

# Concerted Evolution of Two *Mhc* Class II *B* Loci in Pheasants and Domestic Chickens

Håkan Wittzell,<sup>1</sup> Alain Bernot,<sup>2</sup> Charles Auffray, and Rima Zoorob

Centre National de la Recherche Scientifique, UPR 420, Génétique Moléculaire et Biologie du Développement, Villejuif, France

The major histocompatibility complex (*Mhc*) of the ring-necked pheasant contains two polymorphic *Mhc* class II *B* genes. We show here, by screening of a cDNA library and RT-PCR from RNA, that both of these loci, *Phco-DAB1* and *Phco-DAB2*, normally are transcribed in the spleen. They differ mainly in the 3' untranslated (UT) region, with the transcript lengths, not including the poly(A) tails, being 1,100 nt for *DAB1* and 955 nt for *DAB2*. These two loci are orthologous to the *B-LBI* and *B-LBII* loci of the domestic chicken, respectively. *DAB1* and *DAB2* therefore seem to have evolved from a duplication before the split of the evolutionary lineages leading to the pheasant and the domestic chicken ca. 20 MYA. This is the first report of an orthologous relationship between avian *Mhc* genes. Yet, the third exons of *DAB1* and *DAB2* were identical in all available sequences and differed at 10 positions from the exon 3 sequences of *B-LBI/B-LBII*. The species-specific exon 3 suggests that *DAB1* and *DAB2* are subject to concerted evolution, i.e., interlocus genetic exchange. The exon 2 sequences show characteristic polymorphism, with hypervariable segments occurring in different combinations in different alleles. Given the divergence in the 3'UT region, the finding of the same exon 2 sequence at both the *DAB1* and the *DAB2* loci in one of the pheasant haplotypes also suggests that interlocus genetic exchange does occur. Accordingly, the exon 2 sequences tended to cluster irrespective of locus in the phylogenetic analyses. Genetic exchange simultaneously involving both exon 2 and exon 3 may be facilitated by the short length of the intervening intron (<100 bp) in pheasants and domestic chickens compared with, e.g., humans (about 3 kb).

## Introduction

The class I and class II molecules of the major histocompatibility complex (*Mhc*) play a central role in the immune system by binding and presenting peptides from degraded foreign proteins to cytotoxic T cells and helper T cells, respectively. The multitude of pathogens is thought to have led to extensive *Mhc* variation (Klein 1986), and this variation is probably enhanced by heterozygous advantage (Hughes and Nei 1988, 1989). For example, in the human *Mhc* there are loci with more than 150 known alleles (Parham and Ohta 1996). The *Mhc* has been studied for a large number of mammalian species, but comparatively little is known about *Mhc* structure and evolution in other vertebrates.

The *Mhc* of the domestic chicken, *Gallus gallus domesticus*, has so far been the only well-characterized avian *Mhc* (Guillemot et al. 1988; Kroemer et al. 1990). It deviates from the mammalian *Mhc* by its compact size and by consisting of two independently assorting clusters of *Mhc* genes: the serologically defined *Mhc* (*B* complex) and the so-called *Rfp-Y* system (Guillemot et al. 1988; Miller et al. 1994, 1996). Both the *B* complex and the *Rfp-Y* system contain class I and class II *B* genes, but the function of the *Rfp-Y* system is not known. The class II *B* genes of the *Rfp-Y* system belong to the so-called *B-LBIII* gene family, which is only weakly polymorphic and not represented in the *B* complex (Zoorob et al. 1993; Juul-Madsen et al. 1997). A

similar *Mhc* organization seems to characterize the *Mhc* (*MhcPhco*) of the ring-necked pheasant, *Phasianus colchicus* (Wittzell et al. 1994; Jarvi et al. 1996; von Schantz et al. 1996), but the second *Mhc* cluster, which possibly is homologous to the *Rfp-Y* system, is so far only known to contain class II *B*-like sequences (Wittzell et al. 1995).

The class II molecules are  $\alpha\beta$  heterodimers expressed on the surfaces of B cells, macrophages, and dendritic cells. Both the  $\alpha$  and the  $\beta$  chains contribute to peptide-binding, but the polymorphism resides primarily in the  $\beta$  chain (Klein 1986). No account of the  $\alpha$  chain encoding class II A genes has been published for birds so far, but there is an increasing number of shorter or longer class II *B* sequences available, especially from passerine songbirds (e.g., Edwards, Grahm, and Potts 1995; Edwards, Wakeland, and Potts 1995; Vincek et al. 1995, 1997). Edwards, Wakeland, and Potts (1995) found that, as expected, class II *B* sequences from three species of songbirds showed an excess of nonsynonymous nucleotide substitutions in the parts of exon 2 that code for the antigen-binding groove, whereas synonymous substitutions dominated in the remaining codons. However, they were unable to discern any orthologous relationships between the sequences from different songbird species. This was interpreted as suggesting a higher rate of concerted evolution, i.e., interlocus genetic exchange, in the class II *B* genes of birds than in those of mammals. In mammals, it is possible to recognize orthologous loci in such distantly related species as humans and mice (Trowsdale 1995). Interlocus genetic exchange leads to the homogenization of loci within species such that the orthology is lost. This process occurs in many multigene families (Li 1997), but it remains a controversial issue in *Mhc* evolution. Recently, Nei, Gu, and Sitnikova (1997) concluded that there was little evidence for concerted evolution in *Mhc* genes.

<sup>1</sup> Present address: Department of Theoretical Ecology, Lund University, Lund, Sweden.

<sup>2</sup> Present address: Centre National de Séquençage, Evry, France.

Key words: major histocompatibility complex, interlocus genetic exchange, gene conversion, transcription, ring-necked pheasant, birds.

Address for correspondence and reprints: Håkan Wittzell, Department of Theoretical Ecology, Lund University, Ecology Building, S-223 62 Lund, Sweden. E-mail: hakan.wittzell@teorekol.lu.se.

*Mol. Biol. Evol.* 16(4):479–490. 1999

© 1999 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

We previously characterized three class II *B* haplotypes from the *Mhc* (*B* complex) of pheasants by restriction fragment length polymorphism (RFLP) analysis (Wittzell et al. 1994). We also showed by PCR and DNA sequencing that two polymorphic class II *B* sequences cosegregated with each haplotype. Here, we study the transcription of these class II *B* genes and investigate their orthology with class II *B* genes from the domestic chicken.

## Materials and Methods

### Animals

The studied pheasants came from a farm population in the region of Tours, western France. They were brought in for autopsy as part of a routine survey of the local breeding stock of game birds. All birds appeared to be healthy. A blood sample was taken from each bird, and the spleen was removed and immediately frozen on liquid nitrogen right at the start of the autopsy.

### DNA Isolation and RFLP Typing

We isolated genomic DNA from whole blood using standard phenol-chloroform extraction (Sambrook, Fritsch, and Maniatis 1989). The *Mhc* genotype was established by RFLP analysis of the *Mhc* class II *B* and *Mhc* class I genes with the restriction enzymes *PvuII* for class II and *TaqI* for class I, and the cDNA probes *B-LBII* (Zoorob et al. 1990) and *F10* (Guillemot et al. 1988) from the domestic chicken, respectively. Southern blots were performed as previously described (Wittzell et al. 1994). The RFLP haplotypes have been defined by Wittzell et al. (1994) and von Schantz et al. (1996) unless otherwise stated.

### RNA Isolation and cDNA Library

Total RNA was isolated from the spleen using a kit from Stratagene (La Jolla, Calif.) based on the guanidinium thiocyanate method. The frozen spleen was homogenized in the denaturation solution followed by phenol-chloroform extraction. Messenger RNA was subsequently isolated with the Poly(A) Quik mRNA kit (Stratagene).

The cDNA library was created using the Superscript kit for cDNA synthesis and plasmid cloning (Gibco BRL, Gaithersburg, Md.) according to the manufacturer's protocol but with the Superscript Plus reverse transcriptase (Gibco BRL). This enzyme is a modified form of the cloned reverse transcriptase from the Moloney murine leukemia virus. Five micrograms of mRNA was primed for the first strand synthesis with a (dT)<sub>15</sub> primer-adaptor containing a restriction site for *NotI* (5'-GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)<sub>15</sub>). A *SalI* adaptor was added following the second strand synthesis, and the generated cDNA was digested with *NotI* and size-fractionated twice on Sephacryl S-500 HR Chromatography columns before ligating it to the *NotI*-*SalI* site of the plasmid vector pSPORT1. Finally, we introduced the ligated cDNA into *Escherichia coli* cells by electroporation.

