Journal of Heredity 2010:101(4):448–460
doi:10.1093/jhered/esq018
Advance Access publication March 3, 2010

Extensive MHC Class II B Gene Duplication in a Passerine, the Common Yellowthroat (*Geothlypis trichas*)

Jennifer L. Bollmer, Peter O. Dunn, Linda A. Whittingham, and Charles Wimpee

From the Department of Biological Sciences, University of Wisconsin—Milwaukee, PO Box 413, Milwaukee, WI 53201.

Address correspondence to Jennifer L. Bollmer at the address above, or e-mail: bollmer@uwm.edu.

Abstract

The major histocompatibility complex (MHC) is characterized by a birth and death model of evolution involving gene duplication, diversification, loss of function, and deletion. As a result, gene number varies across taxa. Birds have between one and 7 confirmed MHC class II B genes, and the greatest diversity appears to occur in passerines. We used multiple primer sets on both genomic DNA (gDNA) and complementary DNA (cDNA) to characterize the range of class II B genes present in a passerine, the common yellowthroat (*Geothlypis trichas*). We confirmed 39 exon 2 sequences from gDNA in a single individual, indicating the presence of at least 20 class II B loci. From a second individual, we recovered 16 cDNA sequences belonging to at least 8 transcribed loci. Phylogenetic analysis showed that common yellowthroat sequences fell into subgroups consisting of classical loci, as well as at least 4 different clusters of sequences with reduced sequence variability that may represent pseudogenes or nonclassical loci. Data from 2 additional common yellowthroats demonstrated high interindividual variability. Our results reveal that some passerines possess an extraordinary diversity of MHC gene duplications, including both classical and nonclassical loci.

Key words: *balancing selection, common yellowthroat, gene duplication,* Geothlypis trichas, *major histocompatibility complex, recombination*

The major histocompatibility complex (MHC) consists of genes that play a central role in vertebrate immunity, mainly by recognizing foreign peptides and presenting them to T cells, thus initiating the adaptive immune response (Klein 1986). MHC class I genes present primarily endogenous self antigens and antigens originating from intracellular pathogens such as viruses to CD8-bearing T cells, whereas class II genes present exogenous antigens to CD4-bearing T cells. Class I and II genes can be further categorized as classical (high polymorphism and expression), nonclassical (low polymorphism with limited expression), or nonfunctional pseudogenes (Klein and Figueroa 1986). Classical MHC genes have the highest levels of polymorphism known for coding genes (Gaudieri et al. 2000; Robinson et al. 2003). This variability is thought to be produced primarily by recombination (Richman et al. 2003; Reusch and Langefors 2005; Schaschl et al. 2006) and maintained through balancing selection caused by coevolution of hosts and parasites and sexual selection (Doherty and Zinkernagel 1975; Penn and Potts 1999).

MHC gene number varies across taxa, including between closely related species and between haplotypes within a species (Trowsdale 1995; Málaga-Trillo et al. 1998; Kelley et al. 2005; Bryja et al. 2006). The MHC appears to follow a birth-and-death model of evolution, in which new genes arise via gene duplication and then are maintained for long periods, become nonfunctional pseudogenes or are deleted (Nei et al. 1997; Nei and Rooney 2005). Within individuals, a higher number of loci and alleles may be favored if there is a heterozygote advantage; however, evidence for this is mixed (McClelland et al. 2003; Ilmonen et al. 2007; Oliver et al. 2009), and some studies have found that an intermediate number of MHC alleles is optimal (Wegner et al. 2003; Bonneaud, Mazuc, et al. 2004; Kalbe et al. 2009). Woelfing et al. (2009) predicted that the optimal number of MHC molecules (class I and II combined) per individual should be between 3 and 25, the same order of magnitude typically observed in individuals, because higher variability may lead to depletion of the T-cell repertoire (Nowak et al. 1992; Woelfing et al. 2009). The number of expressed MHC class II B genes varies from one to 3 in amphibians (Kobari et al. 1995; Babik et al. 2008; Zeisset and Beebee 2009), one to at least 6 in fish (e.g., Málaga-Trillo et al. 1998; Aguilar and Garza 2007), and at least 4 in a reptile (Miller et al. 2005). Eutherian mammals share conserved class II gene regions of classical (DR, DQ, and DP) and nonclassical

Table 1 Number of MHC class II B loci described in various avian species

Order	Species	Total number of loci	Transcribed loci	Pseudogenes or nonclassical loci	Source
Passeriformes	Common yellowthroat (*Geothlypis trichas*)	20 (31)*	8 (17)	9 (16)	This study
	House sparrow (*Passer domesticus*)	3*	3		Bonneaud, Sorci, et al. (2004)
	Little greenbul (*Andropadus virens*)	8*	7	1	Aguilar et al. (2006)
	Scrub jay (*Aphelocoma coerulescens*)	3*	3		Edwards et al. (1995)
	Great reed warbler (*Acrocephalus arundinaceus*)	4*	4		Westerdahl et al. (2000)
	New Zealand robin (*Petroica australis australis*)	7*	4		Miller and Lambert (2004a, 2004b)
	I′iwi honeycreeper (*Vestiaria coccinea*)	3*		1	Jarvi et al. (2004)
	Small cactus finch (*Geospiza scandens*)	6*		1	Sato et al. (2000), (2001)
	Red-winged blackbird (*Agelaius phoeniceus*)	3		2	Edwards et al. (1998, 2000); Gasper et al. (2001)
Charadriiformes	Great snipe (*Gallinago media*)	3*			Ekblom et al. (2003)
	Least (*Aethia pusilla*), crested (*A. cristatella*) auklets	2*			Walsh and Friesen (2003)
Galliformes	Ring-necked pheasant (*Phasianus colchicus*)	2	2		Wittzell et al. (1999)
	Black grouse (*Tetrao tetrix*)	4*	2	2	Strand et al. (2007)
	Chicken (*Gallus gallus*)	5	5	3	Miller et al. (2004); references therein
	Japanese quail (*Coturnix japonica*)	7	6		Shiina et al. (2004)
	Turkey (*Meleagris gallopavo*)	3			Chaves et al. (2009)
	Peafowl (*Pavo cristatus*)	3*			Hale et al. (2009)
Sphenisciformes	Penguins (9 species)	1	1		Tsuda et al. (2001); Bollmer et al. (2007); Kikkawa et al. (2009)
Ciconiiformes	Crested ibis (*Nipponia nippon*)	1*			Zhang et al. (2006)
Falconiformes	Lesser kestrel (*Falco naumanni*)	1*			Alcaide et al. (2008)
Strigiformes	Barn owl (*Tyto alba*)	2	2		Burri, Niculita-Hirzel, Roulin, and Fumagalli (2008)
Psittaciformes	Parrotlet (*Forpus passerinus*)	1			Hughes et al. (2008)

The total number of loci are given, as are the number of those loci that are transcribed (if cDNA had been analyzed) or are possible pseudogenes or nonclassical loci. For the common yellowthroat, the estimated number of loci based on confirmed sequences is given, as is the number of loci when both confirmed and unconfirmed were considered (in parentheses). The MHC regions have been mapped in only a subset of these species, so most estimates are minima. In many cases (indicated by *), the number of loci was estimated by ourselves or other authors based on the maximum number of sequences recovered from a single individual with the assumption that individuals are heterozygous at each locus.

(DM and DO) loci, each having one or more expressed B genes in most taxa (Hughes and Nei 1990; Takahashi et al. 2000; Kumánovics et al. 2003; Yuhki et al. 2003).

In birds, the number of class II B loci identified per species varies from one to at least 7 confirmed and possibly 8 (Table 1). The best mapped bird species is the domestic chicken (*Gallus gallus*), which has a minimal MHC (few genes with short introns) compared with mammals (Kaufman et al. 1999; Shiina et al. 2007; Hosomichi et al. 2008). Two class II genes occur within the polymorphic chicken B complex, although at least 3 less variable class II genes of unknown function are present at the Y complex (Miller et al. 2004). Some other species in Galliformes appear to have similarly low numbers of class II loci (Wittzell et al. 1999; Strand et al. 2007), as do raptors (Alcaide et al. 2007; Burri, Niculita-Hirzel, Salamin, et al. 2008), penguins (Tsuda et al. 2001; Bollmer et al. 2007), and parrotlets (Hughes et al. 2008) among others (Table 1). Studies of passerine MHC have generally revealed more genes than nonpasserines (Westerdahl 2007; Table 1). Based on sequences amplified from complementary DNA (cDNA), at least 3–7 transcribed loci may be present per species in passerines, whereas only 1–3 transcribed loci appear in nonpasserines (Table 1). The presence of MHC pseudogenes and nonclassical loci with low variability also appears to be more common in passerines (Westerdahl 2007), although the Y complex in Galliformes has similar nonclassical loci. Single probable pseudogenes have been identified in greenbuls (*Andropadus virens*; Aguilar et al. 2006) and house finches (*Carpodacus mexicanus*; Hess et al. 2000), and 2 loci (one probable pseudogene and one probable nonclassical gene) have been identified in red-winged blackbirds (*Agelaius phoeniceus*; Edwards et al. 2000; Gasper et al. 2001). It is unclear whether these types of genes are duplicated within species as are the classical loci.

