boiled in 50 μl of H_2O for 5 min. Two microliters of the boiled colony was used in a PCR containing M13 forward and reverse primers. The resulting product was then sequenced on an ABI 377 using M13 primers. For the second method, plasmid DNA was isolated from positive clones using the Qiagen Plasmid kit (Qiagen Inc., Valencia, CA). The resulting plasmid DNA was then sequenced with M13 forward and reverse primers on an ABI 377 following manufacturer's protocols. Larger inserts often required sequencing with an internal primer (available on request).

Polymorphism Survey

With the acquisition of intron 1 sequences for little greenbul MHC class II β genes, a set of forward primers at the 3′ end of intron 1 [Figure 1: primer 1b (TGC CAT GGA CGC TTA CAC T) and primer 1c (CTG TGT CCT GCA YTC AGG)] were designed to be used in conjunction with primer 2a to amplify two distinct sequence groups. Primer 1b was designed to specifically amplify sequences from a single locus that we designate *Anvi*-DAB1. This primer was based on a conserved motif shared by a divergent group of similar sequences spanning exon 2 (Figure 2, see Results). Similarly, the second forward primer (1c) amplified a set of divergent sequences that we designate as type I which are characterized by unique substitutions and an indel in intron 1 (clones B1, 7, 8, 9, and 10; Figure 2). However, this primer amplified multiple loci and was not used to further investigate variability within our population sample of little greenbuls (see Results).

To assess the total number of alleles amplified by primer combination 1b and 2a, SSCP (Orita et al. 1989) was preformed on PCR products from seven little greenbuls. Briefly, both primers were end labeled with ^{32}P (Sambrook et al. 1989), and these radio-labeled primers were used in a PCR with 10 ng genomic DNA, 20 pmol of each primer, 1 mM dNTPs, 1× PCR buffer (Sigma), 0.5 units of *Taq* polymerase (Sigma), and 1.0 mM $MgCl_2$ in a 25-μl reaction volume. The reaction conditions were as follows: an initial 3-min denaturing step at 94°C, 30 s at 94°C, 30 s at 58°C, 30 s at 72°C, and a final 5-min extension at 72°C. Five microliters of the PCR was mixed with 2 μl of stop solution (95% formamide and 0.05% bromophenol blue) heated for 5 min at 95°C and then cooled immediately in ice water. Two microliters of this mixture was loaded on a 5% nondenaturing polyacrylamide gel containing 5% glycerol (v/v) and run at 20 W for 8–10 h at room temperature. Gels were transferred to 3M Whatman paper, dried, and exposed to autoradiographic film for 12–48 h (depending on the activity of ^{32}P). Unique alleles, identified from SSCP, were isolated from dried gels, reamplified, and sequenced. Bands having the same confirmation were sequenced from multiple individuals to confirm the identity of alleles.

Cross-species amplification of exon 2 from *Anvi*-DAB1 was attempted on individuals from species closely related to *A. virens*, including *Andropadus curvirostris*, *Andropadus latirostris*, *Andropadus gracilis*, *Andropadus montanus*, and *Andropadus tephrolaemus*. PCR conditions were the same as described for *A. virens*.

Confirmation of Transcription

To determine if *Anvi*-DAB1–like sequences were transcribed and to compare transcripts to genomic MHC sequences, class II β fragments were amplified from *A. virens* cDNA. A spleen was taken from a single individual in the field and placed in RNAlater buffer (Ambion, Austin, TX) for transport to −80°C storage. Total RNA was isolated from approximately 0.5 g of spleen using the Qiagen extraction kit. Total cDNA was generated using Superscript II (Gibco-BRL, Carlsbad, CA) following manufacturer's protocols. The isolated cDNA was then digested with ribonuclease, precipitated, and subsequently resuspended in H_2O. A set of previously published primers designed for conserved regions within exon 1 and exon 3 was used to amplify the entire exon 2 from cDNA [Figure 1: primer 1d (GCA CTG GTG GTG CTG GGA GC) and primer 2b (TGG AAC CAC CTC ACC TGG AT)]. Amplification was done in 50-μl volumes with 1 μl of first-round cDNA, 1 unit of Amplitaq Gold (Applied BioSystems, Foster City, CA), 5 μl of 10× Amplitaq Gold buffer (Applied BioSystems), 3.0 mM $MgCl_2$, 20 pmol of forward and reverse primer, and 1 mM of dNTPs. Amplified fragments were cloned into a TA cloning vector (Invitrogen Corp.), and 20 clones were sequenced on a Beckman CEQ2000 using standard M13 forward and reverse primers. All sequences have been deposited in GenBank (AY437889–AY437914).

Data Analysis

Representative sequences of the putative *A. virens* MHC class II β genes were used in BLAST searches (Altschul et al. 1997) on GenBank (www.ncbi.nlm.nih.gov) to assess identity to other published avian MHC sequences. All sequences were imported into SEQUENCHER (GeneCodes, Ann Arbor, MI) and aligned manually to previously published avian MHC sequences. We used published class II β sequences from isolated genes of the domestic chicken (*Gallus gallus*), pheasant (*Phasianus colchis*), red-winged blackbird (*Agelaius phonecius*), house finch (*Carpodacus mexicanus*), scrub jay (*Aphelocoma coerulescens*), Bengalese finch (*Lonchura striata*), and reed warbler (*Acrocephalus arundinaceus*) (Table 1). A sequence from a human class II β gene was used as an outgroup when applicable.