Replica nitrocellulose filters containing  $5 \times 10^5$  clones from the unamplified cDNA library were

screened with the probe 1.5, which is a cloned genomic PCR fragment (407 bp) from a Swedish pheasant. The probe spans 244 bp of the second exon, the entire second intron, and 77 bp of the third exon, including primers, and it stems from the allele *DAB\*03* below. The excised and electroeluted fragment was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP (Amersham, Amersham, U.K.) using an oligolabeling kit (Pharmacia, Uppsala, Sweden). The filters were prehybridized for 4 h at 64°C in  $5 \times$  Denhardt's solution and 0.1% SDS and were thereafter hybridized overnight at 42°C in 50% formamide,  $6 \times$  SSC,  $5 \times$  Denhardt's solution, and 0.1% SDS. They were washed at 64°C with a final stringency of  $0.1 \times$  SSC and 0.1% SDS. The washed filters were exposed to X-ray film at -80°C with two intensifying screens (Eastman Kodak, Rochester, N.Y.).

### Reverse Transcriptase PCR

One microgram of total RNA from the spleen was reverse transcribed with the GeneAmp RNA PCR kit (Perkin Elmer Cetus, Norwalk, Conn.) and an oligo (dT) primer according to the manufacturer's protocol. The resulting cDNA was amplified in 100  $\mu$ l of  $1 \times$  PCR buffer, 0.2 mM of each of the four dNTPs, 2 mM MgCl<sub>2</sub>, 200 ng of each of the two primers, and 2.5 U of *Taq* polymerase using the thermal cycler GeneAmp PCR System 9600 (Perkin Elmer Cetus). The amplification started with heating the sample to 90°C for 90 s, followed by 30 cycles consisting of denaturation for 45 s at 94°C, annealing for 45 s at 60°C, and extension for 45 s at 72°C. The last cycle was followed by an additional extension for 10 min at 72°C. For the reverse transcriptase PCR (RT-PCR), we used the primers 471 (5'-GAGTGCCACTACCTGAACGGCACCGAGCGG) and B3 (5'-CGTCACGTAGCAGCCAGGCGGTC), which are complementary to the 5' parts of the second and third exons, respectively. For selectively amplifying the *DAB1\*06* allele from cDNA, we used the primers B27b (5'-ACCGAGCAGATTAGATTTCTAGAC), which matches the 5' part of exon 2 of this allele, and B7 (5'-GACGCCGTCTGAGCGACTTCT), which is specific for the 3' untranslated (UT) part of the *DAB1* locus.

### PCR from Genomic DNA

As a complement to the RT-PCR, we also made amplifications from genomic DNA of the same birds. For these experiments, we used 1  $\mu$ g of genomic DNA and the GeneAmp PCR kit (Perkin Elmer Cetus). The primers were int3 (5'-GCGGCTGTGTGCCTGACC) and DR1 (5'-GCTCCTCTGCACCGTGAAGGA), which correspond to the 3' part of the first intron and the 3' end of the second exon, respectively (Wittzell et al. 1994). The reaction conditions were the same as those described above for the PCR from reverse-transcribed RNA. For one bird, we also used the primer B44 (5'-ACCCAGCAGGTGAGGCATGTG) in combination with DR1 to check the sequence of the *DAB2\*0102* allele. B44 is complementary to the 5' part of exon 2 (Wittzell et al. 1994), and it preferentially gives the alleles *DAB1\*0101*, *DAB2\*0102*, and *DAB2\*04*. The

PCR fragments were cloned in the plasmid vector pCRII using the TA Cloning System (Invitrogen, San Diego, Calif.).

### Sequencing

Double-stranded plasmid DNA templates were prepared with a kit for miniprep purification (Promega, Madison, Wis.) and sequenced with Sequenase, version 2.0 (U.S. Biochemicals, Cleveland, Ohio), according to the manufacturer's protocol, which is based on the dideoxy chain-termination method (Sanger, Nicklen, and Coulson 1977). The cDNA clones were sequenced using a number of internal primers, which have been described by Zoorob et al. (1993) and Wittzell et al. (1994). In addition, we used the primers B5 (5'-GACAGCTACG TGTGCCGGGTG), B6 (5'-GGATGCTGAATTAGCT GCTGC), and B7 (see above). B5 corresponds to the 3' part of exon 3, and B6 corresponds to exon 6 and the beginning of the 3' untranslated region. Deletions in the cDNA clones were created by religation after excision of the *Pst*I-*Pst*I fragment and, when present, the *Sal*I-*Xho*I fragment. We furthermore subcloned the *Eco*RI-*Xho*I fragment from some cDNA clones in the plasmid vector pCRII.

Cloned PCR fragments were sequenced in both directions. All presented sequences of PCR fragments have been found in at least three identical clones and usually in independent amplifications, thus ruling out errors in the sequences due to the PCR technique.

### Sequence Analysis and Phylogeny

The sequences were analyzed with the Wisconsin software package from GCG (Genetics Computer Group 1994). The phylogenetic analysis is based on Jukes-Cantor distance estimates (Jukes and Cantor 1969) and the neighbor-joining method (Saitou and Nei 1987) in the computer program MEGA (Kumar, Tamura, and Nei 1993).

## Results

### Isolation of cDNA Clones

The cDNA library was created from the spleen of a heterozygous individual in order to isolate as many different full-length cDNA clones as possible. The bird was a male with the combined *Mhc* class II *B* and class I genotype *R<sub>2</sub>i/R<sub>3a</sub>*. The *R<sub>2</sub>i* haplotype has not previously been defined. It consists of the common *Mhc* class II *B* haplotype *R<sub>2</sub>* cosegregating with a new RFLP class I pattern, *i*, which differs from the class I pattern *f* (von Schantz et al. 1996) by lacking the shorter of the two hybridizing class I fragments of approximately 2.2 kb. Screening of 500,000 clones from the unamplified cDNA library showed abundant hybridizing clones, and 14 of these clones were isolated. The clones were preliminarily characterized by sequencing the 5' and 3' parts, including the hypervariable exon 2, and by restriction site analysis. This analysis showed that the isolated clones represented three different sequences, which we here have designated *Phco-DAB1\*0101*, *Phco-DAB2\*0102*, and *Phco-DAB2\*04* (EMBL accession numbers AJ224344, AJ224345, and AJ224348). The

choice of nomenclature will be addressed below. The longest clone from each allele was sequenced in its entire length.

The first allele, *DAB1\*0101*, was represented by nine clones. The longest one (cB9) contained 16 bp of the 5'UT region, an open reading frame of 263 codons, the stop codon TAG, and a 3'UT region consisting of 288 bp, i.e., in all 1,096 bp excepting the poly(A) tail.

The second allele, *DAB2\*0102*, was represented by three clones. The longest one (cB4.2) contained 12 bp of the 5'UT region and a reading frame that differed from that of *DAB1\*0101* only in a single base pair in exon 2. The difference in exon 2 (codon 51 is ATA in cB4.2, but ACA in *DAB1\*0101*) could not be verified in the two other clones from *DAB2\*0102*, as one of them was truncated in codon 56 and the other one was truncated in codon 75. We therefore made three independent amplifications from genomic DNA of the same bird using the primer pair int3-DR1 or B44-DR1. Among 22 sequenced PCR clones corresponding to the exon 2 sequence of *DAB1\*0101* or *DAB2\*0102*, all had ACA in codon 51. The same result was found in eight sequenced *DAB1\*0101/DAB2\*0102* clones from an RT-PCR using the primer pair 471-B3. This suggests that the difference in clone cB4.2 is due to a misincorporation during the cDNA synthesis. Despite the identical reading frame, *DAB2\*0102* differs from *DAB1\*0101* in having a much shorter 3'UT part of 143 bp (fig. 1). This is not simply due to an early termination of the *DAB1\*0101* transcript, since there are seven nucleotide substitutions in the beginning of the 3'UT region, and the last 68 nt of the *DAB2\*0102* transcripts differ completely from those of the corresponding part of *DAB1\*0101*. Furthermore, both *DAB1\*0101* and *DAB2\*0102* transcripts have a normal poly(A) signal (AATAAA) preceding the poly(A) tail.