Here, we employed multiple primer sets on both genomic DNA (gDNA) and cDNA to assess the diversity of MHC class II B loci present in a widespread North American warbler, the common yellowthroat (*Geothlypis trichas*). This is a promising species for further investigation

of selection on the MHC because previous research has shown that female mate choice is based on male plumage ornamentation (Tarof et al. 2005; Dunn et al. 2008), and male ornament size signals greater T-cell mediated and humoral immunity (Garvin et al. 2008; Dunn et al. 2010), which may be associated with MHC variation. Due to the close relationships among duplicated MHC loci within birds, primers often amplify multiple loci at once (Hess and Edwards 2002; Westerdahl et al. 2004). We took advantage of this to estimate the number of loci present in the common yellowthroat. We amplified sequences from gDNA to screen for classical loci as well as possible low variability (nonclassical) loci or pseudogenes, and we amplified sequences from cDNA in order to estimate the number of loci that were actually transcribed. Our study targeted exon 2 because it codes for the peptide-binding region of the molecule and should be under selection. We show that the number of class II B genes in this species appears to be much greater than that found in previous studies of birds.

Materials and Methods

Amplification of MHC from gDNA

For our initial investigation of MHC allelic variation, we used DNA from 3 unrelated male common yellowthroats. Samples were collected at the University of Wisconsin—Milwaukee Field Station in Saukville, WI (lat 43°23′N, long 88°01′W) from individuals that were part of a long-term study of mating behavior in this species (Garvin et al. 2006; Dunn et al. 2008).

First, we amplified sequences from a single individual (WNR) using 2 previously published primer sets: 1) 326/325 (Ekblom et al. 2003), a degenerate primer set which amplified a 210-bp fragment (primers included) within exon 2 and 2) Int1f.7/Int2r.1 (Edwards et al. 1998; Aguilar et al. 2006), which amplified part of intron 1, all of exon 2, and a small portion of intron 2 (525–575 bp including primers). We amplified 326/325 in 50 μl reactions: 1× polymerase chain reaction (PCR) buffer, 1 mM dNTPs, 2.5 mM $MgCl_2$, 1 μM each primer, 1.2% DMSO, and 1.5 U GoTaq Flexi DNA polymerase (Promega, Madison, WI). Reaction conditions were as follows: 94 °C for 2 min; then 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s; and a final extension of 72 °C for 7 min. We also amplified Int1f.7/Int2r.1 in a 50 μl reaction: 1× PCR buffer, 1 mM dNTPs, 2.5 mM $MgCl_2$, 1 μM each primer, and 1.5 U GoTaq. Reaction conditions included 94 °C for 3 min; then 30 cycles of 94 °C for 1 min, 54 °C for 1 min, and 72 °C for 1 min; and a final extension of 72 °C for 5 min. PCR products were blunt ended by treating them with *Pfu*, purified using QIAquick kits (QIAGEN, Valencia, CA), and then cloned using the CloneSmart Blunt Cloning Kit (Lucigen, Middleton, WI). Colonies with inserts of the correct size were sequenced at the University of Chicago Cancer Research Center DNA Sequencing Facility using the vector primers (SL1 and SR2). WNR underwent 2 (326/325) and 3 (Int1f.7/Int2r.1) independent PCR, cloning, and sequencing rounds with these primer sets.

We designed a new forward primer (GetrInt1f; 5′-GCCCTGAGCTGTGTGCTG-3′) because the intron primer set (Int1f.7/Int2r.1) amplified sequences that we suspected belonged to either pseudogenes or nonclassical genes, and we wanted to target likely functional alleles. The GetrInt1f primer was positioned in intron 1 (ending 9 bp upstream of exon 2), and, in combination with Int2r.1, amplified a 327 bp fragment (primers included) encompassing all of exon 2. Reaction conditions for GetrInt1f were the same as for Int1f.7/Int2r.1. Using the GetrInt1f primer, we cloned sequences from the original individual (WNR), as well as 2 other individuals (ERM and GRG). Each individual underwent 3 independent PCR, cloning, and sequencing rounds with the GetrInt1f/Int2r.1 primer set.

Amplification of MHC from cDNA

To amplify transcribed sequences, total RNA was isolated from liver tissue of a common yellowthroat collected in Baton Rouge, Louisiana in January 2009. Liver tissue was stored in RNAlater (QIAGEN) in the field, and RNA was extracted using the RNeasy Mini Kit (QIAGEN) with a DNase treatment. We then used the OneStep RT-PCR Kit (QIAGEN) for both reverse transcription and amplification of our target sequence, following the manufacturer's protocol with an annealing temperature of 57 °C. In order to amplify exon 2, we used primers placed in exons 1 and 3: Song EX1F.2 and Song EX3R.1 (Aguilar et al. 2006). The PCR product was then cloned and sequenced following the methods outlined above. The extracted RNA underwent 3 independent rounds of reverse transcription PCR (RT-PCR) and cloning.

RFLP and Southern Blot Analysis

To obtain an estimate of how many class II B genes might be present in the common yellowthroat, we performed a restriction fragment length polymorphism (RFLP) analysis. We selected 4 unrelated adults from the Wisconsin population and digested 10 μg of gDNA for 3 h with 20 units of *Pvu*II, an enzyme that has been used successfully in other avian MHC studies (e.g., Ekblom et al. 2003; Bonneaud, Sorci, et al. 2004; Miller and Lambert 2004a). The digested DNA was run on a 0.8% agarose gel for 19 h at 60 V in 0.5× Tris-Borate-EDTA buffer and then transferred by Southern blotting to a Hybond (Amersham) membrane following manufacturer's instructions. The blot was prehybridized at room temperature overnight in 50% formamide, 5× standard saline citrate (SSC), 10× Denhardt's solution, and 250 μg/ml yeast RNA. The blot was then hybridized at 42 °C overnight with ^{32}P-labeled *Getr**15 sequence as a probe. The hybridization buffer was the same as the prehybridization, with the addition of dextran sulfate to a concentration of 10%. The blot was washed in 0.5× SSC, 0.1% SDS by rinsing briefly at room temperature, followed by two 30 min washes at 65 °C with a buffer change between each wash. Autoradiography was carried

out for 24 h on Kodak BioMax film at −80 °C using an intensifying screen.

Data Analysis

Spurious sequences can form during PCR due to *Taq* error or the formation of heteroduplexes or chimeras, especially when amplifying multiple sequences in one reaction (Jansen and Ledley 1990; L'Abbé et al. 1992). We considered sequences to be confirmed if they were amplified in at least 2 independent reactions, either from 2 different PCR/cloning rounds from the same individual or from 2 different individuals. Due to the apparently large number of sequences present in each bird (>39 in WNR), it was often difficult to get duplicates of individual sequences in independent reactions, so we have reported unconfirmed sequences as well and labeled them as such. Other sequences that were less than 3 bp different from either confirmed or unconfirmed sequences were discarded as probable PCR error as were sequences that appeared to be chimeras of 2 other sequences. All confirmed sequences were submitted to GenBank (accession numbers GQ247563–GQ247638), and both confirmed and unconfirmed sequences are available online as Supplementary material.

We aligned the sequences manually in BioEdit v. 7.0.9.0 (Hall 1999). To determine relationships among common yellowthroat sequences and to identify subsets of sequences, including possible pseudogenes or nonclassical genes, we constructed a phylogenetic network using the program SplitsTree4 (Huson 1998; Huson and Bryant 2006). We employed the Neighbor-Net method (Bryant and Moulton 2004) using Jukes–Cantor distances. The advantage of phylogenetic networks is that they permit the representation of conflicting signal that arises when processes such as gene duplication and recombination have made evolutionary relationships among sequences more complex than can be demonstrated by a traditional phylogenetic tree (Bryant and Moulton 2004). We used confirmed exon 2 sequences to do this. To assess the relationships between exon 2 sequences of common yellowthroat and other species, we constructed a maximum likelihood (ML) tree using the HKY85 nucleotide substitution model with 500 bootstrap replicates in the program PhyML v. 3.0 (Guindon and Gascuel 2003). We then calculated exon 2 sequence diversity measures in the program DnaSP v. 4.5 (Rozas et al. 2003).

