Genomics and Polymorphism of *Agph-DAB1*, an *Mhc* Class II B Gene in Red-Winged Blackbirds (*Agelaius phoeniceus*)

Scott V. Edwards,*† *Joe Gasper*,* *and Mariam March*[*1]

*Department of Zoology and †Burke Museum, University of Washington

To further our understanding of the evolution of avian *Mhc* genes at the genomic level, we screened a cosmid library made from a red-winged blackbird (*Agelaius phoeniceus*) with a blackbird cDNA probe and subcloned from one of the *Mhc*-containing cosmids a gene which we designate *Agph-DAB1*. The structure of the gene is similar to that found for chicken class II B genes, except that the introns are surprisingly large, ranging from 98 to over 600 bp, making this the longest avian class II B gene to date. Using primers targeted toward the introns flanking the peptide-binding region (PBR), we amplified the entirety of the second exon and determined nucleotide sequences of 41 PCR products from eight individual blackbirds. The 10 sequence types found, among which were two probable pseudogene sequences, exhibit the classic hallmarks for evolution of PBRs, namely, an excess of nonsynonymous over synonymous substitutions and evidence of gene conversion events in polymorphic subdomains. Despite these patterns and our use of intron primers, the distribution of sequences among individuals suggests that more than one locus was amplified in most individuals, and the bushlike tree of sequences provides little information as to locus-specific clusters. These results imply a complex history of gene conversion, recent duplication, or possibly, concerted evolution among multiple loci, although *Agph-DAB1*, the first genomic *Mhc* sequence from a bird other than chicken, provides important clues in the quest for locus-specific *Mhc* primers in birds.

Introduction

The major histocompatibility complex (*Mhc*) is a dynamic multigene family found in vertebrates encoding immunoglobulinlike receptors which bind foreign peptides (short protein fragments) for presentation to T-cells (Klein 1986; Klein, Satta, and O'hUigin 1993). An *Mhc* molecule, T-cell receptor, and peptide are the three key players required to initiate cellular immune responses to infectious agents (Barber and Parham 1993; Germain and Margulies 1993). In the past, the structure and function of *Mhc* molecules were principally of interest to immunologists, but as ecologists and evolutionary biologists have begun to appreciate the prominent role of parasites in the ecology, evolution, behavior, and conservation of vertebrates, *Mhc* genes have drawn attention as a major immunological mechanism against invasion by parasites (Hedrick and Miller 1995; Edwards and Potts 1996; Sheldon and Verhulst 1996). Additionally, deciphering the genetic forces of mutation, selection, recombination, and gene conversion underlying the extraordinary polymorphism of some *Mhc* genes has proven to be a challenge to population geneticists (Hedrick and Thomson 1983; Hughes and Nei 1989; Takahata and Nei 1990; Erlich and Gyllensten 1991; Takahata, Satta, and Klein 1992; Satta et al. 1993, 1994; Hedrick 1994; Edwards et al. 1997; Hedrick and Kim 1998). The *Mhc*s of humans and other mammals are the best known, but recent molecular studies have made inroads to understanding the *Mhc*s of natural populations of fish, amphibians, birds, and reptiles (e.g., Dixon et al. 1995; Edwards 1996; Radtkey et al. 1996; Sammut, Laurens, and Tournefier 1997).

The evolution of *Mhc* genes in birds is of particular interest for a variety of reasons. As oviparious (egg-laying) vertebrates, birds provide critical tests of some hypotheses for the maintenance of Mhc polymorphism proposed for mammals (e.g., maternal-fetal interaction hypotheses; Hedrick and Kim 1998). Indeed, we know little about levels and patterns of polymorphism in natural populations of any avian *Mhc* genes. The chicken *Mhc* is known to be much smaller and contain far less noncoding DNA than the *Mhc*s of many mammals (Trowsdale 1995), and it would be useful to know if this situation characterizes other birds, and why. Other genes of immunological importance, such as immunoglobulin VL genes, are thought to have evolved and are known to somatically hypermutate in ways completely different from those of mammals (Reynaud et al. 1987; McCormack, Tjoelker, and Thompson 1991), warranting investigation of other immunoglobulin superfamily genes of birds. Finally, based on cDNA sequences, Edwards, Wakeland, and Potts (1995) suggested that the long-term history of avian *Mhc* genes may be characterized by higher rates of concerted evolution, or more recent gene duplications, than that of mammalian *Mhc* genes, a suggestion that should be corroborated at the genomic level.

One practical goal to understanding *Mhc* evolution in a variety of contexts is to characterize *Mhc* genes in enough detail that the targets of selection on *Mhc* genes—e.g., those domains that bind foreign peptides, the peptide-binding regions (PBRs)—can be rapidly surveyed for allelic polymorphism at the sequence level in a large number of individuals. In the case of class II *Mhc* genes, which differ from class I and III genes in their structure, expression, and mode of evolution, the PBR is encoded by a single exon, the second exon, about 90 amino acids long. The ability to amplify this exon by PCR from individual class II genes is preferable to restriction fragment analysis, so that alleles from orthologous loci can be compared rapidly at the nucleotide level (Edwards and Potts 1996). Clearly, the mode of

[1] Present address: Genome Center, University of Washington.

Key words: major histocompatibility complex, introns, balancing selection, gene conversion, pseudogene, birds.

Address for correspondence and reprints: Scott V. Edwards, Department of Zoology, University of Washington, Box 351800, Seattle, Washington 98195. E-mail: edwards@zoology.washington.edu.

Mol. Biol. Evol. 15(3):236–250. 1998

evolution of the *Mhc* multigene family will dictate the ease with which this can be accomplished. If, as in mammals, the class II genes evolve primarily in a divergent manner in which different *Mhc* genes remain distinct and diverge from one another after relatively ancient gene duplications (Nei and Hughes 1992), then orthologous genes and their flanking regions will evolve unique sequences that distinguish them from their paralogs in the same genomes. In contrast, it has been suggested that the class II genes of birds evolve in a concerted manner or have duplicated recently, resulting in substantial similarity among different genes in both coding and noncoding regions (Edwards, Wakeland, and Potts 1995; Edwards 1996). Although this mode of evolution needs to be confirmed and characterized in more detail for birds, in principle, it would make it more difficult to study the evolution of individual *Mhc* genes, because there are fewer sequence motifs distinguishing *Mhc* paralogs from orthologs.

Researchers have taken a variety of approaches to attack these problems. In well-characterized *Mhc*s in which intron sequences are known, it seems best to place PCR primers in flanking introns. This way, the entire PBR can be characterized, and if genes are evolving divergently, it is likely that variable introns will differ among genes. Thus, for example, the individual class II genes of a variety of mammals (e.g., She et al. 1990; Figueroa et al. 1994; Brunsberg et al. 1996; Edwards et al. 1997) and fish (Graser et al. 1996; Sato et al. 1996) have been amplified using primers targeted toward the introns flanking the second exon. However, this method does not guarantee clean amplification of single loci in some species, as evidenced by surveys of chicken class II genes, in which intron primers sometimes amplify more than one locus (Zoorob et al. 1993). In principle, locus-specific motifs can help distinguish the loci amplified, although some interpretation is necessary here (Zoorob et al. 1993). More commonly, class II *Mhc* genes of less well characterized genomes have been accessed by placing primers in two conserved blocks internal to the second exon (e.g., Edwards, Grahn, and Potts 1995; Vincek et al. 1997). In mammals, this approach is efficient and usually amplifies single loci (e.g., Murray, Malik, and White 1995), although occasionally related loci are coamplified, as in the case of primate DQB and DXB genes (Gyllensten, Lashkari, and Erlich 1990). However, functionally important polymorphisms are missed by this method, since the primers are internal to the PBR. As molecular ecological studies of the *Mhc* become more detailed, linking variation in individual amino acids in the PBR to phenotypic variation, as has been extensively explored in humans (e.g., Apple and Erlich 1992; Hill et al. 1992), the data resulting from this omission will be less satisfactory. In nonmammalian vertebrates, primers internal to the PBR almost inevitably amplify more than one *Mhc* locus and occasionally non-*Mhc* loci (Edwards, Grahn, and Potts 1995; Vincek et al. 1997), with the result that the assignments of sequences to individual loci, if attempted, occur in a post hoc fashion once the sequence polymorphisms have been studied. Again, this method can work in principle as long as different genes are distinguished by motifs, but it is unlikely that this approach will satisfy geneticists used to exploring *Mhc* structure in genetically well characterized model species with inbred lines and pedigreed populations (Kaufman et al. 1991; Flajnik et al. 1993; Miller et al. 1994, 1996).

With these ideas in mind, we have genomically characterized an *Mhc* class II gene from the red-winged blackbird (*Agelaius phoeniceus*). Red-winged blackbirds are a common North American songbird whose extreme sexual dimorphism and variable mating system have made it a model system for ecological, evolutionary, and phylogenetic studies (Orians and Beletsky 1989; Weatherhead 1990; Lanyon 1994). There are some data on parasite prevalence in natural populations (Weatherhead 1990; Weatherhead and Bennett 1991), and experimental field manipulations have helped identify some of the targets of selection at the organismal level (Weatherhead and Clark 1994; Searcy and Yasukawa 1995; Rohwer, Langston, and Gori 1996). Gibbs et al. (1991) investigated *Mhc* RFLPs in this species by probing genomic DNA with a mouse class II probe but was unconvinced that the polymorphism detected was due to *Mhc* variation and not a polymorphic, non-*Mhc* locus. Thus, a major goal of this study was to begin to characterize the structure of *Mhc* genes in blackbirds to aid in surveys of *Mhc* polymorphism. No *Mhc* genes from birds other than chickens have been characterized at the genomic level, although coding sequences for a number of class I and II *Mhc* genes are known for several birds (mostly game birds), including red-winged blackbirds (Wittzell et al. 1994; Edwards, Grahn, and Potts 1995; Edwards, Wakeland, and Potts 1995; Shiina et al. 1995; Vincek et al. 1995). Red-winged blackbirds appear to contain a large number of class II *Mhc* or *Mhc*-like sequences as judged by Southern blots (unpublished data), making it particularly important to characterize individual genes in detail.