The third allele, *DAB2\*04*, was represented by two isolated cDNA clones. The longest one (cB4.1) was truncated in codon 22 and lacked 51 bp of exon 2. The sequence of the 5' part of exon 2 of *DAB2\*04* was obtained by PCR from genomic DNA of the same individual using the primer pair int3-DR1. It could thus be established that *DAB2\*04* is identical to the previously characterized sequence 3.44, which segregates with the RFLP haplotype *R<sub>2</sub>* in Swedish pheasants (Wittzell et al. 1994). The 3'UT region of *DAB2\*04* is of the same length as that in *DAB2\*0102* and only differs in a single base pair (fig. 1). This difference, and another one in exon 4, could be verified in both cDNA clones from *DAB2\*04*.

The isolation of cDNA clones was supplemented by RT-PCR from mRNA from the spleen of the same *R<sub>2</sub>i/R<sub>3a</sub>* individual using the primers 471 (exon 2) and B3 (exon 3). Sequencing of 24 cloned RT-PCR fragments showed that a fourth class II *B* allele also was transcribed in this bird. The fourth allele is here designated *Phco-DAB1\*06* (EMBL accession number AJ224349). It corresponds to the sequence 3.27, which cosegregates with *DAB2\*04* in Swedish birds of the *R<sub>2</sub>* haplotype (Wittzell et al. 1994). This led us to amplify a longer fragment of *DAB1\*06* from mRNA using the



	10	20	30	40	50	60
<i>DAB1*0101</i>	<u>TAGCTGCTGCCCCGC*CGATCCGCTCCACCCG**CACCCACGCATCCCC*CACCGTCTCTCT</u>					
<i>DAB1*06</i>	.....*...G.....**.....C.....*G.....					
<i>B-LBI B<sup>6</sup></i>	.....AGT.GA..CTT...G.....TCA.....C.CGC..T..*GG.....G...					
<i>B-LBI pRP2-3</i>	.....*...G.....G.....**.....C...*..T..CGG.....G...					
<i>DAB2*0102</i>	.....*...G.....**.....G...CT.....*G.....CG...					
<i>DAB2*04</i>	.....*...G.....**.....G...CT.....*G.....CG...					
<i>B-LBII B<sup>6</sup></i>	.....*...G.....**.....C...*..T..*GG...CTG...					
<i>B-LBII B<sup>12</sup></i>	.....*...G.....**.....C...*..T..CGG.....G...					
<i>B-LBII pRP2-5</i>	.....*...G.....**.....C...*..T..CGG.....G...					
	70	80	90	100	120	130
<i>DAB1*0101</i>	CGGCTCTCCCTCGGGCTGCCCCCCC*CTGTGCCGCCCTCTGCCGCCGC*AGCGCCCGCTCT					
<i>DAB1*06</i>	.....*T.....*					
<i>B-LBI B<sup>6</sup></i>	.....CTCCA.....*..C.....T.....A..T.....					
<i>B-LBI pRP2-3</i>	.....CTCCA.....C..C.....T.....*..T.....					
<i>DAB2*0102</i>	.....ATTG...***T..C..CT.GAG...T..A..A.*GAT..A..T..T..					
<i>DAB2*04</i>	.....ATTG...***..C..CT.GAG...T..A..A.*GAT..A..T..T..					
<i>B-LBII B<sup>6</sup></i>	.....A..G...***..C..GTTGGAGA..T...A.*GAT..A...TC					
<i>B-LBII B<sup>12</sup></i>	.....A..G...***..C..GTTGGAGA..T...A.*GAT..A...TC					
<i>B-LBII pRP2-5</i>	.....A..G...***..C..GTTGGAGA..T...A.*GAT..A...TC					
	140	150	160	170	180	190
<i>DAB1*0101</i>	GTGCCCTCCCCAAGAAGAAGTCGCTCAGACGGCGTCCCGTTGTCTGCACATCCTTGGGG*TC					
<i>DAB1*06</i>	.....					
<i>B-LBI B<sup>6</sup></i>	..A.....***.....*.....G.....C....AA.					
<i>B-LBI pRP2-3</i>	.....G***.....G.....C....*A.					
<i>DAB2*0102</i>	..C...AT..T..T..ACGCG.TGA.TTTG					
<i>DAB2*04</i>	..C...AT..T..T..ACGCG.TGA.TTTG					
<i>B-LBII B<sup>6</sup></i>	..C...AT..T..T..ACGCG.TGA.TTTG					
<i>B-LBII B<sup>12</sup></i>	..C...AT..T..T..ACGCG.TGA.TTTG					
<i>B-LBII pRP2-5</i>	..*...AT..T..T..ACGCG.TGA.TTTG					
	200	210	220	230	240	250
<i>DAB1*0101</i>	CTTCTGTCTGCGACAG*****CAGTAGAGGGAAGCGGGAGGTCTGTGCTCTTCTATTCTCC					
<i>B-LBI B<sup>6</sup></i>	.G.....T.....G...TGACAGG.G..G..T.GT...C.....G...*					
<i>B-LBI pRP2-3</i>	.G.....T.....G...*****..GG.....G.....C.....C...*					
	260	270	280	290	300	310
<i>DAB1*0101</i>	TTCAGTGCCAGGTGGTGGTTTGGGTTCTTTAACGGAATAAATTCTTTTATT**GCAT					
<i>B-LBI B<sup>6</sup></i>	.....A.A..AA.....CA...T.C.....T..TT....					
<i>B-LBI pRP2-3</i>	.....A.A..AA.....CA...T.C.....T..TT....					

FIG. 1.—Nucleotide sequence of the 3'UT region of the *Mhc* class II *B* loci *DAB1* and *DAB2* in the ring-necked pheasant compared with the *B-LBI* and *B-LBII* loci in the domestic chicken. The *DAB1\*06* sequence was obtained by RT-PCR and is therefore not complete. The chicken sequences are two genomic clones from the haplotype *B<sup>6</sup>* (Xu et al. 1989), a cDNA sequence from the haplotype *B<sup>12</sup>* (Zoorob et al. 1990), and the cDNA clones pRP2-3 and pRP2-5 from undefined haplotypes (Pharr, Bacon, and Dodgson 1993). The stop codon is underlined. Gaps in the alignment are indicated by asterisk.

allele-specific primer B27b (exon 2) and the locus-specific primer B7 (3'UT region). The *DAB1\*06* sequence presented in figures 1 and 2 is a consensus sequence based on three sequenced clones. The 5' part of exon 2 was obtained by PCR from genomic DNA using the primers int3-DR1. The 3'UT region of *DAB1\*06* is homologous to that of *DAB1\*0101*, with only 4 bp differing in the part amplified here (fig. 1).

The presence of both *DAB1\*06* and *DAB2\*04*, which cosegregate in *R<sub>2</sub>* haplotypes (Wittzell et al. 1994), suggests that the two remaining sequences (*DAB1\*0101* and *DAB2\*0102*) segregate with the *R<sub>3a</sub>* haplotype.

#### Nomenclature

The four class II *B* sequences transcribed in the *R<sub>2i</sub>*/*R<sub>3a</sub>* individual belong to a single gene family, which we

designate *Phco-DAB* in accordance with a proposal to standardise the *Mhc* nomenclature for all vertebrates (Klein et al. 1990). This gene family consists of two loci, *Phco-DAB1* and *Phco-DAB2*, with a transcript length excluding the poly(A) tail of 1,100 nt and 955 nt, respectively, assuming that the complete 5'UT region is 20 nt as has been established by S1 nuclease mapping of a chicken *B-LBII* gene (Zoorob et al. 1990). The alleles from these two loci will be numbered in the order of discovery and irrespective of locus, since the sequence of the 3'UT region often will not be known, as in, e.g., *DAB\*02* below.