We tested for the presence of recombination (or gene conversion) using the program GENECONV v. 1.81 (Sawyer 1999). GENECONV employs a substitution method that has a high probability of identifying recombination when it is present (with a low rate of false positives) and is able to analyze homologous sequences from the same locus or multiple loci (Posada 2002). The program compares sequences in a pairwise fashion and searches for stretches of sequence (fragments) that are more similar than would be expected by chance. Each fragment is compared with all possible fragments in the alignment in a permutation process, and global P values (corrected for multiple comparisons) are calculated. We ran 10 000 permutations and allowed zero mismatches. P values <0.05 were considered evidence of recombination.

We tested for historical evidence of balancing selection on the peptide-binding region by calculating nonsynonymous (d_N) and synonymous (d_S) substitution rates. A d_N/d_S ratio of $\omega = 1$ is expected under neutral evolution, $\omega < 1$ under purifying selection, and $\omega > 1$ under positive selection. First, we calculated the substitution rates using the Nei and Gojobori (1986) method with the Jukes–Cantor correction in MEGA v. 4.1 (Tamura et al. 2007). Rates were calculated separately for both hypothesized peptide-binding and non–peptide-binding codons as determined by Brown et al. (1993) for human class II molecules. We tested for positive selection using a Z-test (Nei and Kumar 2000). Second, we tested for positive selection using the ML method implemented in CODEML in the package PAML v. 4 (Yang et al. 2000, 2005). The advantage to this method is that we do not make an a priori assumption about which codons may be peptide binding. To test for selection, we used likelihood ratio tests of neutral models and models with selection. In particular we compared: 1) neutral model M1a ($\omega_0 < 1$, $\omega_1 = 1$) with positive selection model M2a ($\omega_2 > 1$) and 2) neutral model M7 ($0 < \omega < 1$) with positive selection model M8 ($0 < \omega < 1$, $\omega > 1$). This second model comparison (M7 and M8) is most robust to the effect of recombination, which may cause false positives (Anisimova et al. 2003). Positively selected codons with a $\omega > 1$ were identified using the Bayes empirical Bayes approach (Yang et al. 2005).

Results

Number of gDNA and cDNA Exon 2 Sequences

We recovered a large number of gDNA class II exon 2 sequences per individual. From one individual (WNR), we sequenced 42–95 colonies containing MHC inserts for each of the primer sets (Table 2). From these colonies, we were able to confirm a total of 39 different sequences, indicating a minimum of 20 class II B loci. There were an additional 22 unconfirmed sequences from WNR. From the 2 additional individuals (GRG and ERM) screened using only the GetrInt1f/Int2r.1 primer set, we confirmed 12 and 14 sequences from 88 and 84 colonies, respectively. These individuals also appeared to have a large number of unconfirmed sequences. Overall, we confirmed a total of 62 sequences from these 3 individuals, with only 3 of these sequences confirmed in 2 different individuals, indicating high interindividual allelic variability (Supplementary Figure S1).

We also found a large number of transcribed loci, with at least 8 confirmed. From a single common yellowthroat individual, we sequenced a total of 155 clones from 3 independent RT-PCR/cloning rounds and confirmed 16 sequences (Table 2). One confirmed cDNA sequence matched a previously unconfirmed sequence (*Getr**62) recovered from WNR and another matched a confirmed

Table 2 Number of MHC class II B exon 2 sequences amplified from gDNA and cDNA of common yellowthroats

		326/325	Int1f/Int2r	GetrInt1f/Int2r	Total	cDNA
Number of colonies sequenced		42	84	95	221	155
Confirmed	Noncluster	10	3	18	22	16
	Cluster 1		5		5	
	Cluster 2	5		4	9	
	Cluster 3	2	1		3	
Total confirmed		17	9	22	39	16
Unconfirmed	Noncluster		3	5	8	14
	Cluster 1	1	11		12	1
	Cluster 2			2	2	
	Cluster 3					3
Total unconfirmed		1	14	7	22	18
Total number of sequences		18	23	29	61	34

The gDNA sequences were amplified from a single individual (WNR) using 3 different primer sets (326/325, Int1f/Int2r, and GetrInt1f/Int2r), and the "Total" column reports the total number of sequences across primer sets (after removing redundant sequences) amplified from that individual. The cDNA sequences were amplified from a second individual. Numbers of confirmed and unconfirmed sequences are presented separately and are further classified into subgroups based on sequence similarity.

sequence (*Getr**34) from ERM. The 16 confirmed cDNA sequences indicated at least 8 loci were transcribed; however, the number may be much higher as we recovered an additional 18 unconfirmed sequences (Table 2).

RFLP Analysis

The RFLP analysis provided further evidence that the common yellowthroat has a large number of class II B genes. Each of the 4 individuals had between 19 and 27 bands in the range of 0.5–3.5 kb (Supplementary Figure S2). Also, all 4 individuals had different genotypes.

Phylogenetic Relationships

A phylogenetic network of confirmed gDNA sequences from WNR and confirmed cDNA sequences from a second individual showed a pattern of structuring, possibly corresponding to different groups of loci (Supplementary Figure S3, Figure 1). A subset of sequences fell into clusters divergent from the rest of the sequences but with low within-cluster sequence divergence. Different primer sets amplified different clusters: Int1f.7/Int2r.1 amplified Cluster 1, 326/325 and GetrInt1f/Int2r.1 amplified Cluster 2, and 326/325 and Int1f.7/Int2r.1 amplified Cluster 3. More than 2 gDNA sequences fell into each cluster, indicating that the cluster sequences originated from multiple loci. Two sequences (*Getr**01 and 08, both amplified by 326/325) were especially divergent from the rest and may represent separate loci. The remaining sequences were amplified by one or more of the 3 primer sets. A neighbor joining (NJ) tree generated using amino acid sequences produced the same pattern of clustering as the DNA sequences (Bollmer JL, Dunn PO, Whittingham LA, unpublished data). Because introns are not under selection, they may give more accurate information about locus identity. In a separate NJ tree, the corresponding intron 1 sequences supported these subgroups, further evidence that the cluster sequences are from divergent groups of loci (Supplementary Figure S4). All the confirmed cDNA sequences were interspersed among the noncluster gDNA sequences (Supplementary Figure S3). Similar to the confirmed gDNA sequences, the unconfirmed sequences from WNR also consisted of noncluster and cluster sequences; whereas 4 of the unconfirmed cDNA sequences fell into Clusters 1 and 3 (Table 2; Supplementary Figure S5). Although none of the Cluster 3 cDNA sequences were confirmed, different Cluster 3 sequences were recovered from independent cloning rounds: 2 sequences from the first cloning round and one from the third.

Some common yellowthroat sequences intermixed with sequences from other passerine species (Figure 1). Four of the yellowthroat sequences fell into well-supported clusters with sequences from other passerines. The 3 common yellowthroat Cluster 3 sequences (*Getr**05, 10, 18) formed a cluster with 2 Savannah sparrow (*Passerculus sandwichensis*) sequences. Also, common yellowthroat sequence *Getr**01 (previously one of the divergent, noncluster sequences) fell into a new cluster (Cluster 4) with sequences from 3 other passerine species: the Savannah sparrow, red-winged blackbird (*A. phoeniceus*), and common Amakihi (*Hemignathus virens*), a Hawaiian honeycreeper. In contrast, no sequences from other species fell into the common yellowthroat Clusters 1 and 2, and no common yellowthroat sequences were closely related to putative pseudogenes or nonclassical genes previously identified in other passerine species (e.g., Came-DAB1, Anvi-DAB1, Agph-DAB2, and Agph-DAB3; Figure 1).