Materials and Methods

Cosmid Library Construction and Screening

Isolation of high-molecular-weight genomic DNA from blackbird blood and construction and screening of a cosmid library was performed as in Longmire et al. (1993). We isolated genomic DNA from a frozen blood sample of a female blackbird from Washington state by proteinase K digestion and gentle phenol-chloroform extraction, followed by dialysis in 1 × TE overnight. No RNAse incubation was necessary, as the level of RNA in blood-derived DNA from birds is minimal. Four micrograms of genomic DNA was partially digested with *Sau*3A so that most DNA was cut into fragments of >23 kb, dephosphorylated and ligated into vector arms of sCos-1 vector. 90,000 colonies were screened using Hybond N+ filters and the ECL nonradioactive screening kit (Amersham) as per manufacturer's instructions and hybridization conditions. Probes spanning exons 1–4 of a red-winged blackbird class II RT-PCR product were derived by amplification of clone 2.1 with primers 8 and 10 of Edwards, Wakeland, and Potts (1995). Prospective cosmid clones were

replated and rescreened twice. Once isolated, the presence of *Mhc* sequences on these clones was confirmed by Southern hybridization (ECL), amplification, and direct sequencing of conserved segments of exons 2 and 3 using primers Rwex2f (5′ → 3′: TTCATTAACGGCACGGAG) and Rwex2r (GGCGGAACACCTCGTAGT) for exon 2 and primers Rwex3f and Rwex3r (GTGATGGATTTCTACCCTGC and CAGCTGTAGGTGAGCCC, respectively) for exon 3.

Shotgun Subcloning, Sequencing, and Survey of Polymorphism

As part of a large scale sequencing study of avian *Mhc* genes, we prepared a subclone library from one of the *Mhc*-bearing cosmids, Rwcos3, using standard shotgun subcloning methods (Wilson et al. 1992; Rowen, Koop, and Hood 1996) that will be outlined in detail elsewhere. Briefly, 1.5–2.5-kb fragments of Rwcos3 were generated by sonication, and then the sheared ends of these fragments were repaired by DNA polymerase and blunt-end-ligated to pUC18. To focus immediately on those subclones bearing *Mhc* genes, as opposed to noncoding or other DNA, several thousand subclone colonies were screened using a probe generated by PCR from one of the blackbird cDNAs using primers Rwex2f and Rwex3r. Although subcloning of *Mhc*-bearing fragments directly from cosmids in this way can be less effective in obtaining subclones spanning entire genes than subcloning restriction fragments, in practice, fewer novel sequencing primers need to be designed, because subclone ends overlap rather than abut and are not confined to restriction sites, thus facilitating a stratified random sequencing approach (for a similar approach applied to *Mhc* genes, see Robertson and McMaster 1985). Subclones from eight positive colonies were then isolated and sequenced using standard plasmid primers and dye-terminator chemistry on an ABI 373A automated sequencer. Sequences were assembled and aligned to one another and to other bird sequences manually using Genetic Data Environment (GDE) software.

To survey polymorphism in the PBR of the gene thus sequenced, we obtained blood and tissue samples from eight blackbirds from a single locality near Gainesville, Fla. (locality details available on request). PCR products were generated from genomic DNA of these birds using primers targeted to introns 1 and 2 flanking exon 2 (fig. 3) and were cloned in PCR 2.1 and pNoTA/.T7 vectors (Invitrogen and 5′ → >3′ kits, respectively). The 41 resulting sequences were aligned by eye using GDE. These and the sequence of *Agph-DAB1* have been deposited in GenBank under accession numbers AF030987–AF030997.

Phylogenetic and Statistical Analysis

Neighbor-joining trees (Saitou and Nei 1987) with bootstrapping (Felsenstein 1985) were generated using PHYLIP (Felsenstein 1994). The numbers of synonymous and nonsynonymous substitutions per site (d_N and d_S) were estimated using the method of Nei and Gojobori (1986) in the computer package MEGA (Kumar, Tamura, and Nei 1993). We did not use the methods of Li (1993) for estimating d_N and d_S, because the high-transition bias that makes it preferable to use this method is not prevalent at those *Mhc* genes that have been examined (Imanishi and Gojobori 1992). Segments of exon 2 that exhibited anomalous likelihoods relative to other segments (‘‘spatial phylogenetic variation,’’ or SPV) were determined using the program PLATO (partial likelihoods assessed through optimization) of Grassly and Holmes (1997). This method calculates a value Q which is essentially the average per-site likelihood for a given region on a given tree divided by the per-site likelihood of all sites outside that region. The program adjusts segment lengths to maximize Q and then compares this value to a null distribution to determine significance. DNA segments resulting in high values of Q exhibit anomalous likelihoods relative to other segments and can be the result of recombination or gene conversion, although anomalously high mutation rates can also yield a similar patterns.

Results

Cosmid Cloning

The screening of the blackbird library using blackbird cDNAs as probes yielded five unique cosmid clones containing *Mhc* sequences after tertiary screening and PCR analysis (see *Materials and Methods*). Partial exon 2 and 3 sequences that could easily be aligned with blackbird and chicken *Mhc* class II B sequences were amplified and sequenced from each of these clones (not shown). Preliminary restriction mapping indicated that cosmids 3 and 7, and possibly 10, overlapped partially, with no detectable overlap in cosmids 1 and 5 (fig. 1). Because of the large number of *Mhc*-hybridizing bands detected in digests of clone 3, this clone was chosen for further analysis (fig. 1). Since chicken *Mhc* genes can occur within 5 kb of one another, we reasoned that a large number of bands could indicate multiple *Mhc* genes on the same cosmid clone.

The cosmid 3 subclones generated after sonication, repair, and ligation (see *Materials and Methods*) were plated and screened with blackbird cDNA sequences, and of the many positive clones, 10 were picked for further analysis. Sequencing of the ends of each of these clones immediately revealed the presence of an *Mhc* gene(s) (fig. 2). Subclone ends were found to be spread across the gene but nonetheless clustered, a phenomenon which has been observed in other experiments of this type. Sequencing from the subclone ends alone permitted a total of nearly 2 kb of unique *Mhc* sequence occurring in three clusters to be determined. All overlapping sequences matched each other exactly without ambiguities, suggesting that one gene had been subcloned. Inspection of the final sequence revealed the presence of a single *Pst* I restriction site in exon 2 (fig. 3). Although our probing of Rwcos3 cut with *Pst* I revealed three fragments, the middle, fainter fragment may be the result of spurious hybridization, and the number of *Pst* I sites found in our sequence is nonetheless consistent with fragment sizes (>2 kb) in our blot data (fig. 1). Although the possibility of two very similar, if not iden-

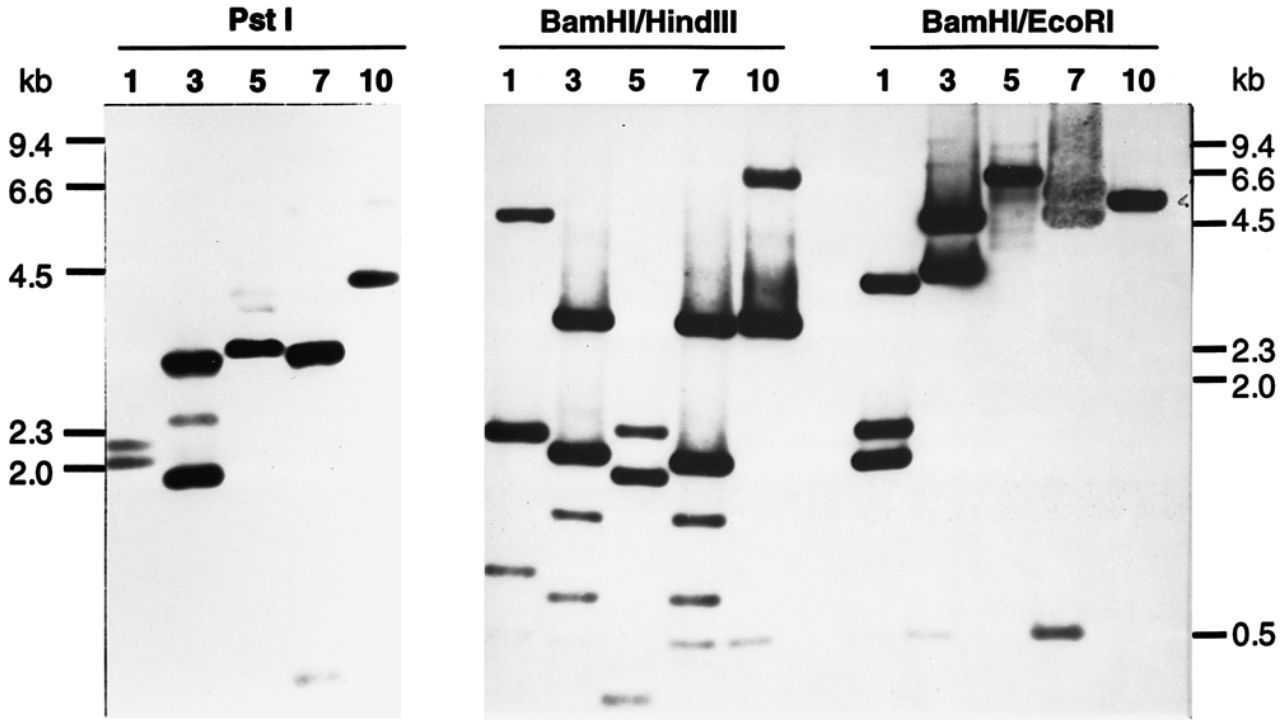

FIG. 1.—Probing of five *Mhc*-containing cosmid clones (1, 3, 5, 7, and 10) with a blackbird cDNA probe. *Left panel,* Hybridization patterns of clones cut with Pst I. *Right panel,* Hybridization pattern of clones cut with *Bam*HI double digests. Size markers in kilobases are indicated.

tical, genes on cosmid 3 remains, from a practical standpoint we gain little information for primer design, since two of our sublcones (2 and 4) span the second exon and flanking regions.