#### Comparison with the Domestic Chicken

The chicken *Mhc* class II *B* genes are called *B-LB* genes for historical reasons (Guillemot et al. 1988; Kroemer et al. 1990), and the loci are named after the

		$\beta_1$ domain (Exon 2)					
		10	20	30	40	50	60
DAB1*0101		FFLHGVI FECH FVNGTQQVRHVERDIHNRQQYAHFDSVDVGKYVADTPLGELQAEYWNNTTEY					
DAB2*0102		.....					
DAB*02		..Y.....	LLD.Q.Y.....	V.....	R.....	AL	
DAB*03		..D.....	ER..LL.....	F.....	AI		
DAB2*04		..D.....	Y.....	LL.....			
DAB*05		..R.MRY.....	FLD.Q.Y.....	TS.....	F.....	AI	
DAB1*06		..YYSTTS.....	Q...E.I.FLD.YFY.....	FT.....	F.....	PV..Q..S..QT	
DAB*07		.....Q.....	Y.H.NNY.....	LL.....	AL		
DAB*08		.....Q.....	Y.H.Q.Y.....		AL		
DAB*09		..Y.M.A.....	LLD.Q.Y.....	R.....	AL		
B-LBI B <sup>6</sup>		..QWTFKA.....	YL...ERA.FL...H.Y.....	FM.....	R...I...S.A.I		
B-LBI B <sup>12</sup>		..FC.A....YL...ER..YLD.E.Y.....	P.....	S.A.F			
B-LBI pRP2-3		..QWTFKA.....	YL...ERA.FL...H.Y.....	FM.....	R...I...S.A.I		
B-LBII B <sup>6</sup>		..FY.A.G....YL...ER..YL.E.Y.....	F.....	P.....	S.A.L		
B-LBII B <sup>12</sup>		..FC.A.S....YL...ER..YLQ.Y.Y.....	FT.....	F...S...P.....	S.A.L		
B-LBII pRP2-5		..FC.A....YL...ER..YLQ.Y.Y.....	LV.....	F.....	P.....	S.A.L	
		$\beta_2$ domain (Exon 3)					
		70	80	90	100	110	120
DAB1*0101		MEYRRGEVDRYCRHNYEGVESFTVQRR					
DAB2*0102		.....					
DAB*02		L.Q.MNA...T.....	VG.....				
DAB*03		L.Q.NN...TV.....	I.....	S			
DAB2*04		.....	GVF.....				
DAB*05		L.Q.MN...TV.....	I..				
DAB1*06		L..A..A..W.....					
DAB*07		L.Q.MHA.....	GVF.....				
DAB*08		L.Q.MHA.....	I.....				
DAB*09		L.Q.MN...TV.....	VG.....	S			
B-LBI B <sup>6</sup>		L.DEMNA...TF.....	GVG.....	S	.....+	.....+	.....+
B-LBI B <sup>12</sup>		..N.MN...TF.....	GVG.....	S	.....+	.....+	.....+
B-LBI pRP2-3		L.DEMNA...TF.....	GVG.....	S	.....+	.....+	.....+
B-LBII B <sup>6</sup>		L.NLMNIA.GP.....	GIL.....	S	.....+	.....+	.....+
B-LBII B <sup>12</sup>		L.N.MN...F.....	G.....	S	.....+	.....+	.....+
B-LBII pRP2-5		L.NIMNIA.GF.....	GIL.....	S	.....+	.....+	.....+
		130	140	150	160	170	180
DAB1*0101		EIEVKWFHNGREETERVVSTDMQNGDWTYQVLVVLEAVPRRGDSYVCRVEHASLRQPVSQPW					
DAB2*0102		.....					
DAB2*04		.....					
DAB1*06		.....					
B-LBI B <sup>6</sup>		.....L..+	.....	T.....	I..A.		
B-LBI B <sup>12</sup>		.....L..+	.....	T.....	I..A.		
B-LBI pRP2-3		.....L..+	.....	T.....	I..A.		
B-LBII B <sup>6</sup>		.....L..+	.....	T.....	I..A.		
B-LBII B <sup>12</sup>		.....L..+	.....	T.....	I..A.		
B-LBII pRP2-5		.....L..+	.....	T.....	I..A.		
		CP (Exon 4) Transmembrane domain				Cytoplasmic domain	
		190	200	210	220	230	
DAB1*0101		EPPADAGRSKLLTGVGGFVLGLVFLALGLVVFLRGHK				GRPVAAAP	GMLN
DAB2*0102		.....				.....	.....
DAB2*04		.....				.....	.....
DAB1*06		.....				.....	.....
B-LBI B <sup>6</sup>		.....	.....	F.....	Q.....	.....	.....
B-LBI B <sup>12</sup>		.....	.....			.....	.....
B-LBI pRP2-3		.....	.....	F.....	Q.....	.....	.....
B-LBII B <sup>6</sup>		.....	.....	F.....	Q.....	.....	.....
B-LBII B <sup>12</sup>		.....	.....	F.....	Q.....	.....	.....
B-LBII pRP2-5		.....	.....	F.....	Q.....	.....	.....

FIG. 2.—Predicted protein sequence for exons 2–6 of pheasant *DAB* genes in one-letter code compared with *B-LBI* and *B-LBII* genes in the domestic chicken. Residues are numbered as in the mature protein. Codons distinguished by a synonymous substitution compared with *DAB1\*0101* are indicated by plus signs in exons 3–6. Intron-exon boundaries are predicted from chicken class II *B* genes (Zoorob et al. 1990). *DAB\*05* corresponds to the previously described sequence 3.50 (Wittzell et al. 1994). It is not known whether this sequence is transcribed. The chicken sequences are the same as those in figure 1, along with the *B-LBI* gene from the haplotype *B*<sup>12</sup> (Zoorob et al. 1993). The latter gene is designated *B-LBI* by definition, although the 3'UT region has never been sequenced. CP = connecting peptide.

number of the gene in the chicken haplotype  $B^{12}$  (*B-LBI*, *B-LBII*, etc.) and not based on the gene family it belongs to. The present pheasant sequences are homologous to the so-called *B-LBII* gene family (Zoorob et al. 1993) among the class II *B* genes of the domestic chicken (Wittzell et al. 1994). Beside numerous PCR fragments of exon 2 (Zoorob et al. 1993), the *B-LBII* family is known from three full-length genomic clones (Xu et al. 1989; Zoorob et al. 1990) and a few cDNA clones (Zoorob et al. 1990; Moon Sung et al. 1993; Pharr, Bacon, and Dodgson 1993). The pheasant sequences are very similar to that of the prototypic *B-LBII* gene of the chicken haplotype  $B^{12}$  (Zoorob et al. 1990). The differences reside mainly in the hypervariable exon 2, which codes for the antigen-binding part of the *Mhc* class II molecule. In contrast, there are no differences from the *B-LBII* gene in exon 1 (results not shown) or in exons 5 and 6 (fig. 2). Exon 3, coding for the extracellular  $\beta_2$  domain, differs by 10 bp in eight codons between the pheasant sequences and the *B-LBII* gene, leading to only 4 amino acid substitutions (fig. 2).

The most interesting observation stems from a comparison of the 3'UT regions of the *Mhc* class II *B* genes in both species. The same variation with short and long transcripts has also been seen in the domestic chicken (Zoorob et al. 1990; Moon Sung et al. 1993; Pharr, Bacon, and Dodgson 1993). The 3'UT region is short in the *B-LBII* locus of the  $B^{12}$  haplotype (Zoorob et al. 1990), and Pharr, Bacon, and Dodgson (1993) assumed that long transcripts emanate from the *B-LBI* locus. However, only in one case have both kinds of 3'UT regions been cloned from a chicken which was known to be homozygous ( $B^6/B^6$ ) at the *Mhc* (Xu et al. 1989). The genomic clones in question can easily be aligned to the 3'UT regions from the pheasant genes (fig. 1). The 3'UT region of the chicken *B-LBI* locus (genomic clone CCII-4 in Xu et al. 1989) is thus orthologous to the two *DAB1* alleles found in the pheasant cDNA library, whereas the 3'UT region of the chicken *B-LBII* locus (genomic clone CCII-7) corresponds to the two pheasant *DAB2* alleles (fig. 1). The results presented here generalize the results from the chicken haplotype  $B^6$  and suggest that both the *Mhc* of the domestic chicken and that of the ring-necked pheasant contain two class II *B* loci of the *DAB* gene family, viz. the *DAB1* (*B-LBI*) locus and the *DAB2* (*B-LBII*) locus, which are very similar except for the 3'UT region.

#### Transcription in an $R_1$ Homozygote

Gene number and transcription were also studied for a pheasant homozygous for the previously characterized *Mhc* class II *B* and class I haplotype  $R_1a$ . In the RT-PCR, we used total RNA extracted from the spleen. Twenty cloned PCR fragments were sequenced, and these clones represented two alleles, *Phco-DAB\*02* and *Phco-DAB\*03* (EMBL accession numbers AJ224346 and AJ224347), corresponding to the sequences 3.2 and 3.3, respectively, which have already been found in Swedish birds with the same RFLP haplotype (Wittzell et al. 1994). The lack of intron sequence ruled out contamination from genomic DNA also in this case, and

supplementary amplifications from genomic DNA confirmed that the 5' part of the second exon did not differ from what we established previously. Furthermore, no other sequences were found during the genomic amplifications, suggesting that this haplotype is only associated with these two class II *B* genes, which are both transcribed.