Sequence Variation, Recombination, and Selection

Sequences in Clusters 1, 2, and 3 were less variable than the remaining, nonclustered and transcribed sequences (Table 3). Using only confirmed sequences, the number of polymorphic sites within each of the clusters was $\leq$18, whereas the nonclustered and transcribed sequences had 110 and 69 variable sites, respectively (within the 160-bp fragment where sequences from all 3 primer sets overlap). Nucleotide diversity and average intersequence divergence were correspondingly low within the 3 clusters. The 76 confirmed exon 2 sequences translated into 74 unique amino acid sequences (Supplementary Figure S1). Among the gDNA sequences, 10 had a single codon deletion, one (*Getr**61) had a single codon insertion, and 2 (*Getr**12 and 20) had a 4 bp deletion resulting in a stop codon (Supplementary Figure S1). No other sequences had stop codons. Among the confirmed cDNA sequences, *Getr**68 had a single codon insertion, and none had deletions.

The program GENECONV found evidence of putative recombination events among confirmed common yellowthroat sequences. This analysis was restricted to the 65 sequences derived from the intron primer sets that spanned the entire exon 2. It identified 4 possible fragments (60–89 bp

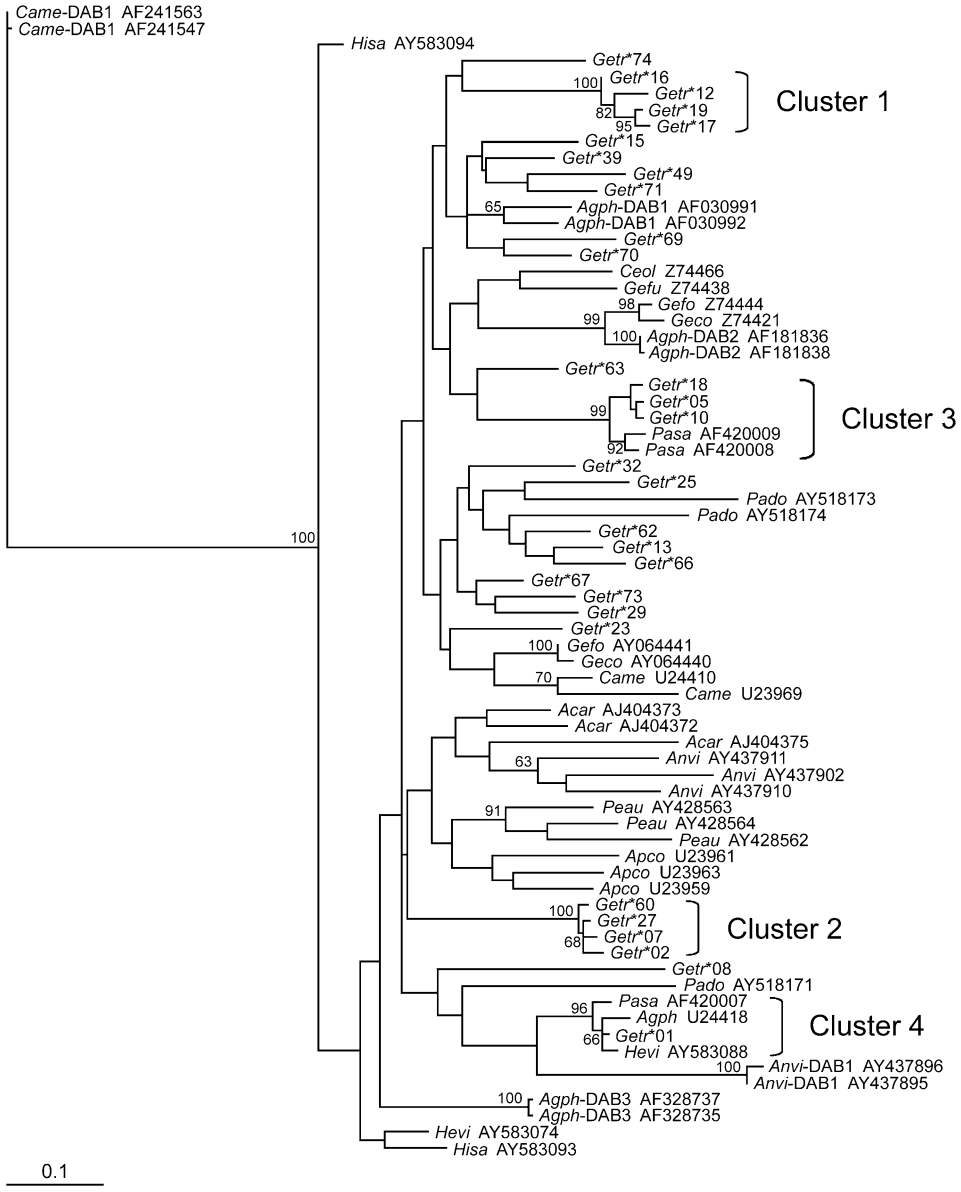

Figure 1. ML tree of MHC class II B exon 2 sequences (based on 133 bp) from common yellowthroats and other passerine species. Bootstrap values above 60 are shown. A subset of yellowthroat sequences (Getr, *Geothlypis trichas*) were used, including a representative sample of sequences from Clusters 1, 2, and 3. (Acar, *Acrocephalus arundinaceus*; Agph, *Agelaius phoeniceus*; Anvi, *Andropadus virens*; Apco, *Aphelocoma coerulescens*; Came, *Carpodacus mexicanus*; Ceol, *Certhidea olivacea*; Geco, *Geospiza conirostris*; Gefo, *Geospiza fortis*; Gefu, *Geospiza fuliginosa*; Hevi, *Hemignathus virens*; Hisa, *Himatione sanguinea*; Pado, *Passer domesticus*; Pasa, *Passerculus sandwichensis*; Peau, *Petroica australis*).

long) that were likely produced by recombination. (Table 4). Seven different sequences were involved in these 4 fragments, and all were noncluster sequences. We found no evidence for recombination between noncluster and cluster sequences, between sequences from different clusters, or between sequences within clusters. The high sequence similarities within clusters, however, make the detection of conversion events within clusters unlikely.

Analysis of nucleotide substitution rates provided evidence of positive selection on exon 2. Using MEGA, there was evidence of positive selection on the peptide-binding codons of the confirmed noncluster and transcribed sequences but not on Cluster 1 and 3 sequences (Table 5). The noncluster gDNA and transcribed sequences had similarly high nonsynonymous mutation rates, whereas both d_N and d_S were much lower among the sequences in Clusters 1, 2, and 3. Cluster 2 sequences did show a significant excess of nonsynonymous mutations, but this result is questionable because of the low mutation rates. We found no evidence of selection acting on the non–peptide-binding codons for any

Table 3 Sequence diversity at common yellowthroat MHC class II B loci

Sequence set	Number of sequences	Number of polymorphic sites	π	Number of bp differences between sequences (average ± standard error)
All gDNA sequences	62	121	0.184	29.4 ± 2.6
Cluster 1	5	16	0.050	8.0 ± 1.9
Cluster 2	12	17	0.035	3.9 ± 1.0
Cluster 3	3	4	0.017	2.7 ± 1.3
Noncluster (gDNA)	42	109	0.172	27.5 ± 2.5
Transcribed	16	68	0.152	24.3 ± 2.7

The number of sequences included for each analysis, number of polymorphic sites, nucleotide diversity (π), and average distance between sequences are reported. Only confirmed sequences were used, and the analyses were limited to the 160 bp for which all primer sets overlapped.

of the sequence sets. The above analyses were limited to confirmed sequences only, and the Cluster 1 sequences with frameshift mutations (*Getr**12 and 20) were excluded. There was also evidence of positive selection using the ML method implemented in CODEML. We did separate analyses for all the confirmed sequences together, the noncluster sequences only, and cluster sequences only. These analyses were restricted to the 65 sequences derived from the intron primer sets that spanned the entire exon 2. Both selection models (M2a and M8) provided a better fit than their respective neutral models (M1a and M7) when we analyzed all sequences, as well as just the noncluster sequences ($P < 0.001$; Supplementary Table S1). When we analyzed just the cluster sequences, M2a was not significantly better than M1a ($P > 0.10$), but M8 was a better fit than M7 ($P < 0.05$). The M2a and M8 models identified 13 and 15 sites, respectively as being under positive selection when all sequences were analyzed and 22 sites each for noncluster sequences only (Supplementary Table S1). Twelve of the 22 sites identified by CODEML as being under positive selection were designated by Brown et al. (1993) as being peptide binding (Supplementary Figure 1). Analyses of the cluster sequences did not identify any sites with posterior probabilities >0.99.