Nucleotide substitution models were evaluated for the exon 2, exon 3, and intron 1 alignments, with the likelihood ratio test implemented in MODELTEST3.5 (Posada and Crandall 1998). Phylogenetic trees were constructed using a Bayesian framework with the estimated substitution model for exon 2 and exon 3 alignments with MRBAYES3 (Huelsenbeck and Ronquist 2001). Nonsynonymous and synonymous substitutions were calculated using the method of Nei and Gojobori in MEGA3 (Kumar

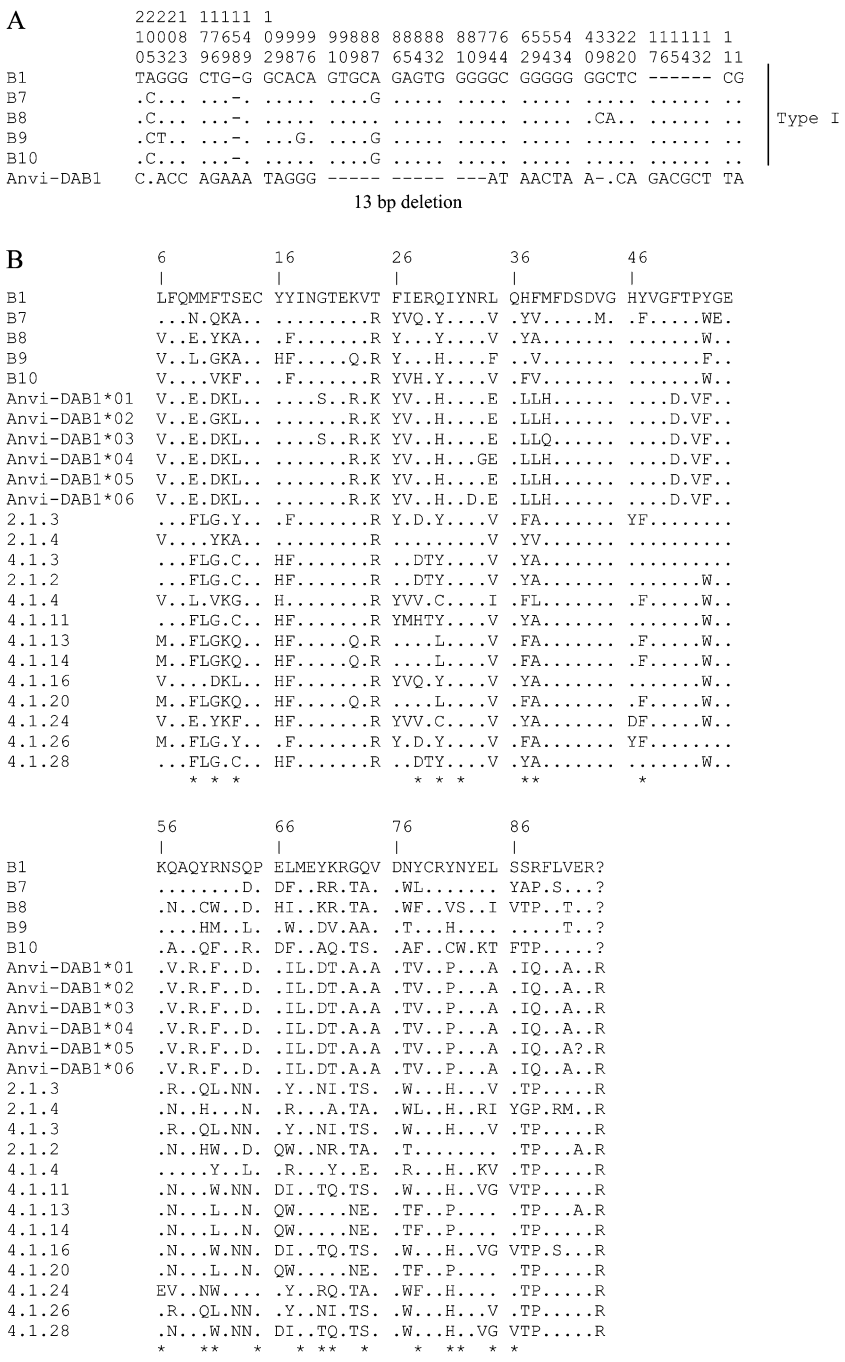

Figure 2. Intron 1 DNA sequences and inferred exon 2 amino acid sequences. (**A**) Aligned variable portions from intron 1. Only variable sites are shown, and they are given as the number of base pairs upstream from the beginning of exon 2. Clones B1, B7, B8, B9, B10, and *Anvi*-DAB1 are from a single *Andropadus virens* individual. (**B**) Translated exon 2 sequences. Clones B1, B7, B8, B9, B10, and *Anvi*-DAB1 were isolated from genomic DNA, whereas the remainder of the sequences were unique cDNA sequences. Codons involved in antigen presentation (from humans, Brown et al. 1993) are denoted by an "*," whereas a "?" indicates an ambiguous site and a "." indicates identity in nucleotide position.

et al. 2004) separately for codons in the antigen-binding sites (ABS) and non-ABS (Brown et al. 1993). A *Z* test was performed for the null hypothesis that the nonsynonymous divergence equals the synonymous divergence (Nei and Kumar 2000). The average number of pairwise differences or nucleotide diversity (π) and Tajima's *D* (Tajima 1989) were calculated using DNASP (v3.53) (Rosas J and Rosas R 1999).

