Heredity (2004) 92, 534–542

www.nature.com/hdy

MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE

H Westerdahl[1], H Wittzell[2], T von Schantz[1] and S Bensch[1]
[1]*Department of Animal Ecology, Ecology Building, Lund University, S-223 62 Lund, Sweden;* [2]*Botanical Museum, Lund University, S-223 61 Lund, Sweden*

The major histocompatibility complex (MHC) has a central role in the specific immune defence of vertebrates. Exon 3 of MHC class I genes encodes the domain that binds and presents peptides from pathogens that trigger immune reactions. Here we develop a fast population screening method for detecting genetic variation in the MHC class I genes of birds. We found evidence of at least 15 exon 3 sequences in the investigated great reed warbler individual. The organisation of the great reed warbler MHC class I genes suggested that a locus-specific screening protocol is impractical due to the high similarity between alleles across loci, including the introns flanking exon 3. Therefore, we used motif-specific PCR to amplify two subsets of alleles (exon 3 sequences) that were separated with by DGGE. The motif-specific primers amplify a substantial proportion of the transcribed class I alleles (2–12 alleles per individual) from as many as six class I loci. Although not exhaustive, this gives a reliable estimate of the class I variation. The method is highly repeatable and more sensitive in detecting genetic variation than the RFLP method. The motif-specific primers also allow us to avoid screening pseudogenes. In our study population of great reed warblers, we found a high level of genetic variation in MHC class I, and no less than 234 DGGE genotypes were detected among 248 screened individuals.
Heredity (2004) **92,** 534–542. doi:10.1038/sj.hdy.6800450

Keywords: great reed warbler; passerine; MHC class I; motif-specific amplification; DGGE; genetic variation

Introduction

In vertebrates the major histocompatibility complex (MHC) plays a central role in the specific immune defence against various pathogens. The MHC encodes cell-surface proteins that bind peptides within its peptide-binding region (PBR) (Brown *et al*, 1993), and an appropriate immune reaction will follow when the bound peptide is non-self. Compared with other coding genes the MHC genes exhibit an extremely high level of polymorphism. Balancing selection is thought to maintain this polymorphism either through frequency-dependent or through overdominant (heterozygote advantage) selection, pathogens being the major selective agent (Hughes and Nei, 1988, 1989; Potts and Wakeland, 1990; Takahata *et al*, 1992; Satta, 1993; Hughes and Hughes, 1995).

Understanding the redundancy of MHC alleles and exploring the evolutionary history of these genes is very important in evolutionary and ecological studies using them as genetic markers. Polymorphic mammalian MHC genes are thought to evolve by the birth and death process; new genes are created by gene duplication and then some genes are maintained in the genome while others become nonfunctional (Nei *et al*, 1997). In humans MHC alleles from the same locus form a single cluster. Moreover, alleles from different loci have different antigen specificity (Nei *et al*, 1997; Gu and Nei, 1999). This facilitates PCR-based single-locus protocols, as locus-specific primers can be designed. In birds, on the other hand, concerted evolution seems to be operating through inter-locus gene conversion and/or unequal crossing-over of repeated gene sequences. This homogenises alleles across loci and makes it difficult to study one locus at a time (Edwards *et al*, 1995, 2000; Wittzell *et al*, 1999b).

The MHC consists of two major classes of genes, class I and class II. In earlier studies we have sequenced transcribed class I and class II genes from a cDNA library constructed from a single great reed warbler *Acrocephalus arundinaceus* (nestling 692), and we detected a large number of unique sequences within this individual (Westerdahl *et al*, 1999, 2000). These and other studies (Edwards *et al*, 1998; Wittzell *et al*, 1999a; Hess *et al*, 2000; Gasper *et al*, 2001) show that the passerine MHC is more variable and complex than that of most galliform birds (Kaufman *et al*, 1995, 1999; von Schantz *et al*, 1996; Shiina *et al*, 1999).