Table 4 Putative recombination events between common yellowthroat MHC class II B sequences

Seq 1	Seq 2	Sim *P*	Begin	End	Length
*Getr**64	*Getr**68	0.023	100	159	60
*Getr**37	*Getr**39	0.032	188	276	89
*Getr**34	*Getr**76	0.040	104	167	64
*Getr**65	*Getr**68	0.044	100	167	68

All fragments are globally significant. The 2 sequences involved, the simulated *P* value, beginning position, ending position, and length (bp) of each fragment are given.

Discussion

The MHC is characterized by gene duplication and deletion, resulting in variation in gene number between even closely related taxa. Passerines are thought to have a greater number of MHC loci than other bird taxa; however, the data on gene number in passerines are more ambiguous than for other taxa. This study provides the best evidence to date of a large number of gene duplications in a passerine. We recovered 39 confirmed exon 2 sequences from a single individual, which is evidence of at least 20 class II B loci in the common yellowthroat. Furthermore, we identified another 22 sequences that were unconfirmed, which, if confirmed by replication, indicates this individual could have a minimum of 31 loci. As mentioned earlier, MHC gene number can vary among haplotypes within a species, so other individuals with different haplotypes could have fewer or more genes. These loci included both those having characteristics typical of classical loci, as well as 3 low variability clusters that may represent either nonclassical loci or pseudogenes, although that remains uncertain. We also recovered a large number of transcribed alleles from a different individual. We were able to confirm 16 sequences but amplified 18 additional unconfirmed sequences, so a minimum of 8 (but as many as 17 if additional sequences are confirmed) loci are transcribed. This is likely an underestimate as gene expression varies across tissues, is low for some genes, and depends on disease state (e.g., Shiina et al. 2004; Hughes 2005). An RFLP analysis revealed a large number of bands, corroborating the high number of sequences recovered from cloning. A few common yellowthroat sequences fell into well-supported clusters with sequences from other passerine species (Figure 1), suggesting the presence of loci that predate the divergences of these species. We also found evidence of selection on the peptide-binding codons of the noncluster and transcribed sequences; in contrast, 2 clusters did not differ from neutrality. Overall, we confirmed a total of 76 sequences from 4 yellowthroat individuals, which is an indication of high variability both within and between individuals.

Number of Loci

Across bird species, MHC class II B gene number varies from one to at least 7 confirmed loci (Table 1). The MHC regions of a few bird species have been mapped to varying degrees, providing the most accurate estimates of overall gene number. Real and substantial differences in gene number do exist: the 2 classical class II B genes mapped in the domestic chicken correspond to 7 loci in the Japanese quail (Shiina et al. 2004). The domestic turkey has 3 class II B loci mapped (Chaves et al. 2009), and there is strong evidence that owls have only 2 loci (Burri, Niculita-Hirzel, Salamin, et al. 2008; Burri, Niculita-Hirzel, Roulin, and Fumagalli 2008) and a parrotlet has only one (Hughes et al. 2008). Some passerines appear to have a larger number of class II B loci than other taxa; however, the data are more ambiguous. A few genes have been characterized in 3 species (Edwards et al. 1998, 2000; Hess et al. 2000; Gasper et al. 2001; Aguilar et al. 2006), but most information

Table 5 Comparison of rates of nonsynonymous (d_N) and synonymous (d_S) substitutions calculated separately for peptide-binding codons and non–peptide-binding codons of common yellowthroat MHC class II B sequences

		Peptide-binding codons			Non–peptide-binding codons		
Sequence set	Number of sequences	$d_N \pm$ SE	$d_S \pm$ SE	d_N/d_S	$d_N \pm$ SE	$d_S \pm$ SE	d_N/d_S
All gDNA sequences	60	0.608 ± 0.120	0.211 ± 0.111	2.88[a]	0.135 ± 0.034	0.137 ± 0.034	0.99
Cluster 1	3	0.000 ± 0.000	0.053 ± 0.066	0.00	0.045 ± 0.017	0.058 ± 0.041	0.78
Cluster 2	12	0.034 ± 0.018	0.000 ± 0.000	—[a]	0.024 ± 0.008	0.024 ± 0.021	1.00
Cluster 3	3	0.000 ± 0.000	0.000 ± 0.000	0.00	0.022 ± 0.013	0.027 ± 0.027	0.81
Noncluster (gDNA)	42	0.553 ± 0.110	0.132 ± 0.092	4.19[a]	0.134 ± 0.033	0.129 ± 0.031	1.04
Transcribed	16	0.534 ± 0.125	0.180 ± 0.079	2.97[a]	0.105 ± 0.030	0.082 ± 0.032	1.28

Substitution rates were calculated separately for the different subsets of sequences. Fourteen peptide-binding and 39 non–peptide-binding codons were analyzed. SE, standard error.

[a] Z-test for positive selection (H_A: $d_N > d_S$) was significant ($P < 0.05$).

about locus number has to be inferred from the number of sequences recovered from single individuals or from RFLP analyses. For example, Miller and Lambert (2004b) recovered up to 14 exon 2 sequences per individual from New Zealand robin (*Petroica australis*) gDNA, and Aguilar et al. (2006) recovered 14 unique sequences from the cDNA of a little greenbul, suggesting that these species may have a minimum of 7 loci (8 in the greenbul when a pseudogene is included). RFLP analyses also provide a rough estimate of gene number, and passerines have had up to 30 bands per individual, indicating the presence of a large number of class II genes in some species (reviewed in Westerdahl 2007).

Based on RFLP and cloning data, the common yellowthroat has a high number of class II B genes compared with other bird species studied to date. The 4 common yellowthroats had between 19 and 27 RFLP bands per individual, a high number similar to the willow warbler (*Phylloscopus trochilus*) and Savannah sparrow (*P. sandwichensis*), which both had more than 20 bands per individual (Westerdahl et al. 2000; Freeman-Gallant et al. 2002). In contrast, great reed warblers (*Acrocephalus arundinaceus*) had 13–17 RFLP bands (Westerdahl et al. 2000), house sparrows (*Passer domesticus*) had 6–12 (Bonneaud, Sorci, et al. 2004), and Florida scrub jays (*Aphelocoma coerulescens*) had only 2–4 (Edwards et al. 1995). The cloning data suggested a similarly high number of loci compared with other species. The confirmation of 39 sequences from a single individual implies that the common yellowthroat has at least 20 class II B loci, and, thus, evidence of extensive gene duplication. Pseudogene or nonclassical sequences (i.e., low variability) have been recovered from other passerines (e.g., Edwards et al. 2000; Hess et al. 2000; Jarvi et al. 2004; Aguilar et al. 2006) but not in the high numbers that seem to be present in the common yellowthroat. Two different low variability loci were recovered from the red-winged blackbird (Edwards et al. 2000; Gasper et al. 2001), whereas in the common yellowthroat, we recovered confirmed sequences that fell into 4 different low variability clusters, 3 of which contained more than 2 sequences, indicating they originated from multiple loci: Cluster 1 had 5 sequences, Cluster 2 had 9, Cluster 3 had 3, and Cluster 4 had one. So, not only are the classical loci duplicated in the common yellowthroat but possible nonclassical or pseudogene loci are likely duplicated as well.

The greater number of loci found in the common yellowthroat compared with other bird species may be attributed to our use of multiple primer sets. In multigene families, particular primers might selectively amplify some genes or alleles but not others (Wagner et al. 1994). We used 3 different primer sets on the gDNA, and each of them amplified sequences not amplified by the other 2 primer sets. Cluster 1 sequences, for example, were amplified almost exclusively by the intron primer combination Int1f.7/Int2r.1, whereas this primer set did not amplify any sequences from Cluster 2. The implication of this finding is that MHC studies may be underestimating the amount of variation present depending on the primers used. In addition to different detection capabilities of different primer sets, the same primer set may detect a different number of loci in different species. The 326/325 exon primer set appeared to amplify only a single locus in the great snipe (*Gallinago media*; Ekblom et al. (2003), whereas it amplified 17 confirmed sequences in a single common yellowthroat (WNR). Also, Aguilar et al. (2006) used the Int1f.7/Int2r.1 intron primer set on the little greenbul and amplified at least 3 apparently classical loci and a single pseudogene, whereas in the common yellowthroat, we amplified at least 2 classical loci and at least 3 nonclassical or pseudogene loci with these primers. Much of the apparent variation in gene number among bird species is likely due to gene duplication, but variation in the primers' ability to detect loci apparently contributes as well.