Table 1. DNA sequences of avian and nonavian taxa used in the phylogenetic analyses

Species	GenBank accession numbers	Code	Portion of gene	
			Exon 2	Exon 3
Agelaius phoeniceus	U23970	*Agph*-1.1	×	×
	U23971	*Agph*-1.2	×	×
	U23967	*Agph*-1.3	×	×
	AF328738	*Agph*-DAB1	×	×
	AF170972	*Agph*-DAB2	×	×
	AF328738	*Agph*-DAB3	×	—
Aphelocoma coerulesens	U23972	*Apco*-1.1	×	×
	U23973	*Apco*-1.2	×	×
	U23958	*Apco*-2.2	×	—
	U23960	*Apco*-2.4	×	×
	U23963	*Apco*-3.3	×	×
	U23965	*Apco*-3.5	×	×
	U23966	*Apco*-3.6	×	×
Acrocephalus arundinaceus	AJ404372	*Acar*-c01	×	×
	AJ404371	*Acar*-c02	×	×
	AJ404373	*Acar*-c022	×	×
	AJ404374	*Acar*-c023	×	×
Carpodacus mexicanus	U23968	*Came*-1.1	×	×
	U23969	*Came*-1.2	×	×
	U23976	*Came*-2.1	×	×
	AF205032	*Came*-DAB1	×	×
Gallus gallus	U91530	B-LBI	×	—
	AJ248573	B-LBII	×	×
	AJ248574	B-LB4	×	×
	AJ248576	B-LB12	—	×
	AJ248580	B-LB14	×	×
	AJ248584	B-LB19	—	×
Lonchura striata	L42335	*Lost*-05	×	×
	L42334	*Lost*-26	×	×
Phasianus colchicus	AJ224349	*Phco*-DAB1	×	×
	AJ224348	*Phco*-DAB2	×	×
Homo sapiens	NM002124	HLA-DRB1	×	—

A "×" indicates that the sequence was used, and a "—" indicates that it was not used in the phylogenetic analysis.

Results

We amplified two fragments, A and B, in *A. virens* (Figure 1) and found a subset of sequences that defined a single 1109-bp contig that contained partial intron 1, exon 2, intron 2, and partial exon 3 sequences (Figure 3). This gene fragment contained a 13-bp deletion in intron 1 that was not found in other gene fragments (Figure 2). Consequently, we designated this locus as *Anvi*-DAB1 according to the established procedures of MHC nomenclature (Klein et al. 1990). We designed a locus-specific primer 1b in intron 1 that was used in combination with primer 2a to amplify exon 2 of *Anvi*-DAB1 (Figure 1). Using SSCP, we found six distinct alleles in seven *A. virens*. Only one or two alleles were found for each of the seven individuals, suggesting that this primer pair amplified a single locus (Figures 4 and 5). One allele had a frameshift mutation within exon 2 (*Anvi*-DAB1*05: Figure 4). However, no stop codons were found within the reading frame for any of the alleles. Exon 2 of *Anvi*-DAB1 had a low level of nucleotide diversity (π) per site equal to 0.007. In contrast, nucleotide diversities of 0.03–0.18 have been observed for most avian MHC class II loci (Hess et al. 2000; Zoorob et al. 1993). A Tajima's D value of -0.46 suggested an excess of rare mutations in exon 2 of *Anvi*-DAB1. However, this D value is not statistically different from the neutral expectation.

Attempts to amplify exon 2 of *Anvi*-DAB1 from other closely related *Andropadus* species by lowering PCR stringency through a decrease in annealing temperature were not successful. This failure to amplify may indicate that exon 2 of *Anvi*-DAB1 is not present in these taxa or has diverged in flanking intron sequences.

Other MHC Class II β Gene Fragments

The sequences of five clones of fragment A shared unique similarities and were divergent from fragment A sequences of *Anvi*-DAB1 (Figure 2). We refer to this divergent group of sequences as type I sequences. Type I and *Anvi*-DAB1 intron 1 sequences differed by 47 sites with 21 of these being insertions/deletions of one or more contiguous bases (Figure 2).

The likelihood ratio test implemented in MODELTEST indicated that the Jukes-Cantor (JC) model of evolution was appropriate for the intron 1 sequences. The JC model with a proportion of invariant sites (I) equal to 0.119 and a gamma distribution shape parameter (Γ) equal to 0.999 was found to be the best model for the exon 2 sequences. Felsenstein 81 (F81) with a proportion of invariant sites equal to 0.273 was found to be the most likely model for the exon 3 sequences. Mean pairwise sequence divergence between the type I and *Anvi*-DAB1 intron sequences was 0.135 (0.027 SE). The six *Anvi*-DAB1 exon 2 alleles had a mean sequence divergence of 0.474 (0.062 SE) from all other *A. virens* MHC sequences. This divergence value was greater than the mean within-group divergence of 0.242 (0.024 SE) for all non–*Anvi*-DAB1 exon 2 sequences. The mean within-group sequence divergence for the *Anvi*-DAB1 exon 2 alleles was 0.011 (0.004 SE).

cDNA Analyses

From one *A. virens*, we amplified and sequenced clones of exons 1 through 3 using primers 1d and 2b. A total of 14 unique sequences were recovered from the cDNA. We found 108 variable sites in exon 2 defining 14 alleles. Only two unique alleles were found in exon 1 and three unique alleles in exon 3. None of these sequences showed more than 74.5% sequence similarity with those from *Anvi*-DAB1. However, given that we sequenced 20 clones and obtained 14 unique sequences, it is possible we could have missed sequences that resembled *Anvi*-DAB1. The cDNA sequences were closer phylogenetically to type I exon 2 sequences (Figure 6).