This sequence spans from approximately 500 bp upstream of the start codon to just under 30 bp before the stop codon, assuming an arrangement similar to a chicken class II B gene (fig. 3). We designate this gene *Agph-DAB1* as per the nomenclatural suggestions for *Mhc* genes of Klein et al. (1990) (see also Sato et al. 1996).

The most striking finding in the comparison of *Agph-DAB1* with a chicken BL-B gene is the large intron sizes of the blackbird gene. All four introns of *Agph-DAB1* examined are larger than their chicken counterparts, ranging from 1.1 (intron 4) to 7.3 times (intron 3) as large, making the gene the largest avian class II B gene cloned to date, at over 2.2 kb. The intron-exon structure of *Agph-DAB1* is otherwise similar to that of chicken BL-B genes, with a similarly sized leader peptide and β1 and β2 exons corresponding to the PBR and immunoglobulin-like domains, respectively (Klein

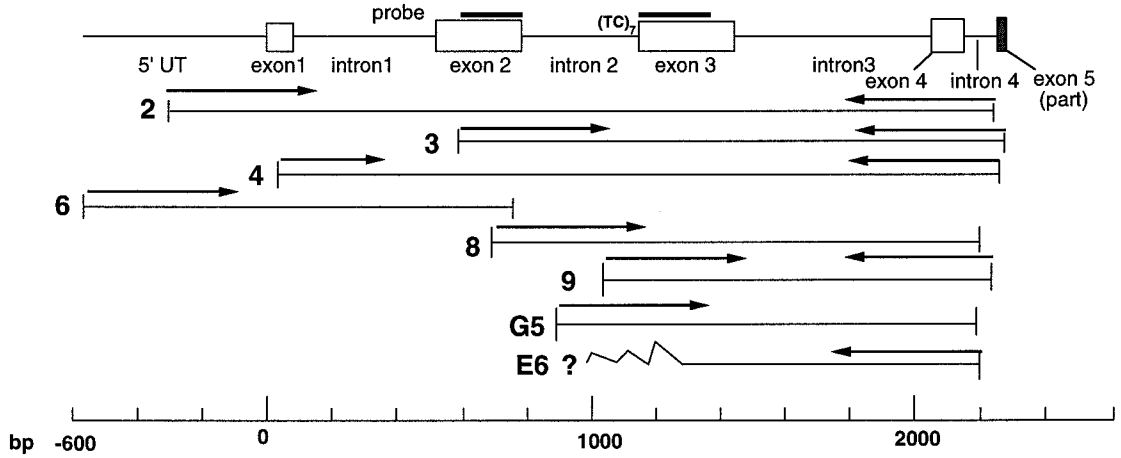

FIG. 2.—Scale diagram of subclone coverage of the blackbird class II B gene *Agph-DAB1*. Intron/exon sizes of the gene are to scale according to the sequence in Fig. 3. The numbering begins at 0 at the start (ATG) codon. Subclone borders are indicated by vertical lines, with subclone numbers indicated to the left. The location of a $(TC)_7$ microsatellite is indicated. The coverage of the cDNA probe is indicated by a solid bar above the gene diagram. The lengths of sequence determined by sequencing the ends of each subclone with vector primers are indicated by arrows. The question mark next to clone E6 indicates that the end of this subclone was not determined.

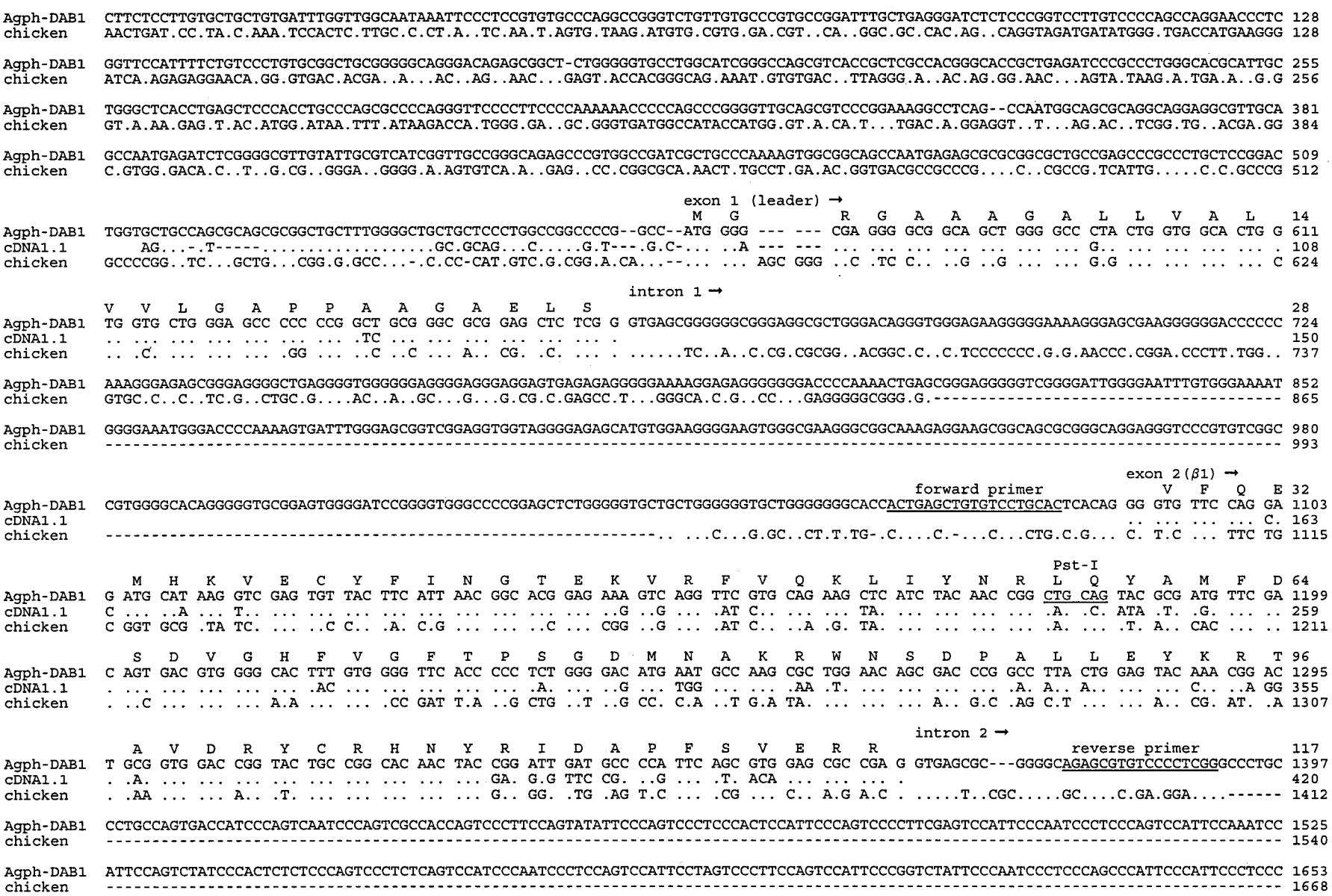

FIG. 3.—Genomic DNA sequence of *Agph-DAB1* compared with a chicken class II B gene (BL-BII; Zoorob et al. 1990) and a previously published red-winged blackbird cDNA (‘cDNA1.1; clone 1.1 of Edwards, Wakeland, and Potts 1995). The amino acid sequence of *Agph-DAB1* is indicated above coding regions, and cumulative numbers of each sequence, including gaps, are indicated at the right. Individual exons and introns and the single *Pst* I site are indicated, as are the forward and reverse primers used to amplify exon 2 (5′ → 3′: Rwcos3I1f.1: ACTGAGCTGTGTCCTGCACT; Rwcos3I2r.1: CCGAGGGGACACGCTCT, respectively). In intron 3, the repeated sequence motif TGGGAGGG is underlined.