Our results suggest the presence of at least 20 class II genes in the common yellowthroat, but there may be as many as 31 loci if unconfirmed sequences are included as well. It is known that artifact sequences may arise during PCR due to *Taq* error or the formation of recombinants (Jansen and Ledley 1990; L'Abbé et al. 1992), and this problem may be exacerbated when the number of target sequences is large (Qiu et al. 2001). For this reason, most

MHC studies verify their sequences by confirming them in independent PCRs. Cloning and sequencing can be an inefficient method when confirming sequences because the questionable sequences cannot be targeted for recloning (all sequences are reamplified and recloned), and some sequences appear to be more easily amplified than others. In our case, the large number of sequences present in these individuals made it difficult to get duplicates of individual sequences in independent reactions. Although we identified and removed sequences that were likely PCR artifacts, we cannot rule out the possibility that others exist among the unconfirmed sequences. In turn, we may have inadvertently discarded true sequences as PCR error. Nevertheless, within each individual, the unconfirmed sequences we retained are all $\geq$3 bp different from other unconfirmed and confirmed sequences in that individual, suggesting that many of these represent additional alleles.

Low Variability Loci

Clusters of sequences with low variation (e.g., 1, 2, and 3 in Supplementary Figure S3, Figure 1), which may be individual loci or groups of loci, have been identified in a number of taxa, but their origins and functions vary. Some low variability loci appear to be pseudogenes, having characteristics such as frameshift mutations, stop codons, or mutations in promoter regions that prevent them from producing functional transcripts (e.g., Moon et al. 2005). Among birds, a few low variability class II loci in passerines have been identified as probable pseudogenes (Edwards et al. 2000; Hess et al. 2000; Aguilar et al. 2006), whereas others do not have characteristics of pseudogenes (at least at the portions sequenced), but still lack polymorphism (Gasper et al. 2001; Sato et al. 2001; Jarvi et al. 2004). These other low variability loci may be nonclassical genes, which have been described predominantly at MHC class I but also at class II in mammals (DM and DO regions). In addition to low polymorphism, nonclassical genes have limited expression (e.g., tissue specific) and may have functions differing from classical MHC genes (Klein and Figueroa 1986). These genes might arise when a polymorphic classical locus specializes to present certain types of antigens (Fischer Lindahl et al. 1997) or to perform other specialized functions (e.g., Kropshofer et al. 1998; Zarutskie et al. 2001). In other cases, nonclassical loci are transcribed, but their function remains unknown. This is the case in house sparrows with low variability class I loci (Bonneaud, Sorci, et al. 2004) and in the chicken with nonclassical class II loci found in the Y complex (Zoorob et al. 1993; Miller et al. 2004).

Common yellowthroat sequences fell into multiple low variability clusters, at least 2 of which may contain nonclassical genes. We hypothesize that the clusters correspond to separate loci or groups of loci, due to the degree of sequence divergence at exon 2 between clusters and between the clusters and noncluster sequences. Intron sequences can be better indicators of gene identity than exon 2 because balancing selection on exon 2 may obscure phylogenetic history. In our study, the pattern of clustering at intron 1 supported the pattern of clustering at exon 2, suggesting that sequences from Clusters 1 and 3, as well as noncluster sequences, came from different groups of loci. However, more data are needed to confirm that the clusters represent separate loci. Clusters 1 and 2 (Figure 1) were composed of common yellowthroat sequences only, suggesting these alleles arose more recently. In contrast, Clusters 3 and 4 (which was originally identified by Jarvi et al. [2004]) did include sequences from multiple species, which indicates these allelic lineages are likely older, predating the divergences of those species, although convergent evolution is also a possible explanation. In birds, few orthologous loci have been identified and only between relatively closely related species (Wittzell et al. 1999; Miller and Lambert 2004a; Strand et al. 2007; Burri, Niculita-Hirzel, Salamin, et al. 2008). Interestingly, both Clusters 3 and 4 consist of fewer sequences and thus likely originate from fewer loci within the common yellowthroat than do Clusters 1 and 2 and the noncluster sequences (Supplementary Figure S3; Table 2). It appears likely that Clusters 1 and 3 contain nonclassical genes. Our cDNA sequencing revealed one unconfirmed sequence from Cluster 1, as well as 3 unconfirmed sequences from Cluster 3. This suggests that at least some of these loci are functional and not pseudogenes. We confirmed 2 sequences with frameshift mutations within Cluster 1, so either these are pseudoalleles at functional loci, or Cluster 1 contains a mix of functional and nonfunctional genes. In contrast, sequences within Cluster 2 showed no evidence of being from pseudogenes (although we sequenced only a small portion of the genes), but we did not recover any of these from the cDNA. They could be expressed in a different tissue or they are possible pseudogenes. It is possible that the clusters we have chosen to recognize as probable low variability loci are spurious because they result from a single individual or are relatively invariant because of the low number of sequences in each cluster (3, 5, and 12). However, the corresponding sequence divergence at intron 1 supports the idea that these sequences originate from different groups of loci, and the sequences at our smallest cluster (3 sequences in Cluster 3) are not only highly similar to each other but also show high similarity to 2 sequences from another species, suggesting it is not an artifact of sequence number.

Evidence of Recombination and Selection

We confirmed a total of 76 exon 2 sequences from only 4 common yellowthroat individuals, an indication of high allelic variability in this species. Two processes responsible for the high polymorphism typically seen at MHC genes are recombination and balancing selection. Recombination shuffles variation, thereby forming new alleles, and it is thought to play a central role in generating MHC polymorphism (Gyllensten et al. 1991; She et al. 1991). Evidence for both intra- and interlocus recombination at the MHC has been found in a number of taxa (e.g., Richman et al. 2003; Miller and Lambert 2004a; Reusch and Langefors

2005; Schaschl et al. 2006). We found evidence that recombination may be contributing to the high allelic diversity within the common yellowthroat, at least among the noncluster and transcribed sequences. Interestingly, we found no evidence of recombination between noncluster and cluster sequences or between sequences from different clusters. Perhaps this is because sequences at the cluster loci have more specialized functions that make recombining with divergent sequences disadvantageous.

Balancing selection is also responsible for high MHC polymorphism as it acts to retain alleles over time. An excess of nonsynonymous mutations ($d_N/d_S > 1$) at putative peptide-binding codons is characteristic of classical MHC genes (Hughes and Nei 1988, 1989; reviewed in Bernatchez and Landry 2003; Garrigan and Hedrick 2003). Among the common yellowthroat sequences, we found strong evidence of positive selection on the noncluster and transcribed sequences. Putative peptide-binding codons showed an excess of nonsynonymous mutations, and CODEML also identified a number of sites under positive selection. In contrast, the sequences within the 3 clusters had low mutation rates or no mutations at all at the peptide-binding codons. Hess et al. (2000) and Aguilar et al. (2006) found d_N/d_S ratios not significantly different from neutral at low variability pseudogenes, whereas others have found d_N/d_S ratios less than 1 at peptide-binding codons of lower variability loci and have speculated that it was the result of purifying selection (Jarvi et al. 2004; Axtner and Sommer 2007). Cluster 2 did have an excess of nonsynonymous mutations at the putative peptide-binding codons; however, the overall mutation rate was very low, so this may be a spurious result. Garrigan and Hedrick (2003) noted that an excess of nonsynonymous mutations is not definitive evidence of current selection because it takes time for sequences to accumulate an excess of nonsynonymous mutations when under selection and it takes a long time for the signal to be erased in the absence of selection.

It has been noted that, although selection appears to favor high MHC variability at the population level, each individual expresses only a small proportion of the existing alleles, which is somewhat unexpected given the high potential for gene duplication (Woelfing et al. 2009). In fact, there is some evidence in polyploid species that duplicate loci have been silenced (Shum et al. 2001; Sammut et al. 2002), suggesting it is advantageous to maintain a lower number of expressed loci. The primary explanation for this is that as the number of MHC alleles (class I and II combined) increases, there is first positive selection for an increased number of T-cell lines that are capable of interacting with MHC/peptide complexes in general, followed by a specific negative selection against T-cell lines that are self-reactive (reviewed in Milinski 2006; Woelfing et al. 2009). Modeling shows that negative selection likely outweighs positive selection (Nowak et al. 1992; Woelfing et al. 2009; but see Borghans et al. 2003) so having too many MHC alleles leads to depletion of the T-cell repertoire and reduced immunocompetence. Woelfing et al. (2009) predicted that the optimal number of MHC molecules should be between 3 and 25 per individual, but a larger optimum could be favored by a high diversity of pathogens. We amplified at least 16 sequences (with possibly as many as 34 different transcribed sequences if confirmed) in a single common yellowthroat from class II alone. It will be interesting to determine whether class I has undergone a similar duplication of transcribed loci, what the optimum number of MHC alleles is in this species, and whether it is also higher than in other taxa.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

Research Growth Initiative grant (101×146) and College of Letters and Science, University of Wisconsin—Milwaukee.