A BLAST search (Altschul et al. 1997) of four representative cDNA clones (2.1.3, 2.1.4, 4.1.4, and 4.1.9) and four representative genomic sequences from exon 2 (Fragment A) (B2, B7, B10, and *Anvi*-DAB1) in GenBank revealed high similarity to other songbird class II β MHC sequences. The highest similarity (90%) was between clone 4.1.9 and the red-winged blackbird sequence, *Agph*-DAB1 (AF030991). The lowest similarity (83%) was between *Anvi*-DAB1 and a Darwin's finch (*Cerithidea olivacea)* MHC sequence (AF164165).

```
1
TCGGGGAGACGCTGGAGGTGGGAGGGGACGGGATGGGAAAGGAGAAATGGGGGAAAGGGTGGAAAAGAGGAAGCGGCAGCGCGGGCAGGAGGGTCCCATG
Intron 1 (fragment)→
101
GCGGCCGTGTGGAGGGGGGGATCCAGGGTGGCCCTCAGGACTCTCTGGGTGCTGAGGGACACCCCCTGACCTGTGCCATGGACGCTTACACTCAGGAGTG
                                                                                                  Exon 2 ↑
201
TTCCAGGAGATGGATAAGCTCGAGTGTTACTACATCAACAGCACCGAGCGGGTAAAGTACGTTGAAAGGCACATCTACAACCGGGAGCAGCTCTTGCACT

301
TCGACAGCGACGTGGGGCACTACGTGGGGGATACCGTATTTGGGGAGAAGGTGGCCCGGTACTTCAACAGCGACCCGGAAATACTGGAGGACACACGGGC

401
TCAGGCTGACACAGTTTGCCGGCCCAACTATGAGGCTTCCATCCAATTCCTCGCGGAGCGCAGAGGTGAGTGTGGTGCAGAGCGTGTCCCCTCGGGCCCT
                                                    Intron 2 ↑
501
GCCCTGGCAATGAGCCTGGAGCCCCTCAAAACCTCCCTGGAAATCACCCCAGAACCCTCAGCCCTCCTTGGGCCAATCCCCGAGGCCTTCTGTCCCTCCC

601
AGCTCATCCCAGTCTCTCCCAGTGCGCCCCCCGCGTGTCCTTCCCGGCCATGCGTGGAGCTGACGCCTCGCAGCCAATCAGAGTGCGTCTCTCTGATGAC

701
GCATCAGTTGCCAGGCAGACCCGCAGCGCCGACTGCCCCGCGCTGAACTCGGCCTCCCTGTGCCGGGCACTTCATTCCCAGTTCCGCCCAGTCCATNCCC

801
AATCTCTCCCAGTCCCTCCCCAGCCCACCCCGCCCCTCCCCTCTCTCTCCCAGCTCCCCCCAGCTGATCCCAGTGTGTCCCCGCTCTCTCCCAGTGCCCC
                                                                                          Exon 3 ↑
901
CCAGCGTTTCCATCTCGCTGCTGCCCTCGAGCTCCCAGCCCGGCCCTGGCCGCCTGCTGTGCTCCGTGCTGGATTTCTACCCTGCCCCCATCCAGGTGAG

1001
GTGGTTCCAGGGCCAGCAGGAGCTCTCGGACCACGTGGTGGCCACCGACGTGGTCCCCAACGGGGACTGGAGCTACCAGCTGCTGGTGCTGCTGGAAACG

1101    1109
CCCCCCCGG
```

Figure 3. Sequenced fragments of *Anvi*-DAB1. The sequence begins at the first base pair following the forward primer and ends at the reverse primer. Putative exons are in bold, and intron/exon boundaries are marked below the sequence by an "↑." Ambiguous sites are indicated by an "N."

Nonsynonymous to synonymous substitution ratios differed among the MHC gene sequences that we investigated. The type I exon 2 and cDNA sequences displayed elevated d_N:d_S ratios for putative antigen-binding codons (Table 2). For example, the d_N:d_S ratio for the type I exon 2 and cDNA sequences was 1.89 and 1.69, respectively, for ABS codons (Table 2). In contrast, only one nonsynonymous substitution (d_N = 0.007) and no synonymous substitutions were found in ABS codons of exon 2 alleles from the *Anvi*-DAB1 locus (Table 2). Further, the d_N:d_S ratios are close to one for the non-ABS codons of *Anvi*-DAB1, suggesting neutrality (Table 2). None of the tests for the inequality of d_N:d_S ratios were significant at $P < .05$, although the cDNA d_N:d_S ratios were significant at $P < .10$.

Concerted evolution is suggested by separate well-supported groupings of scrub jay (Apco) and little greenbul (Anvi) exon 2 alleles (Figure 6). Similarly, although the sequences are short, exon 3 sequences defined clades that generally corresponded to species (Figure 7). All the unique *A. virens* exon 3 sequences, except from *Anvi*-DAB1, were monophyletic with a posterior probability value of 0.57 (Figure 7). The groupings of exon 3 alleles from the reed warbler (Acar), Bengalese finch (Lost), and red-winged blackbird (Agph) had high posterior probabilities in all cases (Figure 7).

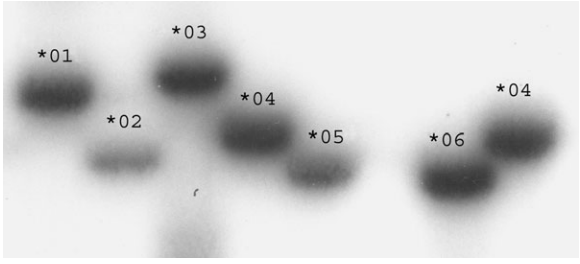

Figure 4. SSCP gel showing the six alleles found in seven little greenbuls. Each allele was cloned, reamplified, and subjected to SSCP analysis to confirm mobility differences (see Methods).