#### Transcription in a Heterozygote

Finally, we also studied a bird which appeared to be homozygous ( $R_3/R_3$ ) according to the RFLP typing of the *Mhc* class II *B* genes. However, typing of the class I genes clearly suggested that the bird was heterozygous at the *Mhc*, but the extended haplotypes cannot be defined here, as they have not been encountered in other investigated pheasants. The RT-PCR from total RNA from the spleen, just as the amplification of genomic DNA, gave four different class II *B* alleles, thus supporting the notion that the bird possessed two different  $R_3$  haplotypes. Three of the alleles (*Phco-DAB\*07*, *Phco-DAB\*08*, and *Phco-DAB\*09*; EMBL accession numbers AJ224350–AJ224352) have not previously been encountered, whereas the fourth, at least in exon 2 and the available part of exon 3, is identical to *DAB2\*04*. The fact that the latter allele has thereby been found in both an  $R_2$  and an  $R_3$  haplotype could be interpreted as a PCR contamination. However, we find this explanation unlikely, since the same result was obtained with both standard PCR and RT-PCR, i.e., with and without intron sequences, respectively. The presence of *DAB2\*04* in two different haplotypes may instead be the result of a genomic recombination.

The RT-PCR showed that both *DAB1* and *DAB2* are transcribed in all studied haplotypes, but it did not allow us to evaluate the relative rates of transcription of the two loci. However, the isolation of zero and nine cDNA clones of the *DAB1* alleles versus two and three cDNA clones of the *DAB2* alleles of the same  $R_2/R_3a$  bird does not suggest that there are any consistent differences in transcription between the loci.

#### Phylogeny

A phylogenetic tree of the second exons of pheasant and domestic chicken alleles suggests that all pheasant sequences except *DAB1\*06* are more closely related to each other than to any of the chicken sequences (fig. 3). The tree was obtained by the neighbor-joining method (Saitou and Nei 1987) with the human *HLA-DRB\*0101* allele as an outgroup. An identical tree with only minor changes of the bootstrap values was obtained using the sequence 2.4 (Edwards, Wakeland, and Potts 1995) from a scrub jay, *Aphelocoma coerulescens*, as the outgroup instead (results not shown). Some chicken alleles from the *B-LBII* locus tend to cluster, but overall there is no clear difference between exon 2 sequences from *DAB1* (*B-LBI*) and *DAB2* (*B-LBII*), neither in the pheasant nor in the domestic chicken. This analysis is restrained by the large number of alleles that still have not been assigned to loci, but it can be seen that only small clusters of two to five sequences (or eight in the pheasant cluster) are supported by the bootstrap test. The

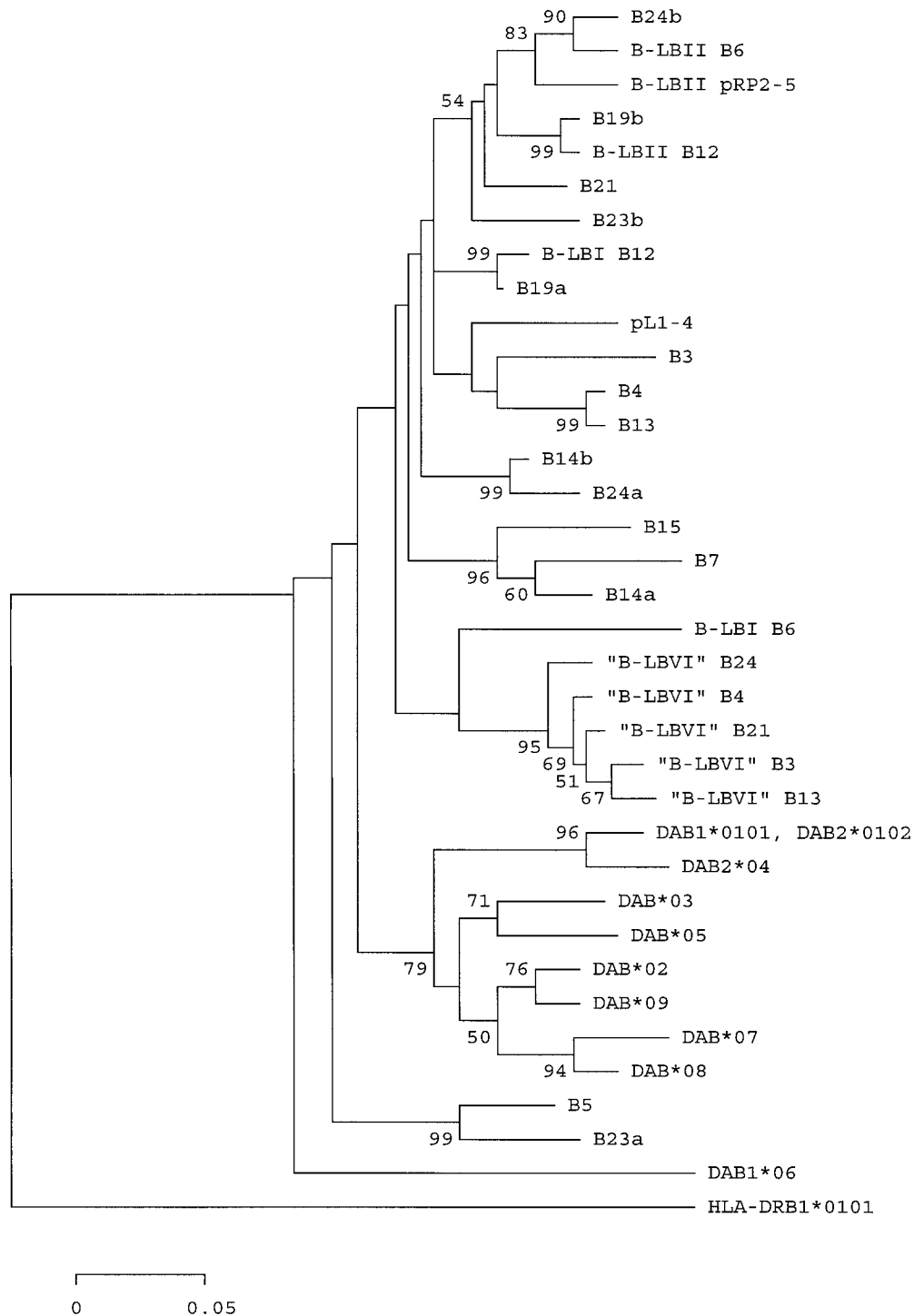


FIG. 3.—Neighbor-joining tree of the hypervariable exon 2 (codons 6–87) of *Mhc* class II *B* genes from ring-necked pheasants (*DAB*) and domestic chickens (others), with a human *DRB* sequence as outgroup (*HLA-DRB1\*0101*; Bell et al. 1987). Numbers at interior branches are bootstrap values based on 500 replications. Values less than 50% are not shown. The scale bar indicates genetic distance in units of nucleotide substitution per site. The chicken sequences are presented in figures 1 and 2, except for the truncated cDNA clone pL1-4 (Pharr, Bacon, and Dodgson 1993) and 21 amplified sequences indicated by the *Mhc* haplotype they stem from (*B*<sup>3</sup>, etc.; Zoorob et al. 1993, corrected for printing errors).

relation between these groups is much more uncertain. It should also be noted that a group of related chicken sequences, which were called the *B-LBVI* gene family by Zoorob et al. (1993) due to their obvious relatedness

to each other, forms a single well-supported cluster, but falling within the variation of the *B-LBI* and *B-LBII* loci.

The phylogenetic tree in figure 3 is based on the entire exon 2 sequence except for the last seven codons.

It is known from mammals that the 5' part of the exon, which codes for the  $\beta$ -sheet, may evolve differently from the 3' part coding for the  $\alpha$ -helix (Gyllenstein, Sundvall, and Erlich 1991; Schwaiger et al. 1993). When the  $\beta$ -sheet and  $\alpha$ -helix coding regions (codons 6–53 and 54–87, respectively) of the sequences in figure 3 are analyzed separately, new relationships emerge in some cases. However, overall, the main branches are short in both trees (results not shown). A more complicated pattern of variation is instead apparent when some of the most polymorphic residues in different parts of exon 2 are considered (table 1). Taken together, the 10 selected residues are unique, except in the chicken sequences  $B^{14b}$  and  $B^{24}$ . Residue 72 was selected for the comparison, since Arg<sub>72</sub> is a salt bridge constituent in human *Mhc* class II molecules (Brown et al. 1993). This renders the residue structurally important, and it is invariable in alleles from humans and mice (Brown et al. 1988). It is nevertheless polymorphic in chicken and pheasant alleles, and the main variants (methionine and arginine) are found in seemingly unrelated alleles in both species (table 1). The other selected residues all line the antigen-binding groove of *HLA-DR1* (Brown et al. 1993). Among the latter ones, residue 47 shows a discrete pattern of variation, being either phenylalanine (codon TTT), tyrosine (TAC), or, in the “*B-LBVI*” sequences, histidine (CAC). TTT and TAC can still be found in sequences from both species. For the rest of the selected residues, the same amino acids are often found in both species, but the codons are more variable, and point mutations may have played a role in generating the variation. Convergent selection has also been operating, as can be seen in residue 13, where serine (AGT) is found in *DAB1\*06* and one chicken sequence, whereas serine (TCC or TCG) is found in some other chicken sequences. It is nevertheless tempting to assume that common ancestry, rather than convergent selection, is the reason for the sequence motif threonine-serine in residue 37–38, which is always encoded ACATCG and shared by *DAB\*05* and three chicken sequences (table 1).