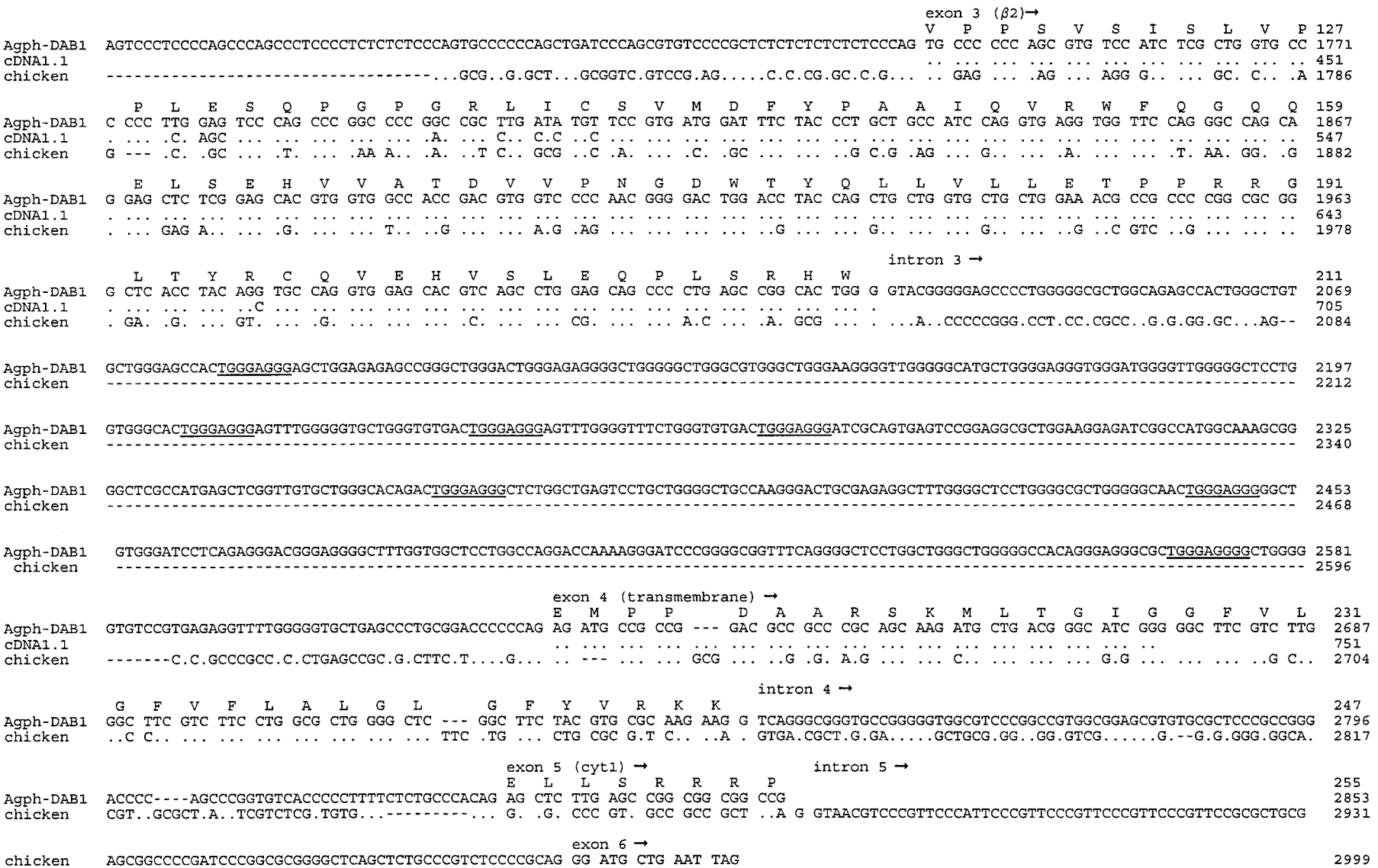

FIG. 3 *(Continued)*

Table 1
Comparison (%) of Base Compositions and Exon/Intron Lengths of Blackbird and Chicken *Mhc* Class II B Genes

	Species	5′	Exon 1			Intron 1	Exon 2			Intron 2	Exon 3			Intron 3	Exon 4			Intron 4	Exon 5		
			1	2	3		1	2	3		1	2	3		1	2	3		1	2	3
		568		84		439		264		380		282		603		105		98		21+	
		571		90		208		264		86		279		82		108		88		42	
A . . .	Blackbird	13.0	3.6	3.6	17.9	20.5	26.1	37.5	5.7	14.7	14.9	25.5	2.1	12.8	22.9	14.3	0.0	7.1	14.3	0.0	0.0
	Chicken	23.5	10.0	0.0	6.7	7.7	23.9	38.6	9.1	8.1	18.3	29.0	1.1	7.3	13.9	11.1	2.8	5.7	7.1	0.0	7.1
T . . .	Blackbird	19.0	3.6	32.1	7.1	10.7	20.5	27.3	11.4	23.7	18.1	29.8	4.3	17.2	17.1	42.9	0.0	16.3	14.3	28.6	0.0
	Chicken	15.2	3.5	26.7	0.0	9.6	22.7	23.9	10.2	9.3	16.1	26.9	1.1	7.3	11.1	44.4	2.8	17.0	7.1	21.4	7.1
C . . .	Blackbird	35.0	28.6	42.9	17.9	15.3	22.7	13.6	43.2	48.4	37.2	24.5	45.7	19.6	22.9	17.1	54.3	38.8	71.4	14.3	28.6
	Chicken	25.6	33.3	43.3	30.0	30.3	21.6	13.6	40.9	39.5	28.0	21.5	36.6	45.1	33.3	16.7	36.1	21.6	57.1	42.9	50.0
G . . .	Blackbird	32.9	64.3	21.4	57.1	53.5	30.7	21.6	39.8	13.2	29.8	20.2	47.9	50.4	37.1	25.7	45.7	37.8	0.0	57.1	71.4
	Chicken	25.7	53.3	24.4	63.3	52.4	31.8	23.9	39.8	43.0	37.6	22.6	61.3	40.2	41.7	27.8	58.3	55.7	28.6	35.7	35.7

NOTE.—In each cell, the number above is the percentage for *Agph-DAB1*, and the number below is that for the BL-BII gene (Zoorob et al. 1990). 1, 2, and 3 in exon columns indicate first, second, and third positions of codons in those exons, respectively. In the top section, numbers above and below are lengths in base pairs for the indicated region in blackbird and chicken, respectively.

Table 2
Distribution of Blackbird *Mhc* Class II B Sequence Types Among Individuals in the PCR Survey

	Sequence Type									
Individual	1a	1b	2	3	4a	4b	5	6	7a	7b	Total
1.	3				1		1				5
2.	5			1		1		2	1		10
3.	3	1	1				1				6
4.	4										4
5.			1				2				3
6.	1	2					2				5
7.	1							1	1	1	4
8.	1			1			1		1		4
Total.	18	3	2	2	1	1	7	3	3	1	41

and O'hUigin 1993). If the cytoplasmic tail of *Agph-DAB1* is similar to that in chickens, as it appears to be on the basis of the short intron 4 and exon 5 (Zoorob et al. 1990), then our clones covered to within 100 bp of the stop codon. The middles of the introns have many repeated sequence variants (e.g., TGGGAGGG repeated nontandemly 7 times in intron 3; TGGGA repeated 18 times nontandemly throughout gene), and a $(TC)_7$ microsatellite appears just 5′ of exon 3 (figs. 2 and 3).

The coding regions of *Agph-DAB1* contain several features expected for functional *Mhc* genes, such as conserved cysteine codons and all of the conserved residues involved in salt bonds, glycosylation, disulfide bonds, the PBR, and peptide-binding listed in tables 1 and 3 of Kaufman, Salomonsen, and Flajnik (1994). The base compositions of the blackbird gene and a chicken gene are presented in table 1. Like chicken class II genes (Kaufman, Salomonsen, and Skoedt 1991; Kaufman 1995) *Agph-DAB1* is rich in G's and C's, particularly in introns and in the third positions of codons. We searched for regulatory sequences (e.g., X and Y boxes) found in chicken BL-B β2-microglobulin genes (Kroemer et al. 1990; Riegert et al. 1996) using blast searches of the 5′ region of *Agph-DAB1*, but the best match was a weak similarity to exonic chicken myosin smooth muscle light chain mRNA (GenBank accession number GGMLC2), a doubtful homology.

Survey of Polymorphism in the PBR

We designed two primers to amplify the PBR encoded by exon 2 (figs. 2 and 3). These primers were placed close to the exon borders so that they would amplify a small product that could eventually be analyzed by indirect methods of mutation detection and so that the amplified DNA would contain fewer polymorphisms found outside the PBR. The primers faithfully amplified a clean 289-bp product from eight blackbird genomic DNAs. Three to ten products per individual were cloned (table 2) and sequenced (fig. 4). The large number of variable sites was clustered in subdomains of the PBR known to be variable in chickens and mammals, with a total of 191 variable sites in the 270-bp coding region (fig. 4). Some pairs of sequences were very closely related (e.g., types 1a and 1b differing by four adjacent nucleotides in the α-helix region of the PBR), whereas other pairs were very divergent (e.g., types 3 and 6,

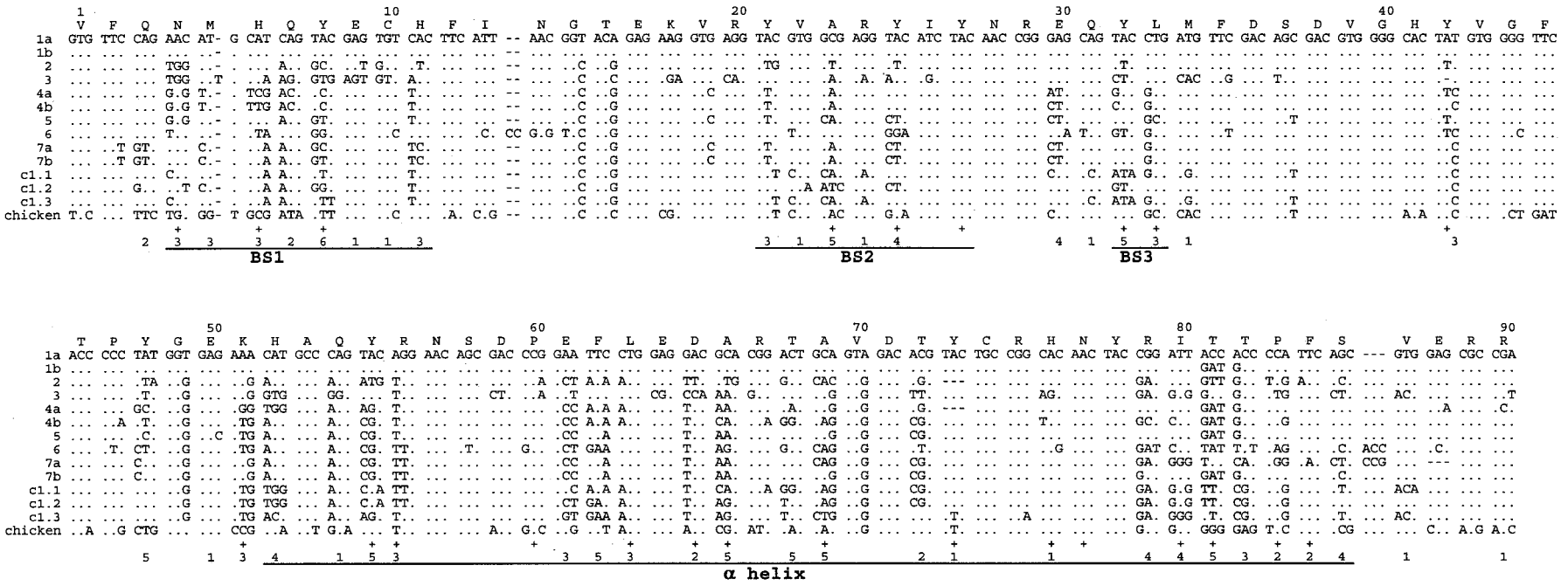

FIG. 4.—PBR sequences determined from eight blackbirds from Florida. Amino acid sequences of type 1a are indicated above the sequence, and codons are num bered beginning with the first PBR codon. The frameshift indels in types 3 and 6 are indicated, with dashes inserted in other sequences to preserve alignment. Four traditionally recognized subdomains (She et al. 1991) are indicated by solid lines. Plus signs indicate the 24 codons observed to contact peptide in a human allele (Brown et al. 1993). Numbers below various codons indicate the numbers of parsimony steps undergone by amino acids at those codons on the tree in fig. 5 after types 3 and 6 were removed (see text). c1.1–3 are blackbird cDNA clones from Edwards, Wakeland, and Potts (1995). The chicken sequence is from the B12 haplotype (Xu et al. 1989).