Discussion

Using a PCR-based approach, we have isolated genomic and transcribed portions of several MHC class II β loci from *A. virens*. We used an approach that utilizes knowledge of

```
                    11  20  33  34  39  54  83  92
                    |   |   |   |   |   |   |   |
Anvi-DAB1*01    GAT AGC AAC CGG CAC GGG TAT GAG
Anvi-DAB1*02    .G. G.. ... ... ... ... ..C ...
Anvi-DAB1*03    ... ... ... ... ..G ... ... ...
Anvi-DAB1*04    ... G.. ... G.. ... ... ..C ...
Anvi-DAB1*05    ... G.. ... ... ... ..C ..C -..
Anvi-DAB1*06    ... G.. G.. ... ... ... ..C ...
```

Figure 5. Variable codons of the exon 2 sequences for six cloned alleles from the *Anvi*-DAB1 locus.

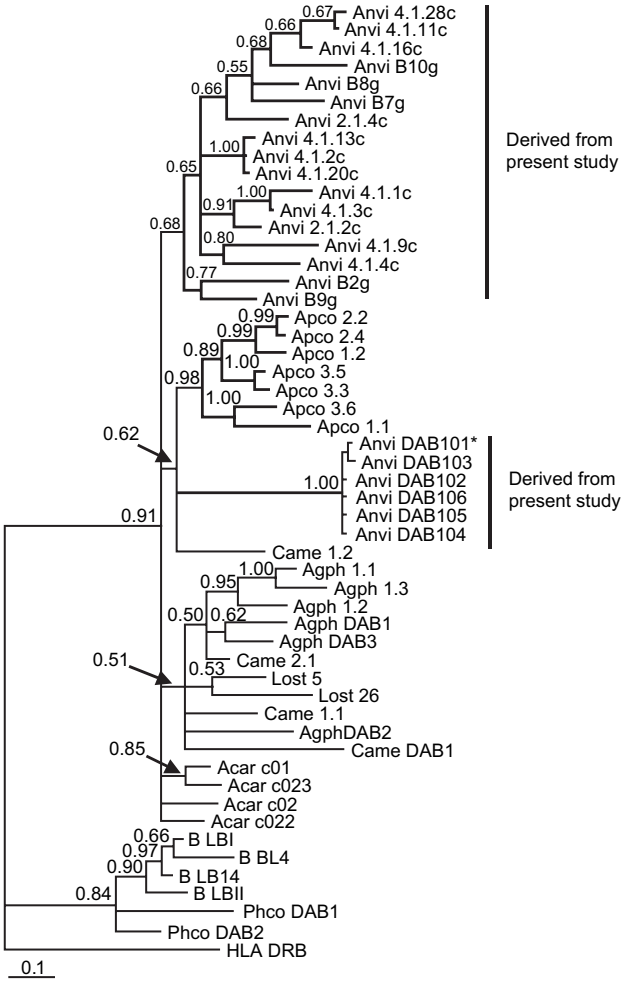

Figure 6. Bayesian tree of the exon 2 sequences from select avian taxa. Numbers above nodes indicate their posterior support. Sequences followed by a "c" are from cDNA and "g" are type I genomic sequences, and the *Anvi*-DAB1 sequence is followed by an "*." Solid vertical line indicates sequences obtained in this study. Posterior probabilities above 50% are shown. Species codes: Acar, *Acrocephalus arundinaceus*; BLB, *Gallus gallus*; Agph, *Agelaius phoeniceus*; Apco, *Aphelocoma coerulesens*; Anvi, *Andropadus virens*; Came, *Carpodacus mexicanus*; HLA, *Homo sapiens*; Lost, *Lonchura striata*; Phco, *Phasianus colchicus*.

conserved regions within the avian MHC to amplify unique exons and introns. This approach may prove useful in the isolation of MHC genes from other nonmodel organisms. The locus for which we have the most information, *Anvi*-DAB1, possessed low levels of polymorphism and was highly divergent from other MHC class II β genes. Similarly, *Anvi*-DAB1–like sequences were not found in a survey of MHC class II β transcripts.

Is *Anvi*-DAB1 a Pseudogene?

Our results suggest that *Anvi*-DAB1 is a nonfunctional or nonclassical gene. First, we observed low levels of polymorphism and divergence among alleles, a result that is not expected for a "classical" MHC gene (Edwards and Hedrick 1998; Hedrick 1994). Similarly, Hess et al. (2000) observed a decrease in polymorphism for an MHC class II pseudogene in house finches and chickens when compared to polymorphism in classical MHC genes. Second, the d_N:d_S ratio for *Anvi*-DAB1 was lower than would be expected for a typical MHC gene as most show an elevated level of nonsynonymous to synonymous substitution reflecting strong balancing selection (Hughes and Nei 1988, 1989). Third, d_N:d_S ratios of *Anvi*-DAB1 alleles for ABS and non-ABS codons were similar. Fourth, we discovered a frameshift mutation consistent with this gene being a pseudogene. Alternatively, the allele containing the frameshift mutation may be nonfunctional, a pseudoallele. Finally, we found a lack of *Anvi*-DAB1–like alleles in the cDNA analysis. However, this result may be due to transcription of this locus at low levels or nonexistent expression in the spleen. Furthermore, we may have missed an *Anvi*-DAB1 transcript from the spleen cDNA, given we only obtained 14 unique sequences from 20 sequenced clones.