## Discussion

### Concerted Evolution

Although generally accepted and well documented for other multigene families (Li 1997), concerted evolution has remained a controversial issue in the case of *Mhc* genes. One reason may be the apparent contradiction between the homogenizing effect of concerted evolution and the unrivaled polymorphism of *Mhc* genes. To create polymorphism in *Mhc* class II *B* genes rather than extinguishing it, genetic exchange should mainly encompass short sequences and, in particular, only parts of the hypervariable exon 2.

The presence of a *DAB1* locus and a *DAB2* locus with distinct 3'UT regions in both pheasants and domestic chickens suggests that these two loci arose through a duplication before the split of the evolutionary lineages leading to these species. The split may be estimated to have taken place about 20 MYA (Helm-By-

chowski and Wilson 1986). The observation that the *DAB* exon 3 sequences nevertheless are species-specific and not locus-specific indicates that the loci are subject to interlocus genetic exchange. The occurrence of the same exon 2 sequence in two different alleles (*DAB1\*0101* and *DAB2\*0102*) with distinctly different 3'UT regions is also compatible with concerted evolution. Edwards, Wakeland, and Potts (1995) found that the exon 3 sequences clustered in a species-specific manner when they studied exons 2 and 3 of the class II *B* genes of three species of passerine songbirds. The lack of orthology was interpreted as indicating a much higher level of concerted evolution in the class II *B* genes of birds than in those of mammals. The results presented here suggest that it still may be possible to distinguish orthologous loci in different bird species if the 3'UT region is characterized.

The general conclusion of our study and the one by Edwards, Wakeland, and Potts (1995) is nevertheless that concerted evolution is an important feature in the evolution of avian class II *B* genes. Recently, however, Nei, Gu, and Sitnikova (1997) found little evidence that concerted evolution played a role in the evolution of the *Mhc* and immunoglobulin multigene families when the genes from different classes of vertebrates were compared. Instead, they suggested that the death-and-birth process of evolution, by which some *Mhc* genes are deleted and other ones duplicated and subsequently differentiated, is a much more prominent force than concerted evolution. Klein and O'hUigin (1995) had previously proposed that contraction and expansion of the number of genes in the same manner characterize the evolution of the *DRB* gene family in mammals. We cannot see that concerted evolution and the death-and-birth process should be mutually exclusive. Rather, it may be a question of timescale, with concerted evolution operating on closely related loci such as *DAB1* and *DAB2* in pheasants and domestic chickens, whereas the birth-and-death process in the long run will lead to a reorganization and, eventually, to a loss of orthology, as can be seen for the class II *B* genes in birds versus mammals. In this perspective, the suggested interlocus genetic exchange between, e.g., a functional pig *DRB* gene and two putative *DRB* pseudogenes (Brunsberg et al. 1996) is not necessarily unreconcilable with the claim that the death-and-birth process has played an important role in the long-term evolution of the mammalian *DRB* gene family.

Nei, Gu, and Sitnikova (1997) did not completely rule out the possibility that short gene segments can be exchanged between related loci, but they stressed that human *Mhc* loci from different gene families such as *DRB*, *DQB*, and *DPB* almost always produce monophyletic clusters. A possible explanation is that the probability of gene conversion declines with increasing sequence divergence (Liskay, Letsou, and Stachelek 1987). Concerted evolution can therefore be expected to play a role primarily within gene families. Accordingly, there is no indication that the so-called *B-LBIII* gene family of the chicken *Rfp-Y* system, which is separated from the *Mhc* (*B*) complex by a genetically unstable



**Table 1**  
**Polymorphic Codons from Different Parts of the Hypervariable Exon 2 of *Mhc* Class II *B* Genes of Pheasants (*DAB*) and Domestic Chickens (others) Compared with a Consensus Sequence**

	CODON								RESIDUE							
	9	11	13	30	37–38	47	72	85–86	9	11	13	30	37	47	72	85
<i>Consensus</i> .....	TGC	GTG	TCT	CAA	TACGCG	TAC	ATG	GTGTGTG	C	A	A	Q	YA	Y	M	VV
<i>B</i> <sup>24b</sup> .....		.C.	GG.	G..	.....	.TT	...	A..CG.	.	.	G	E	..	F	.	IR
<i>B-LBII B</i> <sup>6</sup> .....	.A.	.C.	GG.	G..	.....	.TT	...	A..C..	Y	.	G	E	..	F	.	IL
<i>B-LBII</i> pRP2-5 .....		.C.	.T.	T.C	.TA.T.	.TT	...	A..C..	.	.	F	Y	LV	F	.	IL
<i>B</i> <sup>19b</sup> .....		.C.	..C	T.C	.....	.TT	...	.G....	.	.	S	Y	..	F	.	G.
<i>B-LBII B</i> <sup>12</sup> .....		.C.	..C	T.C	.T.A..	.TT	...	.G....	.	.	S	Y	FT	F	.	G.
<i>B</i> <sup>21</sup> .....		.C.	..C	...	.T....	.TT	...	A..C..	.	.	S	.	F.	F	.	IL
<i>B</i> <sup>23b</sup> .....		.C.	G..	...	.T.A..	.TT	...	...T.C	.	.	.	.	FT	F	.	.F
<i>B-LBI B</i> <sup>12</sup> .....		.C.	.T.	G..	.....	...	...	...G.	.	.	F	E	..	.	.	.G
<i>B</i> <sup>19a</sup> .....		.C.	.T.	G..	.....	...	...	.....	.	.	F	E	..	.	.	..
pL1-4 .....	C..	A..	.A.	...	ACAT..	...	.G.	.....	R	M	Y	.	TS	.	R	..
<i>B</i> <sup>3</sup> .....		.C.	AG.	T.T	.T.A..	.TT	.G.	.G....	.	.	S	Y	FT	F	R	G.
<i>B</i> <sup>4</sup> .....	.A.	AG.	G..	..T	.TAAT.	.TT	.G.	...AAT	Y	R	.	H	LM	F	R	.N
<i>B</i> <sup>13</sup> .....	.A.	AG.	G..	...	.TAAT.	.TT	.G.	...AAT	Y	R	.	.	LM	F	R	.N
<i>B</i> <sup>14b</sup> .....	.A.	A..	G..	...	.T....	...	...	...AAT	Y	M	.	.	F.	.	.	.N
<i>B</i> <sup>24a</sup> .....	.A.	A..	G..	...	.T....	...	...	...AAT	Y	M	.	.	F.	.	.	.N
<i>B</i> <sup>15</sup> .....	.A.	...	.T.	...	A..AT.	...	...	...AT	Y	V	F	.	NM	.	.	.D
<i>B</i> <sup>7</sup> .....	CC.	...	.T.	...	ACAT..	...	...	T.G.G.	P	V	F	.	TS	.	.	LG
<i>B</i> <sup>14a</sup> .....	.A.	A..	.TG	...	ACAT..	...	...	.....	Y	M	L	.	TS	.	.	..
<i>B-LBI B</i> <sup>6</sup> .....	..G	T.T	G.A	..C	.T.AT.	...	...	...G.	W	F	.	H	FM	.	.	.G
“ <i>B-LBVI</i> ” <i>B</i> <sup>24</sup> .....	..G	.G.	G.A	..T	...T..	C..	...	.....	W	G	.	H	.S	H	.	..
“ <i>B-LBVI</i> ” <i>B</i> <sup>4</sup> .....	..G	T.T	G.A	T.T	.....	C..	...	.....	W	F	.	Y	..	H	.	..
“ <i>B-LBVI</i> ” <i>B</i> <sup>21</sup> .....	..G	T..	G.A	T.T	.....	C..	...	.....	W	L	.	Y	..	H	.	..
“ <i>B-LBVI</i> ” <i>B</i> <sup>3</sup> .....	..G	T.T	G.A	T.T	...A..	C..	...	.....	W	F	.	Y	.T	H	.	..
“ <i>B-LBVI</i> ” <i>B</i> <sup>13</sup> .....	...	T.T	G.A	T.T	.....	C..	...	.....	.	F	.	Y	..	H	.	..
<i>DAB1*0101, DAB2*0102</i> ....	CA.	...	.T.	G.C	.....	...	CG.	.G....	H	V	F	D	..	.	R	G.
<i>DAB2*04</i> .....	GA.	...	.T.	G.C	CT.TT.	...	CG.	..GT.T	D	V	F	D	LL	.	R	.F
<i>DAB*03</i> .....	GA.	...	.T.	G.C	.....	.TT	.AT	A....	D	V	F	D	..	F	N	I.
<i>DAB*05</i> .....	C..	A..	.AC	...	ACAT..	.TT	...	A....	R	M	Y	.	TS	F	.	I.
<i>DAB*02</i> .....	.A.	...	.T.	...	...T.	...	...	...G.	Y	V	F	.	.V	.	.	.G
<i>DAB*09</i> .....	.A.	A..	G..	...	.....	...	...	...G.	Y	M	.	.	..	.	.	.G
<i>DAB*07</i> .....	CA.	...	.T.	A.T	CT.TT.	...	...	..GT.T	H	V	F	N	LL	.	.	.F
<i>DAB*08</i> .....	CA.	...	.T.	...	.....	...	...	A....	H	V	F	.	..	.	.	I.
<i>B</i> <sup>5</sup> .....	..G	.C.	...	...	...A.	...	GG.	...T..	W	.	S	.	.E	.	G	.L
<i>B</i> <sup>23a</sup> .....		A..	...	...	.....	...	CG.	...T.C	.	M	S	.	..	.	R	.F
<i>DAB1*06</i> .....	.A.	AC.	AG.	T.T	.T.A.A	.TT	CG.	.G....	Y	T	S	Y	FT	F	R	G.