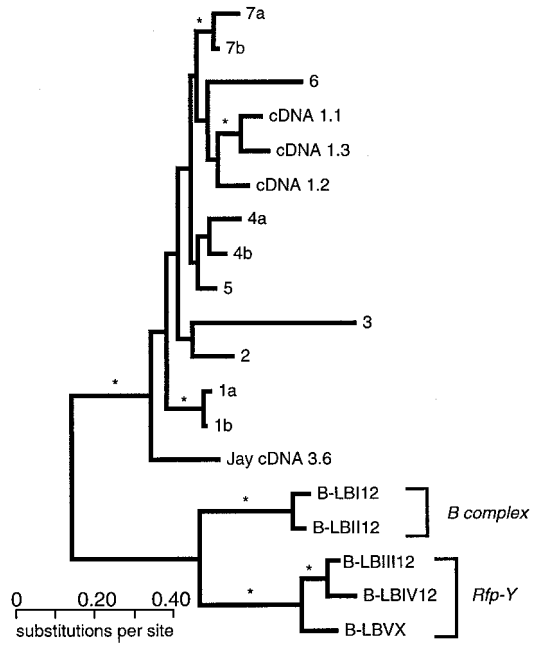

FIG. 5.—Neighbor-joining tree based on nucleotide variability in 10 PBR sequence types found in the PCR survey. The distance method used was the HKY85 model (Hasegawa, Kishino, and Yano 1985) with a gamma distribution ($\alpha = 0.5$), a general, albeit arbitrary, model of nucleotide change. The numbers at branch tips indicate types illustrated in figure 4 and table 2. The cDNA clones of blackbirds (1.1, 1.2, 1.3) and scrub jays (*Aphelocoma caerulescens*; 3.6) are from Edwards, Wakeland, and Potts (1995). The five chicken sequences correspond to class II B genes identified by cloning and PCR methods (Xu et al. 1989; Zoorob et al. 1990, 1993); the assigment of each gene to one of the two linkage groups in the chicken *Mhc*, the B-complex and *R-fpy*, is according to Miller et al. (1994, 1996). Asterisks indicate branches that received greater than 80% bootstrap support in 500 replicates. Branch lengths are proportional to number of nucleotide substitutions.

89.0% corrected divergence; see legend of fig. 4). Our alignment is compatible with two in-frame, single-codon indel polymorphisms near the 3′ ends of the sequences, whereas types 3 and 6, each retrieved multiple times from within and/or between three individuals, contained 1- and 2-bp indels, respectively (fig. 4). Phylogenetic analysis (fig. 5) resulted in a bushlike tree in which sequences did not fall into strongly marked groups separated from one another by long branches, and few clusters were well supported by bootstrap analysis. Previously sequenced cDNAs from red-winged blackbirds fall inside the genomic sequences but do not cluster strongly with any new type, whereas the genomic sequence cloned from individual 7 (fig. 3) corresponds to type 5, which, perhaps surprisingly, was not among four sequences amplified from this individual (figs. 4 and 5). As previously shown for exon 3 (Edwards, Wakeland, and Potts 1995), five chicken PBRs known to correspond to different loci fall into two groups corresponding to the two linkage groups of the chicken *Mhc* (fig. 5). The depth of the clade including these two groups is somewhat deeper than that spanned by the blackbird PBR sequences.

More than two sequences were retrievable from 6 of the 8 individuals (table 2). In all but one of these cases (individual 6), at least three of the amplified sequences were divergent enough from one another (>20%) to make it unlikely that the differences between sequences were solely the result of *Taq* errors or in vitro recombination of PCR products. Sequence type 1a (fig. 4) was the most commonly encountered sequence (18 of 41 clones, or 44%) and was found in all but one individual (table 2).

Table 3
Analysis of Anomolously Evolving Regions of Blackbird Class II *Mhc* Sequences

Beginning Nucleotide	Ending Nucleotide	Codons Spanned	PBR Subdomain	*Z* value
10	24	4–8	BS-1	3.81*
155	159	51–53	α-helix	5.35*
166	170	55–56	α-helix	3.45*
185	209	61–69	α-helix	5.48*
238	248	79–82	α-helix	10.30*

NOTE.—Beginning and ending nucleotides, codons spanned, and locations of PBR subdomains are indicated in figure 4.

* $P < 0.05$.

Patterns of Nucleotide Substitution

An analysis of synonymous and nonsynonymous substitutions was carried out on the 24 amino acids known to contact peptides in the PBR in a crystallized human allele (fig. 4 and table 4). Across all pairwise comparisons of the 10 types, the mean d_S was 0.303 ± 0.086 and the mean d_N was 0.520 ± 0.063 substitutions per site. Ratios of d_N to d_S ranged from 0.71 between the very similar types 1a and 1b to 4.89 between the very divergent types 1b and 5. A plot of d_N on d_S (not shown) indicates that the nonsynonymous substitutions at the 24 PBR sites of blackbird sequences are saturated among the most distant comparisons (maximum d_N = 88%). However, a tree built only on synonymous substitutions on the whole exon (maximum d_S = 34%) shares the characteristics of the tree in figure 5.

To detect regions of the PBR that are evolving in an anomalous fashion, potentially due to recombination, we conducted an SPV analysis using PLATO. Regardless of whether the two sequences with frameshift mutations were included, the anomalous regions identified in the analysis corresponded very closely to functional subdomains of the PBR (table 3). The variation in the BS-1 and α-helix displayed significantly anomalous substitution patterns relative to other regions; BS-2 and BS-3, which appear to be highly variable in the sequences (fig. 4), may have been missed because these segments are short (Grassly and Holmes 1997).

Discussion

Genomic cloning and characterization of *Mhc* genes is useful for molecular ecologists interested in examining immunological aspects of ecology and behavior of genetically poorly known vertebrates. In principle, knowledge of *Mhc* coding should set the stage for further PCR experiments to obtain flanking intron sequences that will facilitate analysis of *Mhc* polymorphisms in such a way that functionally important variability will

Table 4
Estimates of the Numbers of Synonymous (d_S, above diagonal) and Nonsynonymous (d_N, below diagonal) Substitutions per PBR Site and Standard Errors for 10 Types of *Mhc* Class II B Sequence from Red-Winged Blackbirds (type designations as per table 2 and fig. 4)

Type	1a	1b	2	3	4a	4b	5	6	7a	7b
1a		0.055 ± 0.055	0.481 ± 0.231	0.407 ± 0.191	0.378 ± 0.189	0.388 ± 0.185	0.157 ± 0.102	0.374 ± 0.171	0.282 ± 0.141	0.255 ± 0.134
1b	0.039 ± 0.028		0.378 ± 0.194	0.527 ± 0.236	0.288 ± 0.158	0.304 ± 0.158	0.092 ± 0.076	0.299 ± 0.148	0.370 ± 0.172	0.187 ± 0.112
2	0.661 ± 0.164	0.611 ± 0.152		0.508 ± 0.254	0.542 ± 0.277	0.430 ± 0.226	0.365 ± 0.198	0.397 ± 0.200	0.428 ± 0.212	0.455 ± 0.227
3	0.610 ± 0.155	0.604 ± 0.153	0.708 ± 0.178		0.615 ± 0.294	0.627 ± 0.289	0.426 ± 0.205	0.392 ± 0.184	0.178 ± 0.111	0.409 ± 0.196
4a	0.497 ± 0.128	0.423 ± 0.113	0.556 ± 0.137	0.628 ± 0.159		0.075 ± 0.076	0.133 ± 0.103	0.231 ± 0.136	0.262 ± 0.148	0.069 ± 0.070
4b	0.567 ± 0.141	0.489 ± 0.124	0.670 ± 0.163	0.572 ± 0.144	0.185 ± 0.064		0.200 ± 0.125	0.211 ± 0.123	0.279 ± 0.148	0.208 ± 0.125
5	0.524 ± 0.132	0.450 ± 0.117	0.569 ± 0.141	0.556 ± 0.141	0.266 ± 0.080	0.336 ± 0.093		0.231 ± 0.130	0.233 ± 0.130	0.062 ± 0.063
6	0.881 ± 0.223	0.812 ± 0.203	0.850 ± 0.212	0.882 ± 0.226	0.596 ± 0.150	0.495 ± 0.126	0.523 ± 0.132		0.176 ± 0.105	0.117 ± 0.085
7a	0.587 ± 0.148	0.623 ± 0.156	0.779 ± 0.193	0.640 ± 0.162	0.581 ± 0.146	0.465 ± 0.119	0.419 ± 0.111	0.453 ± 0.119		0.151 ± 0.098
7b	0.382 ± 0.104	0.319 ± 0.092	0.623 ± 0.154	0.487 ± 0.127	0.319 ± 0.092	0.335 ± 0.093	0.161 ± 0.059	0.446 ± 0.117	0.202 ± 0.068	

not be overlooked. Although this fast approach works for many species, it has been surprisingly inconclusive for class II genes of nonmodel avian species (personal observation; W. Potts, personal communication). This study presents the first intron/exon structure of an *Mhc* gene from a bird other than chicken. This structure revealed a gene organization similar to that known for chicken class II genes, although the sizes of introns were markedly larger than those of chicken class II B genes. Since the introns of mammalian class II B genes are often several kilobases (e.g., Begovich, Vu, and Jones 1990), the intron sizes found in *Agph-DAB1,* the blackbird class II gene we have cloned, are nonetheless consistent with the pattern that bird *Mhc* genes, and avian genes in general, are much smaller than those of orthologous mammalian genes (Hughes 1995).