Although nonfunctional MHC gene copies are common in songbirds (Edwards et al. 1998, 2000; Hess et al. 2000) and *Anvi*-DAB1 has many qualities that would indicate it is a pseudogene, we cannot rule out the possibility that *Anvi*-DAB1 is a nonclassical class II MHC gene. The nonclassical MHC genes that have been discovered to date are predominately MHC class I genes (Fischer Lindahl et al. 1997). These genes are distinguished from classical MHC genes by having limited amounts of variation and presenting a unique set of epitopes to the immune system (Fischer Lindahl et al. 1997). Sato et al. (2001) have found evidence for a nonclassical class II MHC gene in the Thraupini, a group of South American songbirds including Darwin's finches. However, the observation of an allele with a single frameshift mutation still argues against *Anvi*-DAB1 being a functional nonclassical MHC gene. Further, we have found homozygotes for the frameshift allele (*Anvi*-DAB1*05), supporting the theory that *Anvi*-DAB1 is a nonfunctional gene (Aguilar et al. 2005).

In classical MHC genes of birds and other vertebrates, substitutions in the ABS codons are predominately replacement substitutions, and therefore d_N:d_S ratios are greater than one (Edwards et al. 1995b, 1998; Hughes and Nei 1988, 1989). The elevated d_N:d_S ratios for ABS codons have been attributed to the action of balancing selection (Hughes and Nei 1988, 1989). This pattern is observed for transcribed MHC sequences and type I exon 2 sequences, supporting the conclusion that they are derived from classical MHC loci (Table 2). In contrast, *Anvi*-DAB1 is either a pseudogene or nonclassical MHC gene.

Multiple MHC alleles were amplified from both genomic DNA and cDNA from a single *A. virens*, indicating that many MHC class II loci are present in this species. This supports the view that a minimal essential MHC does not exist in *A. virens* as has been shown in other songbirds (Edwards et al. 2000; Freeman-Gallant et al. 2002; Westerdahl et al. 1999, 2000). The presence of a minimal essential MHC in *A. virens* would have resulted in the amplification of fewer MHC loci.

We isolated genomic sequences from intron 1 and exon 2 of *A. virens* MHC class II β loci. A single type I intron 1

Table 2. Nonsynonymous (d_N) and synonymous (d_S) substitutions per site and their ratios with one standard error (in parentheses) for exon 2 of the pool of MHC class II β loci, cDNA transcripts, and *Anvi*-DAB1. Values are given for ABS and non-ABS whose division is based on the crystallographic structure of Human class II β products (Brown et al. 1993)

	ABS			Non-ABS		
Exon 2 source	d_N (SE)	d_S (SE)	d_N:d_S	d_N (SE)	d_S (SE)	d_N:d_S
Type I pool	0.578 (0.153)	0.306 (0.102)	1.89:1 (ns)	0.164 (0.038)	0.176 (0.039)	0.93:1 (ns)
cDNA	0.454 (0.108)	0.268 (0.071)	1.69:1*	0.117 (0.025)	0.105 (0.025)	1.11:1 (ns)
Anvi-DAB1	0.007 (0.007)	0	0.007:0 (ns)	0.011 (0.006)	0.017 (0.012)	0.65:1 (ns)

ns, not significant; P values $> .10$ for Z test.

* $P = .095$ for Z test.

sequence was associated with at least five unique exon 2 sequences found in a single individual, a result that would suggest a minimum of three loci if the individual was heterozygous for two loci and homozygous for one. Because of the high similarity of type 1 intron 1 sequences, we could not design locus-specific primers to amplify these putative functional genes. In general, the strong similarity of type 1 intron 1 sequences across at least three MHC loci suggests that recent gene duplications or gene conversion is acting to homogenize the type I intron 1 sequences. Other explanations for the low variation of type I intron 1 sequences include the possibility that the locus had low mutation rates or was derived from recombination between intron 1 and exon 2 (Garrigan and Edwards 1999). These alternatives cannot be tested without