Acknowledgments

We thank John Eimes for advice and help in the laboratory and the Louisiana State University Museum of Natural Science for help with collecting common yellowthroats.

References

Aguilar A, Edwards SV, Smith TB, Wayne RK. 2006. Patterns of variation in MHC class II beta loci of the little greenbul (*Andropadus virens*) with comments on MHC evolution in birds. J Hered. 97:133–142.

Aguilar A, Garza JC. 2007. Patterns of historical balancing selection on the salmonid major histocompatibility complex class II beta gene. J Mol Evol. 65:34–43.

Alcaide M, Edwards SV, Negro JJ. 2007. Characterization, polymorphism, and evolution of MHC class IIB genes in birds of prey. J Mol Evol. 65:541–554.

Alcaide M, Edwards SV, Negro JJ, Serrano D, Tella JL. 2008. Extensive polymorphism and geographical variation at a positively selected MHC class IIB gene of the lesser kestrel (*Falco naumanni*). Mol Ecol. 17:2652–2665.

Anisimova M, Nielsen R, Yang Z. 2003. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics. 164:1229–1236.

Axtner J, Sommer S. 2007. Gene duplication, allelic diversity, selection processes and adaptive value of MHC class II *DRB* genes of the bank vole, *Clethrionomys glareolus*. Immunogenetics. 59:417–426.

Babik W, Pabijan M, Radwan J. 2008. Contrasting patterns of variation in MHC loci in the Alpine newt. Mol Ecol. 17:2339–2355.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? J Evol Biol. 16:363–377.

Bollmer JL, Vargas FH, Parker PG. 2007. Low MHC variation in the endangered Galápagos penguin (*Spheniscus mendiculus*). Immunogenetics. 59:593–602.

Bonneaud C, Mazuc J, Chastel O, Westerdahl H, Sorci G. 2004. Terminal investment induced by immune challenge and fitness traits associated with major histocompatibility complex in the house sparrow. Evolution. 58:2823–2830.

Bonneaud C, Sorci G, Morin V, Westerdahl H, Zoorob R, Wittzell H. 2004. Diversity of *Mhc* class I and IIB genes in house sparrows (*Passer domesticus*). Immunogenetics. 55:855–865.

Borghans JAM, Noest AJ, De Boer RJ. 2003. Thymic selection does not limit the individual MHC diversity. Eur J Immunol. 33:3353–3358.

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

Bryant D, Moulton V. 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol. 21:255–265.

Bryja J, Galan M, Charbonnel N, Cosson JF. 2006. Duplication, balancing selection and trans-species evolution explain the high levels of polymorphism of the *DQA* MHC class II gene in voles (Arvicolinae). Immunogenetics. 58:191–202.

Burri R, Niculita-Hirzel H, Roulin A, Fumagalli L. 2008. Isolation and characterization of major histocompatibility complex (MHC) class IIB genes in the Barn owl (Aves: *Tyto alba*). Immunogenetics. 60:543–550.

Burri R, Niculita-Hirzel H, Salamin N, Roulin A, Fumagalli L. 2008. Evolutionary patterns of MHC class IIB in owls and their implications for the understanding of avian MHC evolution. Mol Biol Evol. 25:1180–1191.

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

Doherty PC, Zinkernagel RM. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 complex. Nature. 256:50–52.

Dunn PO, Garvin JC, Whittingham LA, Freeman-Gallant CR, Hasselquist D. 2010. Carotenoid and melanin-based ornaments signal similar aspects of male quality in two populations of the common yellowthroat. Funct Ecol. 24:149–158.

Dunn PO, Whittingham LA, Freeman-Gallant CR, DeCoste J. 2008. Geographic variation in the function of ornaments in the common yellowthroat *Geothlypis trichas*. J Avian Biol. 39:66–72.

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

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

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

Ekblom R, Grahn M, Höglund J. 2003. Patterns of polymorphism in the MHC class II of a non-passerine bird, the great snipe (*Gallinago media*). Immunogenetics. 54:734–741.

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

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

Garrigan D, Hedrick PW. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution. 57:1707–1722.

Garvin JC, Abroe B, Pedersen MC, Dunn PO, Whittingham LA. 2006. Immune response of nestling warblers varies with extra-pair paternity and temperature. Mol Ecol. 15:3833–3840.

Garvin JC, Dunn PO, Whittingham LA, Steeber DA, Hasselquist D. 2008. Do male ornaments signal immunity in the common yellowthroat? Behav Ecol. 19:54–60.

Gasper JS, Shiina T, Inoko H, Edwards SV. 2001. Songbird genomics: analysis of 45 kb upstream of a polymorphic *Mhc* class II gene in red-winged blackbirds (*Agelaius phoeniceus*). Genomics. 75:26–34.

Gaudieri S, Dawkins RL, Habara K, Kulski JK, Gojobori T. 2000. SNP profile within the human major histocompatibility complex reveals an extreme and interrupted level of nucleotide diversity. Genome Res. 10:1579–1586.

Guindon S, Gascuel O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.

Gyllensten UB, Sundvall M, Erlich HA. 1991. Allelic diversity is generated by intraexon sequence exchange at the *DRB1* locus of primates. Proc Natl Acad Sci U S A. 88:3686–3690.

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

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 41:95–98.

Hess CM, Edwards SV. 2002. The evolution of the major histocompatibility complex in birds. Bioscience. 52:423–431.

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

Hosomichi K, Miller MM, Goto GM, Wang Y, Suzuki S, Kulski JK, Nishibori M, Inoko H, Hanzawa K, Shiina T. 2008. Contribution of mutation, recombination, and gene conversion to chicken *Mhc-B* haplotype diversity. J Immunol. 181:3393–3399.

Hughes AL. 2005. Consistent across-tissue signatures of differential gene expression in Crohn's disease. Immunogenetics. 57:709–716.

Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature. 335:167–170.

Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci U S A. 86:958–962.

Hughes AL, Nei M. 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol Biol Evol. 7:491–514.

Hughes CR, Miles S, Walbroehl JM. 2008. Support for the minimal essential MHC hypothesis: a parrot with a single, highly polymorphic MHC class II *B* gene. Immunogenetics. 60:219–231.

Huson D. 1998. SplitsTree: a progam for analyzing and visualizing evolutionary data. Bioinformatics. 14:68–73.

Huson D, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23:254–267.

Ilmonen P, Penn DJ, Damjanovich K, Morrison L, Ghotbi L, Potts WK. 2007. Major histocompatibility complex heterozygosity reduces fitness in experimentally infected mice. Genetics. 176:2501–2508.

Jansen R, Ledley FD. 1990. Disruption of phase during PCR amplification and cloning of heterozygous target sequences. Nucleic Acids Res. 18:5153–5156.

Jarvi SI, Tarr CL, McIntosh CE, Atkinson CT, Fleischer RC. 2004. Natural selection of the major histocompatibility complex (*Mhc*) in Hawaiian honeycreepers (Drepanidinae). Mol Ecol. 13:2157–2168.

Kalbe M, Eizaguirre C, Dankert I, Reusch TBH, Sommerfeld RD, Wegner KM, Milinski M. 2009. Lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. Proc R Soc Lond B Biol Sci. 276:925–934.

Kaufman J, Milne S, Göbel TWF, Walker B, Jacob JP, Auffray C, Zoorob R, Beck S. 1999. The chicken B locus is a minimal essential major histocompatibility complex. Nature. 40:923–925.

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

Kikkawa EF, Tsuda TT, Sumiyama D, Naruse TK, Fukuda M, Kurita M, Wilson RP, LeMaho Y, Miller GD, Tsuda M, et al. 2009. Trans-species

polymorphism of the *Mhc* class II *DRB*-like gene in banded penguins (genus *Spheniscus*). Immunogenetics. 61:341–352.

Klein J. 1986. Natural history of the major histocompatibility complex. New York: Wiley.

Klein J, Figueroa F. 1986. Evolution of the major histocompatibility complex. CRC Crit Rev Immunol. 6:295–386.

Kobari F, Sato K, Shum BP, Tochinai S, Katagiri M, Ishibashi T, Du Pasquier L, Flajnik MF, Kasahara M. 1995. Exon–intron organization of *Xenopus* MHC class II β chain genes. Immunogenetics. 42:376–385.