NOTE.—The translations of the codons are given to the right in one-letter code. The sequences are arranged in the same order as in the phylogenetic tree of exon 2 (fig. 3).

region (Miller et al. 1996), is subject to genetic exchange with *B-LBI/B-LBII* (Zoorob et al. 1990, 1993), despite the fact that all these class II *B* genes seem to have evolved from a common ancestor after the radiation of galliform birds and passerine songbirds (Edwards, Wakeland, and Potts 1995). It should be added that the *B-LBI* and *B-LBII* loci are separated by less than 6 kb (Guillemot et al. 1988). Interlocus genetic exchange involving both exons 2 and 3 may furthermore be facilitated in the *DAB* gene family by the short length of the intervening intron 2, which is, e.g., only 86 bp in *DAB\*03* (unpublished data; cf. probe 1.5 in *Materials and Methods*) and *B-LBII B<sup>12</sup>* (Zoorob et al. 1990). In contrast, the length of the corresponding intron is about 3 kb in human class II *B* genes (Boss and Strominger 1984; Kappes et al. 1984). If both exons 2 and 3 are involved simultaneously in the genetic exchange, then intron 2 is also expected to evolve in a species-specific and not a locus-specific manner. This cannot be judged presently, due to the paucity of available intron sequences from birds. Some species-specific positions in exon 4 suggest that the latter exon also may have been involved in interlocus genetic exchange, but the few available sequences are slightly less homogeneous than exon 3 (fig. 2). The exact mechanism behind the genetic exchange is not known, but there are appropriate means for gene conversion in birds. Hence, there is only one functional gene for the immunoglobulin segment *V<sub>H</sub>* in the domestic chicken, and diversity is created by somatic gene conversion between about 100 *V<sub>H</sub>* pseudogenes and the single functional gene (Reynaud et al. 1989).

#### Interlocus Versus Interallelic Exchange

The possible role of genetic exchange between alleles at the same locus should also be taken into account. Many of the exon 2 sequences presented here show a complicated pattern of polymorphism compatible with gene conversion or repeated crossing-over. This is the case with *DAB\*07* and *DAB\*08*, which are more similar to each other than to any other sequence, and they only differ in three hypervariable segments corresponding to residues 30–31, 37–38, and 84–86 in the mature protein (fig. 2). However, all of the sequence motifs by which they differ from each other except one have already been found in other pheasant alleles. The same pattern can be seen in *DAB1\*0101* and *DAB2\*0102*, which only differ from *DAB2\*04* in the hypervariable residues 9, 32, 37–38, and 84–86. All of the sequence motifs have been found in other combinations in different alleles. The lack of locus designation for most pheasant and chicken alleles makes it impossible to disentangle the relative roles of interallelic and interlocus genetic exchange. Interallelic genetic exchange may in fact be more common than interlocus exchange. Thus, Zangenberg et al. (1995) found in a PCR assay that 1/10,000 sperm possessed a new *HLA-DPB1* exon 2 sequence due to interallelic genetic exchange, i.e., one to two orders of magnitude greater than that reported for interlocus genetic exchange between two class II *B* loci in mouse sperm (Högstrand and Böhme 1994). In practice, it may be

more difficult to verify interallelic genetic exchange by sequence analysis of the patchwork pattern of polymorphism in exon 2 alone, whereas the 3'UT region provides a reliable marker in studies of interlocus genetic exchange.

#### *DAB* Locus Organization

The evidence that each *Mhc* class II *B* haplotype consists of only one *DAB1* locus and one *DAB2* locus is more consistent in ring-necked pheasants than in domestic chickens. Not only did both the *R<sub>2i</sub>* haplotype and the *R<sub>3a</sub>* haplotype consist of one copy of each locus, but RT-PCR from RNA and standard PCR from genomic DNA also gave two *DAB* sequences in three other haplotypes. This is in agreement with RFLP typing of the *DAB* genes of more than 500 Swedish pheasants, which also invariably indicated that there were two genes per haplotype (Wittzell et al. 1994, 1995; von Schantz et al. 1996; unpublished data). For the domestic chicken, the most compelling evidence of the same *Mhc* class II *B* organization as in pheasants comes from the genomic cloning of one *B-LBI* gene and one *B-LBII* gene in the *B<sup>6</sup>* haplotype (Xu et al. 1989). Guillemot et al. (1988) found the same result for the *B<sup>12</sup>* haplotype, but the 3'UT region of the so-called *B-LBI* gene has not been sequenced in this case (Zoorob et al. 1993). However, later, a third gene fragment was amplified from the *B<sup>12</sup>* haplotype using allele-specific PCR primers for the so-called *B-LBVI* family (Zoorob et al. 1993). Three sequences of *B-LBI/B-LBII*/"*B-LBVI*" were also found by PCR in the *B<sup>10</sup>*, *B<sup>14</sup>*, and *B<sup>24</sup>* haplotypes in the same study. The 3'UT region has never been cloned from a *B-LBVI* sequence, and it seems feasible from the phylogenetic analysis that the "*B-LBVI*" group simply consists of some closely related alleles at the *B-LBI* and *B-LBII* loci. Assuming that PCR-related errors can be excluded, the most likely explanation for the third sequence is that some of the investigated birds were heterozygous for the *Mhc* class II *B* region. Standard serological typing of domestic chickens is based on the so called *B-G* antigens and class I antigens on red blood cells and does not normally take into account the class II *B* genes (Briles et al. 1982; Håla et al. 1988). Presently, it is not possible to completely rule out the possibility that some chicken haplotypes do possess three *DAB* genes due to a more recent duplication. This would not change the conclusion that the duplication leading to *DAB1* (*B-LBI*) and *DAB2* (*B-LBII*) predates the radiation of pheasants and domestic chickens, but we note that molecular typing of *Mhc* homozygous strains of domestic chickens using denaturing gradient gel electrophoresis of exon 2 also invariably gave two sequences (M. Grahn, personal communication). The latter study, moreover, revealed two cosegregating *Mhc* class II *B* sequences per haplotype in families of red jungle fowls *Gallus gallus*, i.e., the wild ancestor of the domestic chicken. Therefore, it seems likely that domestic chickens normally possess only two *DAB* genes per haplotype, just as the ring-necked pheasant.

## Acknowledgments

We thank Dr. A. Bréfort and Laprovot S.C.P.V. for kindly providing the pheasants, Dr. A.-M. Chaussé for help in Tours, Dr. S. Edwards for constructive comments on the manuscript, and Dr. M. Grahn for communicating unpublished results. This study was supported by a post-doctoral fellowship to H.W. from the Swedish Research Council for Agriculture and Forestry, by Centre National de la Recherche Scientifique, and in part by grants from Ministère de la Recherche et de l'Enseignement Supérieur, Institut National de la Santé, et de la Recherche Médicale, and Association pour la Recherche sur le Cancer to C.A. and R.Z.