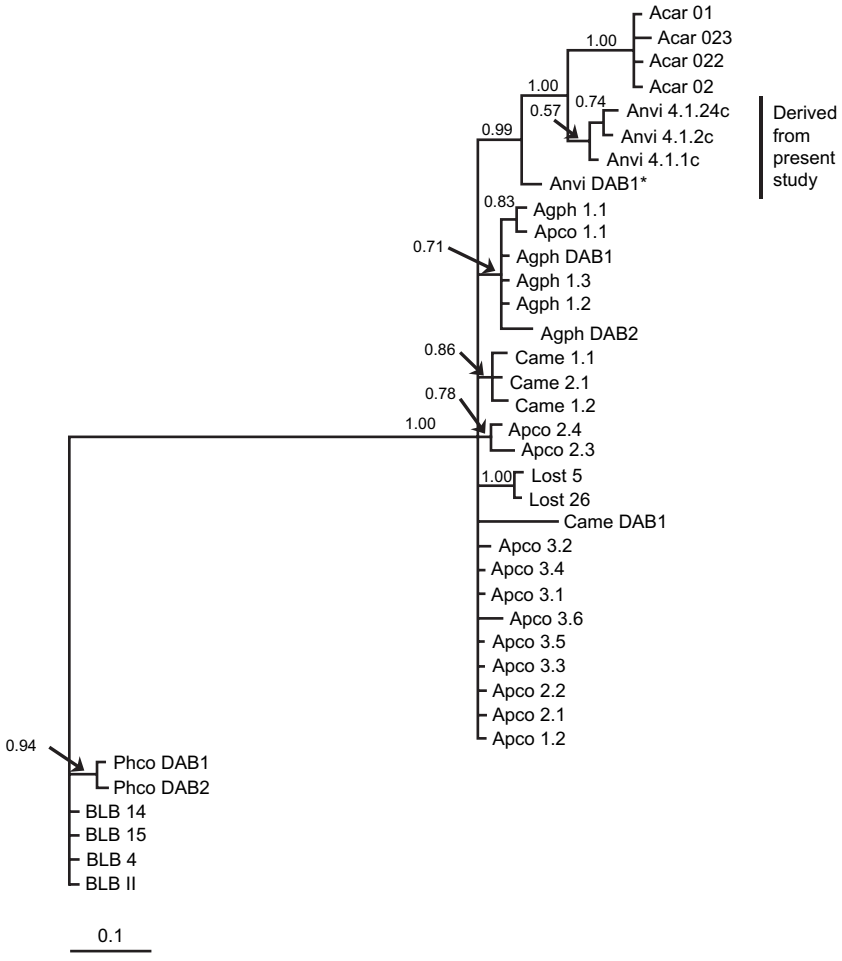

Figure 7. Bayesian tree of the exon 3 sequences from select avian taxa. Numbers above nodes indicate their posterior support if $> 50\%$. Sequences flowed by a "c" are from cDNA, and *Anvi*-DAB1 sequence is followed by an "*." Solid vertical line indicates sequences obtained in this study. See Figure 6 for species codes.

sequence information from other closely related species and knowledge of locus orthology.

The birth-and-death process (Nei et al. 1997) may be operating on the songbird MHC (Edwards et al. 2000). The birth-and-death process incorporates mechanisms such as concerted evolution but allows gene duplications and the formation of pseudogenes (Nei et al. 1997). We find that *A. virens* type I exon 2 and 3 sequences and cDNA transcripts all show a close relationship to one another (Figures 6 and 7), which supports concerted evolution of these gene segments (Figure 6). Studies of exon 1 and larger portions of exon 3 in other avian taxa have found a similar result (Edwards et al. 1995b). In contrast, exon 2 and 3 alleles from the *Anvi*-DAB1 locus were highly divergent from other class II β sequences, indicating that *Anvi*-DAB1 did not experience concerted evolution and was not the result of a recent gene duplication event (Figures 6 and 7). Therefore, in contrast to the concerted evolution that may be operating on the more classical MHC genes in songbirds, *Anvi*-DAB1 may have arisen from an ancient gene duplication event and has remained a genetically independent locus.

Acknowledgments

We thank the Government of the Republic of Cameroon for permission to conduct the field research. This research was supported by grants from the National Geographic Society, National Environmental Research Council, and Royal Society and the National Science Foundation grants DEB-9726425 to T.B.S. and R.K.W. and IRCEB9977072 to T.B.S., an Academic Senate grant (UCLA) to R.K.W., and an NSF-DDIG to A.A. We thank C. S. Baker, S. M. Blankenship, and three anonymous reviewers for comments on the manuscript.

References

Aguilar A, Smith TB, and Wayne RK, 2005. A comparison of variation between an MHC pseudogene and microsatellite loci of the little greenbul (*Andropadus virens*). BMC Evol Biol 5:47.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res 25:3389–3402.

Black FL and Salzano FM, 1981. Evidence for heterosis in the HLA system. Am J Hum Genet 33:894–899.

Brown JH, Jerdetsky T, Gorga JC, Stern LJ, Urban RG, Strominger JL, and Wiley DC, 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33–39.

Edwards SV, Gasper J, Garrigan D, Martindale D, and Koop BF, 2000. A 39-kb sequence around a blackbird *MHC* class II gene: ghost of selection past and songbird genome architecture. Mol Biol Evol 17:1384–1395.

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

Edwards SV, Grahn M, and Potts WK, 1995a. Dynamics of *MHC* evolution in birds and crocodilians: amplification of class II genes with degenerate primers. Mol Ecol 4:719–729.

Edwards SV and Hedrick PW, 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. Trends Ecol Evol 13: 305–311.

Edwards SV, Wakeland EK, and Potts WK, 1995b. Contrasting histories of avian and mammalian *MHC* genes revealed by class II sequences from songbirds. Proc Natl Acad Sci USA 92:12200–12204.

Fischer Lindahl K, Byers DE, Dabhi VM, Hovik R, Jones EP, Smith, GP, Wang C-R, Xiao H, and Yoshino M, 1997. H2-M3, a full service class IB histocompatibility antigen. Annu Rev Immunol 15:851–879.

Freeman-Gallant CR, Johnson EM, Saponara F, and Stanger M, 2002. Variation at the major histocompatibility complex in savannah sparrows. Mol Ecol 11:1125–1130.

Garrigan D and Edwards SV, 1999. Polymorphism across an intron-exon boundary in an avian MHC class II B gene. Mol Biol Evol 16:1599–1606.

Gasper JS, Shiina T, Inoko H, and 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.

Hedrick PW, 1994. Evolutionary genetics of the major histocompatability complex. Am Nat 143:945–964.

Hedrick PW, Lee RN, and Garrigan D, 2002. Major histocompatibility complex variation in red wolves: evidence for common ancestry with coyotes and balancing selection. Mol Ecol 11:1905–1913.

Hedrick PW, Lee RN, and Parker KM, 2000. Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids. Heredity 85:617–624.

Hess CM, Gasper J, Hoekstra HE, Hill CE, and Edwards SV, 2000. MHC class II pseudogene and genomic signature of a 32-kb cosmid in the house finch (*Carpodacus mexicanus*). Genome Res 10:613–623.