Base Composition Suggests a Functional Gene

Is *Agph-DAB1* a functional gene? Without data on expression patterns, this remains ambiguous. However, a number of features suggest that it is at least functional, if not as polymorphic as some highly expressed chicken class II B genes. First, we found no frameshift mutations or premature stop codons throughout our sequence, and several structural features expected for functional class II genes are found in the PBR exon (see *Results*; fig. 3). Second, the base composition of *Agph-DAB1* is similar to (although slightly lower than) that of a functional chicken gene, and, more significantly, the base composition of coding sequences differs markedly from that of introns. For example, the base composition of third positions for *Agph-DAB1* is consistently higher (mean = 90.3 ± 11.0%) than that for introns and untranslated regions (69.0 ± 5.4%) for the blackbird gene. If the base composition of *Mhc*-coding regions is maintained by stabilizing selection (Bulmer 1991), we might expect even third positions to differ in base composition from noncoding sites, which are likely less constrained. However, if *Agph-DAB1* is nonfunctional and its base composition is primarily the result of directional mutation pressure without selection, then base composition of coding and noncoding regions would be expected to be similar. We also note a consistent trend toward increasing GC% as one moves from the beginning to the end of *Agph-DAB1,* although the basis for this trend is obscure. Third, others have noted a positive correlation between expression and polymorphisms of avian class II B genes (Zoorob et al. 1993). Although we cannot unambiguously quantify polymorphism in *Agph-DAB1,* since the amplified sequences are difficult to assign to a particular locus (see below), this pattern may not apply here.

Phylogenetic Consequences of Interlocus Gene Conversion

That the primers we designed to amplify the second exon appeared to amplify more than one *Mhc* locus is not unprecedented. More than one class II B locus can be amplified from the chicken *Mhc* using intron primers (Zoorob et al. 1993), although coamplification of loci is rarer in mammalian *Mhc* studies (see *Introduction*). In

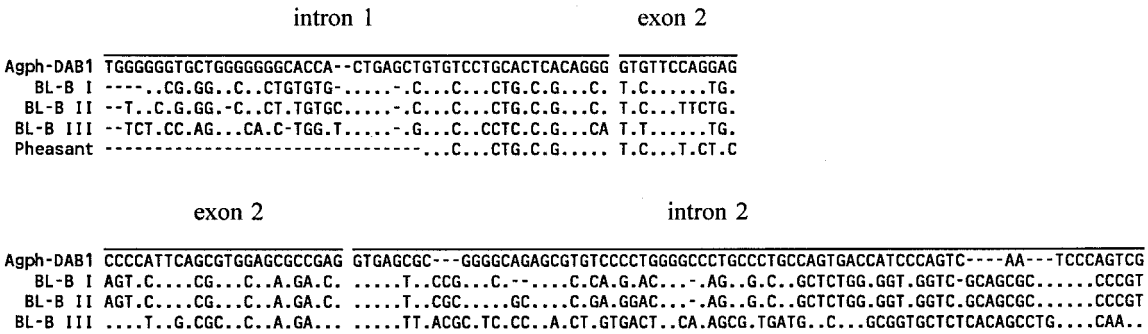

FIG. 6.—Comparison of intron sequences flanking exon 2 in various bird class II B genes. *Agph-DAB1* is from this study; BLB-I is from Xu et al. (1989); BLB-II and -III are from (Zoorob et al. 1990). The pheasant sequence in the top panel is from Wittzell et al. (1994). Note the similarity in intron sequences among different chicken genes (BLBI-III). Alignment was by hand.

six individual blackbirds, we retrieved up to three sequences, a result also consistent with the amplification of a minimum of two loci. This result becomes even less surprising when one compares all known avian intron sequences immediately flanking exon 2 (fig. 6). These flanking sequences, particularly intron 1, display a conservatism between blackbirds, chickens, and pheasants (Wittzell et al. 1994) in the first 30–40 bp that is unexpected if intron sequences are free to vary. Preliminary data from house finches (*Carpodacus mexicanus*), another songbird species, reveal sequences even more similar to those flanking *Agph-DAB1* (unpublished data). Such a pattern could result from selective constraints on the sequences at the ends of the introns or, since these sequences were cloned at random with respect to orthologs from other species, other forces promoting intergenic similarity, such as concerted evolution or recent duplication. Data supporting the latter interpretation come from the observation that similar sequences flank different paralogous chicken genes when compared to one another (Zoorob et al. 1990) (fig. 6). This suggests a pattern in which class II B genes of birds have duplicated rapidly, long after these species have diverged from common ancestors, and have retained flanking sequences from a gene in the common ancestor of songbirds and chickens (Edwards, Wakeland, and Potts 1995). The bushlike pattern observed in the tree of exon 2 sequences is similar to that found in a previous analysis of blackbird PBR diversity (Edwards, Wakeland, and Potts 1995) and also favors a model of recent common ancestry of loci. In figure 5, the blackbird portion of the tree is notably more bushlike than the chicken portion, a result that is robust to the inclusion of more chicken sequences in the analysis. Nonetheless, the short branches uniting genes within each of the two linkage groups defined in the chicken, the B complex and Rfp-y, suggest that these sequences display a similarly bushlike pattern (Edwards, Wakeland, and Potts 1995) (fig. 5). In contrast, phylogenetic analysis of PBR sequences from multiple mammalian genes and those of some birds (Vincek et al. 1997) reveals a distinctly nonbushlike pattern in which the PBRs of each gene are separated from one another by long branches (Gustaffson et al. 1991).

If all the amplified sequences corresponded to alleles at a single locus, a bushlike tree containing no long branches might be expected for alleles in a single randomly mating population, although in fact trees in which long branches occur near the base are more probable under these conditions (Felsenstein 1992; Hudson 1990). Similar genealogies would be expected for loci under balancing selection, such as *Mhc* loci, since the genealogy produced under balancing selection is simply that for a neutral gene on a longer timescale (Takahata 1990). However, because the sequences likely derive from multiple loci, a bushlike pattern could be produced by the steady occurrence of interlocus gene conversion, concerted evolution, or recent and very rapid successive gene duplications. There are few quantitative models available that specify the genealogical patterns of multiple loci experiencing interlocus gene conversion (Takahata 1994), but it is likely that, depending on the frequency and nature of the conversion events, this process would increase similarity among loci and hence yield shorter branches in phylogenetic trees of multiple loci (e.g., Drouin and Dover 1990; Dover 1993; Imanishi 1995). The results from phylogenetic analysis of chicken class II exons are consistent with the idea that frequent interlocus gene conversion occurs within each of the two linkage groups (Edwards, Wakeland, and Potts 1995) (fig. 5). Red-winged blackbirds have a large number of *Mhc* or *Mhc*-like sequences in their genomes, as judged by Southern blots (unpublished data). Until the blackbird class II B gene family is better defined, the hypothesis of frequent interlocus gene conversion within linkage groups must rest on less direct evidence.

We considered the possibility that the majority of patterns produced in the tree and sequence analyses are artifactual and plagued by PCR and cloning errors (L'Abbe et al. 1992; Bradley and Hillis 1997). We have quantified the extent of such errors for bird *Mhc* genes and various primers and have found that point mutations occur at a rate of 1 per 1,000 nucleotides, and recombination events occur in about 1 in 20 templates (Edwards, Grahn, and Potts 1995). However, we reject the idea that the bushlike nature of our tree is solely artifactual, because exactly the same techniques used on mammalian and avian *Mhc* loci nonetheless produced

recognizable clusters corresponding to different loci. Indeed, the vast majority of surveys of nucleotide polymorphism at *Mhc* loci are still performed using similar techniques. The large number of *Mhc*-like sequences in the blackbird genome undoubtedly contributes to the complexity of the patterns observed (unpublished data).

The two sequence types in which frameshift mutations occurred (types 3 and 6) probably represent real variants, and hence pseudogenes, since exactly the same sequences were retrieved from the same or different individuals (although a precise bias in *Taq* and cloning errors cannot be ruled out). Other *Mhc* pseudogenes have been discovered using similar methods (Brunsberg et al. 1996). When the 2-bp indel is removed from type 6 sequences, a premature stop codon (TAG) appears. The relatively long branches associated with these types (statistically significant [$\delta = 0.128$, $Z = 3.36$, $P < 0.01$] in the case of type 3 by a two-cluster test; Takezaki, Rzhetsky, and Nei 1995) are also consistent with those expected and found for nonfunctional avian sequences (Ota and Nei 1995). No class II pseudogenes are known from the chicken *Mhc,* although *Mhc*-like pseudogenes are numerous in mammalian genomes, and some chicken class II genes are expressed at very low levels.

Pseudogenes and Recombination

The patterns of nucleotide substitution displayed by the sequences suggest that they are influenced by balancing selection and or recombination/gene conversion. The high d_N/d_S ratios indicative of balancing selection are a well-known hallmark of *Mhc* genes, and the class II genes of birds, including chicken, appear to display similar patterns despite the difficulty of knowing allelic relationships. The SPV analysis identified at least two noncontiguous regions displaying anomalous nucleotide signatures. Although in principle this test alone cannot distinguish between selection and high mutation, if these results are interpreted in the context of segmental exchange, they suggest a gene conversion mechanism, since simple recombination at points within the PBR would have produced a pattern in which a single anomalous region is sequestered in one end of the sequences. Because such regions were separated by multiple regions that apparently evolve concordantly, the analysis suggests a pattern of gene conversion or motif-shuffling similar to that found in class II genes of a variety of mammals (Andersson et al. 1991; Erlich and Gyllensten 1991; Gyllensten, Sundvall, and Erlich 1991; She et al. 1991; Andersson and Mikko 1995). Regardless of mechanisms, the large divergences among the blackbird sequences are even more striking when considered against the backdrop of an extremely low level of mitochondrial DNA variation (<1%) in this species (Ball et al. 1988).