Kropshofer H, Vogt AB, Thery C, Armandola EA, Li B-C, Moldenhauer G, Amigorena S, Hämmerling GJ. 1998. A role for HLA-DO as a co-chaperone of HLA-DM in peptide loading of MHC class II molecules. Embo J. 17:2971–2981.

Kumánovics A, Takada T, Fischer Lindahl K. 2003. Genomic organization of the mammalian *Mhc*. Annu Rev Immunol. 21:629–657.

L'Abbé D, Belmaaza A, Décary F, Chartrand P. 1992. Elimination of heteroduplex artifacts when sequencing *HLA* genes amplified by polymerase chain reaction (PCR). Immunogenetics. 35:395–397.

Málaga-Trillo E, Zaleska-Rutczynska Z, McAndrew B, Vincek V, Figueroa F, Sültmann H, Klein J. 1998. Linkage relationships and haplotype polymorphism among cichlid *Mhc* class II *B* loci. Genetics. 149:1527–1537.

McClelland EE, Penn DJ, Potts WK. 2003. Major histocompatibility complex heterozygote superiority during coinfection. Infect Immun. 71:2079–2086.

Milinski M. 2006. The major histocompatibility complex, sexual selection, and mate choice. Annu Rev Ecol Syst. 37:159–186.

Miller HC, Belov K, Daugherty CH. 2005. Characterization of MHC class II genes from an ancient reptile lineage, *Sphenodon* (tuatara). Immunogenetics. 57:883–891.

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

Miller HC, Lambert DM. 2004b. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). Mol Ecol. 13:3709–3721.

Miller MM, Bacon LD, Hala K, Hunt HD, Ewald SJ, Kaufman J, Zoorob R, Briles WE. 2004. Nomenclature for the chicken major histocompatibility (*B* and *Y*) complex. Immunogenetics. 56:261–279.

Moon DA, Veniamin SM, Parks-Dely JA, Magor KE. 2005. The MHC of the duck (*Anas platyrhynchos*) contains five differentially expressed class I genes. J Immunol. 175:6702–6712.

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

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

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

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 39:121–152.

Nowak MA, Tarczy-Hornoch K, Austyn JM. 1992. The optimal number of major histocompatibility complex molecules in an individual. Proc Natl Acad Sci U S A. 89:10896–10899.

Oliver MK, Telfer S, Piertney SB. 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). Proc R Soc Lond B Biol Sci. 276:1119–1128.

Penn DJ, Potts WK. 1999. The evolution of mating preferences and major histocompatibility complex genes. Am Nat. 153:145–164.

Posada D. 2002. Evaluation of methods for detecting recombination from DNA sequences: empirical data. Mol Biol Evol. 19:708–717.

Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM, Zhou J. 2001. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. Appl Environ Microbiol. 67:880–887.

Reusch TBH, Langefors A. 2005. Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback *Gasterosteus aculeatus*. J Mol Evol. 61:531–U45.

Richman AD, Herrera LG, Nash D, Schierup MH. 2003. Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in *Peromyscus maniculatus*. Genet Res. 82:89–99.

Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SGE. 2003. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res. 31:311–314.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analsyes by the coalescent and other methods. Bioinformatics. 19:2496–2497.

Sammut B, Marcuz A, Du Pasquier L. 2002. The fate of duplicated MHC class Ia genes in a dodecaploid amphibian, *Xenopus ruwenzoriensis*. Eur J Immunol. 32:2698–2709.

Sato A, Figueroa F, Mayer WE, Grant PR, Grant PR, Klein J. 2000. *Mhc* class II genes of Darwin's Finches: divergence by point mutations and reciprocal recombination. In: Kasahara M, editor. Major histocompatibility complex: evolution, structure and function. Tokyo (Japan): Springer-Verlag. p. 518–541.

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

Sawyer SA. 1999. GENECONV: a computer package for the statistical detection of gene conversion *[Internet]*. St. Louis (MO): Distributed by the author, Department of Mathematics, Washington University Available from: http://www.math.wustl.edu/~sawyer

Schaschl H, Wandeler P, Suchentrunk F, Obexer-Ruff G, Goodman SJ. 2006. Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. Heredity. 97:427–437.

She JX, Boehme SA, Wang TW, Bonhomme F, Wakeland EK. 1991. Amplification of major histocompatibility complex class II gene diversity by intraexonic recombination. Proc Natl Acad Sci USA. 88:453–457.

Shiina T, Briles WE, Goto RM, Hosomichi K, Yanagiya K, Shimizu S, Inoko H, Miller MM. 2007. Extended gene map reveals tripartite motif, C-type lectin, and Ig superfanily type genes within a subregion of the chicken MHC-B affecting infectious disease. J Immunol. 178:7162–7172.

Shiina T, Shimizu S, Hosomichi K, Kohara S, Watanabe S, Hanzawa K, Beck S, Kulski JK, Inoko H. 2004. Comparative genomic analysis of two avian (quail and chicken) MHC regions. J Immunol. 172:6751–6763.

Shum BP, Guethlein L, Flodin LR, Adkison MA, Hedrick RP, Nehring RB, Stet RJM, Secombes C, Parham P. 2001. Modes of salmonid MHC class I and II evolution differ from the primate paradigm. J Immunol. 166: 3297–3308.

Strand T, Westerdahl H, Höglund J, Alatalo RV, Siitari H. 2007. The *Mhc* class II of the Black grouse (*Tetrao tetrix*) consists of low numbers of *B* and *Y* genes with variable diversity and expression. Immunogenetics. 59: 725–734.

Takahashi K, Rooney AP, Nei M. 2000. Origins and divergence times of mammalian class II MHC gene clusters. J Hered. 91:198–204.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Tarof SA, Dunn PO, Whittingham LA. 2005. Dual functions of a melanin-based ornament in the common yellowthroat. Proc R Soc Lond B Biol Sci. 272:1121–1127.

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

Tsuda TT, Tsuda M, Naruse T, Kawata H, Ando A, Shiina T, Fukuda M, Kurita M, LeMaho I, Kulski JK, et al. 2001. Phylogenetic analysis of penguin (*Sphheniscidae*) species based on sequence variation in MHC class II genes. Immunogenetics. 53:712–716.

Wagner A, Blackstone N, Cartwright P, Dick M, Misof B, Snow P, Wagner GP, Bartels J, Murtha M, Pendleton J. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Systematic Biol. 43:250–261.

Walsh HE, Friesen VL. 2003. A comparison of intraspecific patterns of DNA sequence variation in mitochondrial DNA, α-enolase, and MHC class II B loci in auklets (Charadriiformes: Alcidae). J Mol Evol. 57: 681–693.

Wegner KM, Reusch TBH, Kalbe M. 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. J Evol Biol. 16:224–232.

Westerdahl H. 2007. Passerine MHC: genetic variation and disease resistance in the wild. J Ornithol. 148:S469–S477.

Westerdahl H, Wittzell H, von Schantz T. 2000. *Mhc* diversity in two passerine birds: no evidence far a minimal essential *Mhc*. Immunogenetics. 52:92–100.

Westerdahl H, Wittzell H, von Schantz T, Bensch S. 2004. MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. Heredity. 92:534–542.

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

Woelfing B, Traulsen A, Milinski M, Boehm T. 2009. Does intra-individual major histocompatibility complex diversity keep a golden mean? Philos Trans R Soc B Biol Sci. 364:117–128.

Yang Z, Nielsen R, Goldman N, Pedersen A-MK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 155:431–449.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22:1107–1118.

Yuhki N, Beck T, Stephens RM, Nishigaki Y, Newmann K, O'Brien SJ. 2003. Comparative genome organization of human, murine, and feline MHC class II region. Genome Res. 13:1169–1179.

Zarutskie JA, Busch R, Zavala-Ruiz Z, Rushe M, Mellins ED, Stern LJ. 2001. The kinetic basis of peptide exchange catalysis by HLA-DM. Proc Natl Acad Sci U S A. 98:12450–12455.

Zeisset I, Beebee TJC. 2009. Molecular characterization of major histocompatibility complex class II alleles in the common frog, *Rana temporaria*. Mol Ecol Resour. 9:738–745.

Zhang B, Fang SG, Xi YM. 2006. Major histocompatibility complex variation in the endangered crested ibis Nipponia nippon and implications for reintroduction. Biochem Genet. 44:113–123.

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

Received October 6, 2009; Revised January 28, 2010; Accepted January 28, 2010

Corresponding Editor: Jerry Dodgson