Huelsenbeck JP and Ronquist F, 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Hughes AL and Nei M, 1988. Pattern of nucleotide substitutions at major histocompatability complex class I loci reveals overdominant selection. Nature 352:167–170.

Hughes AL and Nei M, 1989. Nucleotide substitutions at major histocompatability complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci USA 86:958–962.

Kaufman J, Milne S, Gobel TWF, Walker BA, Jacob JP, Auffray C, Zoorob R, and Beck S, 1999. The chicken B locus in a minimal-essential major histocompatibility complex. Nature 401:923–925.

Kaufman J, Völk H, and Wallny H, 1995. A "minimal-essential MHC" and an "unrecognized MHC": two extremes in selection for polymorphism. Immunol Rev 143:63–88.

Keith S, Urban EK, and Fry CH, 1992. The birds of Africa, vol 4. New York: Academic Press.

Kirkpatrick CE and Smith TB, 1988. Blood parasites of birds in Cameroon. J Parasitol 74:1009–1013.

Klein J, 1986. Natural history of the major histocompatibility complex. New York: John Wiley & Sons.

Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, and Watkins DI, 1990. Nomenclature for the major histocompatibility complexes of different species: a proposal. Immunogenetics 31:217–219.

Klein J, Satta Y, and O'Huigin C, 1993. The molecular descent of the major histocompatibility complex. Annu Rev Immunol 11:269–295.

Kumar S, Tamura K, and Nei M, 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163.

MHC Sequencing Consortium, 1999. Complete sequence and gene map of a human major histocompatibility complex. Nature 401:921–923.

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

Milner CM and Campbell RD, 1992. Genes, genes and more genes in the human major histocompatibility complex. Bioessays 14:565–571.

Nei M, Gu X, and Sitnikova T, 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA 94:7799–7806.

Nei M and Kumar S, 2000. Molecular evolution and phylogenetics. New York: Oxford University Press.

Orita M, Iwahana H, Kanazawa H, Hayashi K, and Sekiya T, 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA 86:2766–2770.

Posada D and Crandall KA, 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Potts WK, 1996. PCR-based cloning across large taxonomic distances and polymorphism detection: MHC as a case study. In: Molecular zoology advances, strategies, and protocols (Ferraris JD and Palumbi SR, eds). New York: Wiley-Liss Inc.; 181–192.

Robinson J, Malik A, Parham P, Bodmer JG, and Marsh SG, 2000. IMGT/HLA database—a sequence database for the human major histocompatibility complex. Tissue Antigens 55:280–287.

Rosas J and Rosas R, 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174–175.

Sambrook J, Fritsch EF, and Maniatis T, 1989. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press.

Sato A, Meyer WE, Tichy HE, Grant PR, Grant RB, and Klein J, 2001. Evolution of MHC class II B genes in Darwin's finches and their closest relatives: birth of a new gene. Immunogenetics 53:792–801.

Sehgal RNM, Jones HI, and Smith TB, 2001. Host specificity and incidence of *Trypanosoma* in some African rainforest birds: a molecular approach. Mol Ecol 10:2319–2327.

Seutin G, Boag BN, and Boag PT, 1991. Preservation of avian blood and tissue samples for DNA analysis. Can J Zool 69:82–90.

Shiina T, Shimizu C, Oka A, Teraoka Y, Imanishi T, Gojobori T, Hanzawa K, Watanabe S, and Inoko H, 1999. Gene organization of the quail major histocompatibility complex (MHCCoja) class I gene region. Immunogenetics 49:384–394.

Smith TB, Schneider CJ, and Holder K, 2001. Refugial isolation versus ecological gradients: testing alternative mechanisms of evolutionary divergence in four rainforest vertebrates. Genetica 112/113:383–398.

Smith TB, Wayne RK, Girman DJ, and Bruford MW, 1997. A role for ecotones in generating rainforest biodiversity. Science 276:1855–1857.

Smith TB, Wayne RK, Girman DJ, and Bruford MW, 2000. Evaluating the divergence-with-gene-flow model in natural populations: the importance of ecotones in rainforests speciation. In: Rainforests past and future (Mortiz C and Bermingham E, eds). Chicago: University of Chicago Press.

Tajima F, 1989. Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Takahata N, 1990. A simple genealogical structure of strongly balanced allelic lines and trans-specific evolution of polymorphism. Proc Natl Acad Sci USA 87:2419–2423.

Trowsdale J, 1995. "Both man and bird and beast": comparative organization of MHC genes. Immunogenetics 41:1–17.

Wagner JL, 2003. Molecular organization of the canine major histocompatibility complex. J Hered 94:23–26.

Westerdahl H, Wittzell H, and Von Schantz T, 1999. Polymorphism and transcription of MHC class I genes in a passerine bird, the great reed warbler. Immunogenetics 49:158–170.

Westerdahl H, Wittzell H, and Von Schantz T, 2000. MHC diversity in two passerine birds: no evidence for a minimal essential MHC. Immunogenetics 52:92–100.

Wittzell H, Bernot A, Auffray C, and Zoorob R, 1999. Concerted evolution of two MHC class II β loci in pheasants and domestic chickens. Mol Biol Evol 16:479–490.

Zoorob R, Bernot A, Renoir DM, Choukri F, and Auffrey C, 1993. Chicken major histocompatability complex class II genes: analysis of interallelic and interlocus sequence variance. Eur J Immunol 23:1139–1145.

Received March 28, 2004
Accepted November 30, 2005

Corresponding Editor: C. Scott Baker