Our sequence analysis suggests several fruitful lines of further inquiry. If we knew which clusters in the tree of sequences corresponded to individual loci, then it might be possible to interpret even complex, multilocus patterns on, say, SSCP gels. Additionally, it is possible that locus-specific primers could be identified based on PBR polymorphisms. This approach would have the advantage of specificity but would yield incomplete PBR sequences for study. Another approach is to place primers farther away from the ends of the PBR exon. Our comparison of blackbird and chicken sequences suggests that this might be a useful approach, since the conservation of sequences in introns 1 and 2 appears to decay as one moves away from the flanking exons, so much so that the blackbird and chicken introns are essentially unalignable (figs. 3 and 6). PCR products generated in this way might be too long for analysis of polymorphism by rapid methods of mutation detection, but this is a small price to pay for locus specificity. On the other hand, if the conserved sequences flanking chicken and blackbird PBRs (fig. 6) prove to be taxonomically widespread, this could greatly facilitate molecular entry into the *Mhc* of genetically unstudied birds. Further studies of intron evolution in bird class II B genes should help in these endeavors.

Acknowledgments

We thank Jon Longmire and the Los Alamos National Laboratories for technical assistance. L. Gray provided the female blackbird sample from which the library was made. D. Garrigan kindly did the PLATO analysis, and H. Hoekstra helped with the molecular clock tests. We thank D. Garrigan, C. Hess, H. Hoekstra, N. Takahata, and two anonymous reviewers for very constructive comments on the manuscript and J. Klein for advice on nomenclature. This research was supported by NSF grant DEB9419738 to S.V.E.

LITERATURE CITED

ANDERSSON, L., K. GUSTAFSSON, A. K. JONSSON, and L. RASK. 1991. Concerted evolution in a segment of the first domain exon polymorphic MHC class II beta loci. Immunogenetics **33**:235–332.

ANDERSSON, L., and S. MIKKO. 1995. Generation of MHC class II diversity by intra- and intergenic recombination. Immunol. Rev. **143**:5–12.

APPLE, R. J., and H. A. ERLICH. 1992. Two new HLA DRB1 alleles found in African Americans: implications for balancing selection at positions 57 and 86. Tissue Antigens **40**: 69–74.

BALL, R. M. JR., S. FREEMAN, F. C. JAMES, E. BERMINGHAM, and J. C. AVISE. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. Proc. Natl. Acad. Sci. USA **85**:1558–1562.

BARBER, L. D., and P. PARHAM. 1993. Peptide binding to major histocompatibility complex molecules. Ann. Rev. Cell Biol. **9**:163–206.

BEGOVICH, A. B., T. H. VU, and P. P. JONES. 1990. Characterization of the molecular defects in the mouse E and E beta q genes. Implications for the origin of MHC polymorphism. J. Immunol. **144**:1957–1964.

BRADLEY, R. D., and D. M. HILLIS. 1997. Recombinant DNA sequences generated by PCR amplification. Mol. Biol. Evol. **14**:592–593.

BROWN, J. H., T. S. JARDETSKY, J. C. GORGA, L. J. STERN, R. G. URBAN, J. L. STROMINGER, and D. C. WILEY. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature **364**:33–39.

BRUNSBERG, U., I. EDFORS-LILJA, L. ANDERSSON, and K. GUSTAFSSON. 1996. Structure and organization of pig *MHC*

class II DRB genes: evidence for genetic exchange between loci. Immunogenetics **44**:1–8.

BULMER, M. 1991. The selection-mutation-drift theory of synonymous codon usage. Genetics **129**:897–907.

DIXON, B., S. H. M. VAN ERP, P. N. RODRIGUES, E. EGBERTS, and R. J. M. STET. 1995. Fish major histocompatibility complex genes: an expansion. Dev. Comp. Immunol. **19**: 109–133.

DOVER, G. A. 1993. Evolution of genetic redundancy for advanced players. Curr. Opin. Genet. Dev. **3**:902–910.

DROUIN, G., and G. A. DOVER. 1990. Independent gene evolution in the potato actin gene family demonstrated by phylogenetic procedures for resolving gene conversions and the phylogeny of angiosperm actin genes. J. Mol. Evol. **31**:132–150.

EDWARDS, S. V. 1996. Short and long term evoluton of Mhc class II B genes in birds: first glimpses. Pp. 169–178 *in* M. NEI and N. TAKAHATA, eds. Current topics on molecular evolution. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, and the Graduate School for Advanced Studies, Hayama, Japan.

EDWARDS, S. V., K. CHESTNUT, Y. SATTA, and E. K. WAKELAND. 1997. Ancestral polymorphism of Mhc class II genes in mice: implications for balancing selection and the mammalian molecular clock. Genetics **146**:655–668.

EDWARDS, S. V., M. GRAHN, and W. K. POTTS. 1995. Dynamics of *Mhc* evolution in birds and crocodilians: amplification of class II genes with degenerate primers. Mol. Ecol. **4**:719–729.

EDWARDS, S. V., and W. K. POTTS. 1996. Polymorphism of genes in the major histocompatibility complex: implications for conservation genetics of vertebrates. Pp. 214–237 *in* T. B. SMITH and R. K. WAYNE, eds. Molecular genetic approaches in conservation. Oxford University Press, Oxford.

EDWARDS, S. V., E. K. WAKELAND, and W. K. POTTS. 1995. Contrasting histories of avian and mammalian *Mhc* genes revealed by class II B sequences from songbirds. Proc. Natl. Acad. Sci. USA **92**:12200–12204.

ERLICH, H. A., and U. B. GYLLENSTEN. 1991. Shared epitopes among HLA class II alleles: gene conversion, common ancestry and balancing selection. Immunol. Today **12**:411–414.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 1992. Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. Genet. Res. **59**: 139–147.

———. 1994. Phylogeny inference package (PHYLIP). Version 3.6. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FIGUEROA, F., C. O'HUIGIN, H. TICHY, and J. KLEIN. 1994. The origin of the primate *Mhc-DRB* genes and allelic lineages deduced from the study of prosimians. J. Immunol. **152**: 4455–4465.

FLAJNIK, M. F., M. KASAHARA, B. P. SHUM, L. SALTER-CID, E. TAYLOR, and L. DU PASQUIER. 1993. A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian *Xenopus.* EMBO J. **12**:4385–4396.

GERMAIN, R. N., and D. H. MARGULIES. 1993. The biochemistry and cell biology of antigen processing and presentation. Annu. Rev. Immunol. **11**:403–450.

GIBBS, H., P. BOAG, B. WHITE, P. WEATHERHEAD, and L. TABAK. 1991. Detection of a hypervariable DNA locus in birds by hybridization with a mouse MHC probe. Mol. Biol. Evol. **8**:433–446.

GRASER, R., C. O'HUIGIN, V. VINCEK, A. MEYER, and J. KLEIN. 1996. Trans-species polymorphism of class II *Mhc* loci in danio fishes. Immunogenetics **44**:36–48.

GRASSLY, N. C., and E. C. HOLMES. 1997. A likelihood method for the detection of selection and recombination using nucleotide sequences. Mol. Biol. Evol. **14**:239–247.

GUSTAFFSON, K., U. BRUNSBERG, S. SIGURDARDOTTIR, and L. ANDERSSON. 1991. A phylogenetic investigation of MHC class II DRB reveals convergent evolution in the antigen binding site. Pp. 119–130 *in* J. KLEIN and D. KLEIN, eds. Molecular evolution of the major histocompatibility complex. Springer Verlag, Heidelberg.

GYLLENSTEN, U. B., D. LASHKARI, and H. A. ERLICH. 1990. Allelic diversification at the class II DQB locus of the mammalian major histocompatibility complex. Proc. Natl. Acad. Sci. USA **87**:1835–1839.

GYLLENSTEN, U., M. SUNDVALL, and H. ERLICH. 1991. Allelic diversity is generated by intra-exon exchange at the DRB locus of primates. Proc. Natl. Acad. Sci. USA **88**:3686–3690.

HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **21**:160–174.

HEDRICK, P. W. 1994. Evolutionary genetics of the major histocompatibility complex. Am. Nat. **143**:945–964.

HEDRICK, P. W., and T. J. KIM. 1998. Genetics of complex polymorphisms: parasites and maintenance of MHC variation. *In* R. S. SINGH and C. K. KRIMBAS, eds. Genetics, Evolution, and Society. Harvard University Press, Cambridge, Mass. (in press).

HEDRICK, P. W., and P. S. MILLER. 1995. Rare alleles, MHC and captive breeding. Pp. 187–204 *in* V. LOESCHCKE, J. TOMIUK, and S. K. JAIN, eds. Conservation genetics. Birkhauser, Basel, Switzerland.

HEDRICK, P. W., and G. THOMSON. 1983. Evidence for balancing selection at HLA. Genetics **104**:449–456.

HILL, A. V., J. ELVIN, A. C. WILLIS et al. (12 co-authors). 1992. Molecular analysis of the association of HLA-B53 and resistance to severe malaria. Nature **360**:434–439.

HUDSON, R. R. 1990. Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. **7**:1–44.

HUGHES, A. L. 1995. Small genomes for better flyers. Nature **377**:391.

HUGHES, A. L., and M. NEI. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci. USA **86**: 958–962.

IMANISHI, T. 1995. DNA polymorphisms shared among different loci of the major histocompatibility complex genes. Pp. 89–95 *in* M. NEI and N. TAKAHATA, eds. Current topics on molecular evolution. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, and the Graduate School for Advanced Studies, Hayama, Japan.

IMANISHI, T., and T. GOJOBORI. 1992. Patterns of nucleotide substitutions inferred from the phylogenies of the class I major histocompatibility complex genes. J. Mol. Evol. **35**: 196–204.

KAUFMAN, J. 1995. A ''minimal essential Mhc'' and an ''unrecognized'' Mhc: two extremes in selection for polymorphism. Immunol. Rev. **143**:63–88.

KAUFMAN, J., J. SALOMONSEN, and M. FLAJNIK. 1994. Evolutionary conservation of MHC class I and class II molecules—different yet the same. Semin. Immunol. **6**:411–424.

KAUFMAN, J., J. SALOMONSEN, and K. SKOEDT. 1991. Evolution of MHC molecules in nonmammalian vertebrates. Pp. 329–341 *in* J. KLEIN and D. KLEIN, eds. Molecular evolu-

tion of the major histocompatibility complex. Springer-Verlag, Berlin.