## LITERATURE CITED

- BELL, J. I., D. DENNEY JR., L. FOSTER, T. BELT, J. A. TODD, and H. O. McDEVITT. 1987. Allelic variation in the DR subregion of the human major histocompatibility complex. *Proc. Natl. Acad. Sci. USA* **84**:6234–6238.
- BOSS, J. M., and J. L. STROMINGER. 1984. Cloning and sequence analysis of the human major histocompatibility complex gene DC-3 $\beta$ . *Proc. Natl. Acad. Sci. USA* **81**:5199–5203.
- BRILES, W. E., N. BUMSTEAD, D. L. EWERT et al. (16 co-authors). 1982. Nomenclature for chicken major histocompatibility (*B*) complex. *Immunogenetics* **15**:441–447.
- BROWN, J. H., T. S. JARDETZKY, 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.
- BROWN, J. H., T. JARDETZKY, M. A. SAPER, B. SAMRAOUI, P. J. BJORKMAN, and D. C. WILEY. 1988. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature* **332**:845–850.
- BRUNSBURG, 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.
- 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., 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.
- GENETICS COMPUTER GROUP. 1994. Wisconsin sequence analysis package. Version 8. Genetics Computer Group, Madison, Wis.
- GUILLEMOT, F., A. BILLAULT, O. POURQUIÉ, G. BÉHAR, A.-M. CHAUSSÉ, R. ZOOROB, G. KREIBICH, and C. AUFRAY. 1988. A molecular map of the chicken major histocompatibility complex: the class II $\beta$  genes are closely linked to the class I genes and the nucleolar organizer. *EMBO J.* **7**:2775–2785.
- GYLLENSTEN, U. B., M. SUNDVALL, and H. A. ERLICH. 1991. Allelic diversity is generated by intraexon sequence exchange at the *DRB1* locus of primates. *Proc. Natl. Acad. Sci. USA* **88**:3686–3690.
- HÁLA, K., A.-M. CHAUSSÉ, Y. BOURLET, O. LASSILA, V. HASLER, and C. AUFRAY. 1988. Attempt to detect recombination between *B-F* and *B-L* genes within the chicken *B* complex by serological typing, in vitro MLR, and RFLP analyses. *Immunogenetics* **28**:433–438.
- HELM-BYCHOWSKI, K. M., and A. C. WILSON. 1986. Rates of nuclear DNA evolution in pheasant-like birds: evidence from restriction maps. *Proc. Natl. Acad. Sci. USA* **83**:688–692.
- HÖGSTRAND, K., and J. BÖHME. 1994. A determination of the frequency of gene conversion in unmanipulated mouse sperm. *Proc. Natl. Acad. Sci. USA* **91**:9921–9925.
- HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**:167–170.
- . 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc. Natl. Acad. Sci. USA* **86**:958–962.
- JARVI, S. I., R. M. GOTO, W. E. BRILES, and M. M. MILLER. 1996. Characterization of *Mhc* genes in a multigenerational family of ring-necked pheasants. *Immunogenetics* **43**:125–135.
- JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp 21–132 in H. N. MUNRO, ed. *Mammalian protein metabolism*. Academic Press, New York.
- JUUL-MADSEN, H. R., R. ZOOROB, C. AUFRAY, K. SKJØDT, and J. E. HEDEMAND. 1997. New chicken *Rfp-Y* haplotypes on the basis of *MHC* class II RFLP and MLC analyses. *Immunogenetics* **45**:345–352.
- KAPPES, D. J., D. ARNOT, K. OKADA, and J. L. STROMINGER. 1984. Structure and polymorphism of the HLA Class II SB light chain genes. *EMBO J.* **3**:2985–2993.
- KLEIN, J. 1986. *Natural history of the major histocompatibility complex*. John 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. 1995. Class II *B Mhc* motifs in an evolutionary perspective. *Immunol. Rev.* **143**:89–111.
- KROEMER, G., A. BERNOT, G. BÉHAR, A.-M. CHAUSSÉ, L.-N. GASTINEL, F. GUILLEMOT, I. PARK, P. THORAVAL, R. ZOOROB, and C. AUFRAY. 1990. Molecular genetics of the chicken MHC: current status and evolutionary aspects. *Immunol. Rev.* **113**:119–145.
- KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.01. Pennsylvania State University, University Park.
- LI, W.-H. 1997. *Molecular evolution*. Sinauer, Sunderland, Mass.
- LISKAY, R. M., A. LETSOU, and J. L. STACHELEK. 1987. Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics* **115**:161–167.
- MILLER, M. M., R. GOTO, A. BERNOT, R. ZOOROB, C. AUFRAY, N. BUMSTEAD, and W. E. 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. AUFRAY, 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* **96**:3958–3962.
- MOON SUNG, A., A. W. NORDSKOG, S. J. LAMONT, and C. M. WARNER. 1993. Isolation and characterization of cDNA clones for chicken major histocompatibility complex class II molecules. *Anim. Genet.* **24**:227–233.

- NEI, M., X. GU, and T. SITNIKOVA. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA* **94**:7799–7806.
- PARHAM, P., and T. OHTA. 1996. Population biology of antigen presentation by MHC class I molecules. *Science* **272**:67–74.
- PHARR, G. T., L. D. BACON, and J. B. DODGSON. 1993. Analysis of *B-L* $\beta$ -chain expression in two chicken cDNA libraries. *Immunogenetics* **37**:381–385.
- REYNAUD, C.-A., A. DAHAN, V. ANQUEZ, and J. C. WEILL. 1989. Somatic hyperconversion diversifies the single  $V_H$  gene of the chicken with a high incidence in the D region. *Cell* **59**:171–183.
- SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. *Molecular cloning: a laboratory manual*. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- SANGER, F., S. NICKLEN, and A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- SCHWAIGER, F.-W., E. WEYERS, C. EPPLER, J. BRÜN, G. RUFF, A. CRAWFORD, and J. T. EPPLER. 1993. The paradox of MHC-DRB exon/intron evolution:  $\alpha$ -helix and  $\beta$ -sheet encoding regions diverge while hypervariable intronic simple repeats coevolve with  $\beta$ -sheet codons. *J. Mol. Evol.* **37**:260–272.
- TROWSDALE, J. 1995. “Both man & bird & beast”: comparative organization of MHC genes. *Immunogenetics* **41**:1–17.
- VINCEK, V., D. KLEIN, R. T. GRASER, F. FIGUEROA, C. O’HUGIN, 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’HUGIN, Y. SATTI, 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 Biol. Sci.* **264**:111–118.
- VON SCHANTZ, T., H. WITZELL, G. GÖRANSSON, M. GRAHN, and K. PERSSON. 1996. MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proc. R. Soc. Lond. B Biol. Sci.* **263**:265–271.
- WITZELL, H., T. VON SCHANTZ, R. ZOOROB, and C. AUFRAY. 1994. Molecular characterization of three *Mhc* class II *B* haplotypes in the ring-necked pheasant. *Immunogenetics* **39**:395–403.
- . 1995. *Rfp-Y*-like sequences assort independently of pheasant *Mhc* genes. *Immunogenetics* **42**:68–71.
- XU, Y., J. PITCOVSKI, L. PETERSON, C. AUFRAY, Y. BOURLET, B. M. GERNDT, A. W. NORDSKOG, S. J. LAMONT, and C. M. WARNER. 1989. Isolation and characterization of three class II MHC genomic clones from the chicken. *J. Immunol.* **142**:2122–2132.
- ZANGENBERG, G., M.-M. HUANG, N. ARNHEIM, and H. ERLICH. 1995. New HLA-DPB1 alleles generated by interallelic gene conversion detected by analysis of sperm. *Nat. Genet.* **10**:407–414.
- ZOOROB, R., G. BÉHAR, G. KROEMER, and C. AUFRAY. 1990. Organization of a functional chicken class II *B* gene. *Immunogenetics* **31**:179–187.
- ZOOROB, R., A. BERNOT, D. MARTEL RENOIR, F. CHOUKRI, and C. AUFRAY. 1993. Chicken major histocompatibility complex class II *B* genes: analysis of interallelic and interlocus sequence variance. *Eur. J. Immunol.* **23**:1139–1145.

CRAIG MORITZ, reviewing editor

Accepted December 22, 1998