KLEIN, J. 1986. Natural history of the major histocompatibility complex. Wiley, New York.

KLEIN, J., R. E. BONTROP, R. L. DAWKINS, H. A. ERLICH, U. B. GYLLENSTEN, E. R. HEISE, P. P. JONES, P. PARHAM, E. K. WAKELAND, and D. I. WATKINS. 1990. Nomenclature for the major histocompatibility complexes of different species: a proposal. Immunogenetics **31**:217–219.

KLEIN, J., and C. O'HUIGIN. 1993. Composite origin of major histocompatibility complex genes. Curr. Opin. Genet. Dev. **3**:923–930.

KLEIN, J., Y. SATTA, and C. O'HUIGIN. 1993. The molecular descent of the major histocompatibility complex. Annu. Rev. Immunol. **11**:269–295.

KROEMER, G., A. BERNOT, G. BEHAR, A.-M. CHAUSSE, L.-N. GASTINEL, F. GUILLEMOT, I. PARK, P. THORAVAL, R. ZOOROB, and C. AUFFREY. 1990. Molecular genetics of the chicken MHC: current status and evolutionary aspects. Immnunol. Rev. **113**:119–145.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetic analysis. Version 1.01. The Pennsylvania State University, University Park.

L'ABBE, D., A. BELMAAZA, F. DECARY, and P. CHARTRAND. 1992. Elimination of heteroduplex artifacts when sequencing HLA genes amplified by polymerase chain reaction (PCR). Immunogenetics **35**:395–397.

LANYON, S. M. 1994. Polyphyly of the blackbird genus *Agelaius* and the importance of monophyly in comparative studies. Evolution **48**:679–693.

LI, W.-H. 1993. Unbiased estimation of the rates of synonymous and non-synonymous substitution. J. Mol. Evol. **36**: 96–99.

LONGMIRE, J. L., N. C. BROWN, L. J. MEINKE, K. L. ALBRIGHT, J. J. FAWCETT, E. W. CAMPBELL, R. K. MOYZIS, C. E. HILDEBRAND, and L. L. DEAVEN. 1993. Construction and characterization of partial digest DNA libraries from flow-sorted human chromosome 16. Genet. Anal. Tech. Appl. **10**:69–76.

MCCORMACK, W. T., L. W. TJOELKER, and C. B. THOMPSON. 1991. Avian B-cell development: generation of an immunoglobulin repertoire by gene conversion. Annu. Rev. Immunol. **9**:219–241.

MILLER, M. M., R. GOTO, A. BERNOT, R. ZOOROB, C. AUFFRAY, N. BUMSTEAD, and W. W. BRILES. 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.

MILLER, M. M., R. M. GOTO, R. L. TAYLOR JR., R. ZOOROB, C. AUFFRAY, R. W. BRILES, W. E. BRILES, and S. E. BLOOM. 1996. Assignment of *Rfp-Y* to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. Proc. Natl. Acad. Sci. USA **93**: 3958–3962.

MURRAY, B. W., S. MALIK, and B. N. WHITE. 1995. Sequence variation at the major histocompatibility complex locus DQβ in Beluga whales. Mol. Biol. Evol. **12**:582–593.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426.

NEI, M., and A. L. HUGHES. 1992. Balanced polymorphism and evolution by the birth-and-death process in the MHC loci. Pp. 27–38 *in* K. TSUJI, M. AIZAWA, and T. SASAZUKI, eds. Proceedings of the 11th histocompatibility workshop and conference. Oxford University Press, Oxford.

ORIANS, G. H., and L. D. BELETSKY. 1989. Red-winged blackbirds. Pp. 183–197 *in* I. NEWTON, ed. Lifetime reproduction in birds. Academic Press, New York.

OTA, T., and M. NEI. 1995. Evolution of immunoglobulin VH pseudogenes in chickens. Mol. Biol. Evol. **12**:94–102.

RADTKEY, R. R., B. BECKER, R. D. MILLER, R. RIBLET, and T. J. CASE. 1996. Variation and evolution of class I Mhc in sexual and parthenogenetic geckos. Proc. R. Soc. Lond. B **263**:1023–1032.

REYNAUD, C.-A., V. ANQUEZ, H. GRIMAL, and J.-C. WEILL. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. Cell **48**:379–388.

RIEGERT, P., R. ANDERSEN, N. BUMSTEAD et al. (12 co-authors). 1996. The chicken beta 2-microglobulin gene is located on a non-major histocompatibility complex microchromosome: a small, G+C-rich gene with X and Y boxes in the promoter. Proc. Natl. Acad. Sci. USA **93**:1243–1248.

ROBERTSON, K. A., and W. R. MCMASTER. 1985. Complete structure of a rat RT1 E beta chain extensive conservation of MHC class II beta chains. J. Immunol. **135**:4095–4099.

ROHWER, S., N. LANGSTON, and D. GORI. 1996. Body size and harem size in male red-winged blackbirds: manipulating selection with sex-specific feeders. Evolution **50**:2049–2065.

ROWEN, L., B. KOOP, and L. HOOD. 1996. The complete 685-kilobase DNA sequence of the human β T cell receptor locus. Science **272**:1755–1762.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAMMUT, B., V. LAURENS, and A. TOURNEFIER. 1997. Isolation of class I cDNAs from the axolotl *Ambystoma mexicanum.* Immunogenetics **45**:285–294.

SATO, A., F. FIGUEROA, C. O'HUIGIN, D. N. REZNICK, and J. KLEIN. 1996. Identification of major histocompatibility complex genes in the guppy, *Poecilia reticulata*. Immunogenetics **43**:38–39.

SATTA, Y., C. O'HUIGIN, N. TAKAHATA, and J. KLEIN. 1993. The synonymous substitution rate at major histocompatibility complex loci in primates. Proc. Natl. Acad. Sci. USA **90**:7480–7484.

———. 1994. Intensity of natural selection at the major histocompatibility complex loci. Proc. Natl. Acad. Sci. USA **91**:7184–7188.

SEARCY, A., and K. YASUKAWA. 1995. Polygyny and sexual selection in red-winged blackbirds. Princeton University Press, Princeton, N.J.

SHE, J. X., S. BOEHME, T. W. WANG, F. BONHOMME, and E. K. WAKELAND. 1990. The generation of MHC class II gene polymorphism in the genus *Mus.* Biol. J. Linn. Soc. **41**:141–161.

———. 1991. Amplification of MHC class II gene polymorphism by intra-exonic recombination. Proc. Natl. Acad. Sci. USA **88**:453–457.

SHELDON, B. C. and S. VERHULST. 1996. Ecological immunology: costly parasite defences and trade-offs in evolutionary ecology. Trends Ecol. Evol. **11**:317–321.

SHIINA, T., A. ANDO, T. IMANISHI, H. KAWATA, K. HANZAWA, T. GOJOBORI, H. INOKO, and S. WATANABE. 1995. Isolation and characterization of cDNA clones for Japanese quail (*Coturnix japonica*) major histocompatibility complex (*MhcCoja*) class I molecules. Immunogenetics **42**:213–216.

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

———. 1994. Comments on the detection of reciprocal recombination or gene conversion. Immunogenetics **39**:146–149.

TAKAHATA, N., and M. NEI. 1990. Overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics **124**:967–978.

TAKAHATA, N., Y. SATTA, and J. KLEIN. 1992. Polymorphism and balancing selection at major histocompatibility complex loci. Genetics **130**:925–938.

TAKEZAKI, N., A. RZHETSKY, and M. NEI. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. **12**:823–833.

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

VINCEK, V., D. KLEIN, R. T. GRASER, F. FIGUEROA, C. O'HUIGIN, and J. KLEIN. 1995. Molecular cloning of major histocompatibility complex class II B gene cDNA from the Bengalese finch *Lonchura striata.* Immunogenetics **42**:262–267.

VINCEK, V., C. O'HUIGIN, Y. SATTA, N. TAKAHATA, P. T. BOAG, P. R. GRANT, B. R. GRANT, and J. KLEIN. 1997. How large was the founding population of Darwin's finches? Proc. R. Soc. Lond. B **264**:111–118.

WEATHERHEAD, P. J. 1990. Secondary sexual traits, parasites, and polygyny in red-winged blackbirds, *Agelaius phoeniceus.* Behav. Ecol. **1**:125–130.

WEATHERHEAD, P. J., and G. F. BENNETT. 1991. Ecology of red-winged blackbird parasitism by haematozoa. Can. J. Zool. **69**:2352–2359.

WEATHERHEAD, P. J., and R. G. CLARK. 1994. Natural selection and sexual size dimorphism in red-winged blackbirds. Evolution **48**:1071–1080.

WILSON, R. K., B. F. KOOP, C. CHEN, N. HALLORAN, R. SCIAMMIS, and L. HOOD. 1992. Nucleotide sequence analysis of 95 kb near the 3′ end of the murine T-cell receptor α/δ chain locus: strategy and methodology. Genomics **13**:1198–1208.

WITTZELL, H., T. VON SHANTZ, R. ZOOROB, and C. AUFFREY. 1994. Molecular characterization of three Mhc class II B haplotypes in the ring-necked pheasant. Immunogenetics **39**:395–403.

XU, Y., J. PITCOVSKY, L. PETERSON, C. AUFFREY, Y. BOURLET, B. M. GERNDT, A. W. NORDSKOG, S. J. LAMONT, and C. M. WARNER. 1989. Isolation and characterization of three class II major histocompatibility complex genomic clones from the chicken. J. Immunol. **142**:2122.

ZOOROB, R., G. BEHAR, G. KROEMER, and C. AUFFREY. 1990. Organization of a functional chicken class II B gene. Immunogenetics **31**:179–187.

ZOOROB, R., A. BERNOT, D. M. RENOIR, F. CHOUKRI, and C. AUFFREY. 1993. Chicken major histocompatibility complex class II B genes: analysis of interallelic and interlocus sequence variance. Eur. J. Immunol. **23**:1139–1145.

NAOYUKI TAKAHATA, reviewing editor

Accepted November 14, 1997