In the present study, we wanted to identify as many MHC class I genes as possible in the great reed warbler genome. Secondly, we aimed at developing a screening method to investigate the genetic variation in the MHC class I genes of the birds in our study population. Our emerging understanding of the organisation of the great reed warbler MHC class I genes suggested that a locus-specific protocol would not work due to the

Correspondence: H Westerdahl, Department of Animal Ecology, Ecology Building, Lund University, S-223 62 Lund, Sweden.
E-mail: Helena.Westerdahl@zooekol.lu.se
Received 23 March 2003; revised 13 June 2003; accepted 20 October 2003

homogenisation of alleles across loci. Therefore, we developed an alternative screening method that allows for the amplification of only a subset of transcribed MHC alleles at a time. We used this method to screen 248 great reed warblers hatched between 1983 and 1993 for MHC class I variation. We also investigated segregation of MHC alleles in families and linkage between MHC alleles among unrelated individuals.

Materials and methods

Study population

The great reed warbler is a long-distant migratory bird, breeding in northern and central Palaearctic and wintering in sub-Saharan Africa. Our study population in Lake Kvismaren (59°10′N, 15°25′E) has been investigated intensively for nearly 20 years (eg Bensch and Hasselquist, 1991; Hasselquist *et al*, 1995). Each year the majority of breeding males and females (>98%) have been caught in mist-nets, measured, weighed and ringed with an aluminium ring and a unique combination of colour rings. The first blood samples were collected in 1987 and, since then, samples have been taken from almost all (>95%) breeding birds and nestlings.

Sequencing of introns that flank exon 3 in nestling 692

Intron 2: To isolate intron 2 sequences, we constructed sequence-specific primers in exon 2 and in exon 3 using prior knowledge from cDNA sequences (Westerdahl *et al*, 1999). We used the GeneAmp PCR kit (Perkin-Elmer, Foster City, CA, USA) when running PCRs and tested six different primer combinations (HN30-HN40, HN30-HN41, HN30-HN42, HN31-HN40, HN31-HN41 and HN31-HN42) on genomic DNA from the same juvenile from which we had obtained the cDNA library (nestling 692). The sequences of primers are given in Table 1. The PCRs were run in a total volume of 40 μl and included 25 ng of genomic DNA, 0.5 μM of each of the primer pair, 1 × PCR buffer, 0.150 mM dNTP, 2.0 mM $MgCl_2$ and 1.0 U of *Taq* polymerase. The thermal profiles were as follows: 35 cycles at 94°C, (58–63)°C, 72°C, each for 30s, using standard procedures in a thermal cycler GeneAmp PCR System 9600 (Perkin Elmer). The PCR product was purified through precipitation in NH_4Ac and used as template in dye terminator sequencing reactions, according to the manufacturer's protocol (Perkin Elmer). Following the cycle sequencing reaction the excess dye terminators were removed using precipitation in NaAc and the sequences were run on an ABI PRISM 310 Genetic Analyser (Perkin Elmer).

Intron 3: When preparing intron 3 sequences, we used two general primers (primers that would amplify intron 3 sequences from all class I loci known so far in great reed warblers), one in exon 3 (HN11) and one in exon 4 (HN22; Table 1). DNA from nestling 692 was used as the template and PCRs were run as above with an annealing temperature of 66°C. The PCR products were cloned into a pGEM-T vector, according to the manufacturer's protocol (pGEM-T Vector Systems, Promega, USA). PCR templates for cycle sequencing were prepared from positive colonies as in Westerdahl *et al* (1999) and the sequences were run as described above.

DNA sequencing of genomic exon 3 sequences in nestling 692 and its family

Two general primers were designed, one in intron 2 (HN34) and one in intron 3 (HN45; Table 1). The PCRs were run at an annealing temperature of 63°C using genomic DNA from nestling 692, its full-sibs and parents (Perkin Elmer). The PCR products were cloned and sequenced, and this procedure was repeated several times to identify as many different exon 3 sequences as possible from each individual (three times for the sibs and twice for the parents). The risk of PCR artefacts is high when sequencing alleles from multigene families like the MHC since several allele sequences contain similar motifs. Therefore, only alleles that were found in at least two independent PCRs were regarded as unique, verified sequences. We will call the verified exon 3 sequences 'alleles' to facilitate future discussions. However, we are aware that they are not true alleles since they may be located at different loci.

Sequence-specific amplification of exon 3 in great reed warblers

Although both the introns adjacent to exon 3 vary considerable in length, no locus-specific sequences have been detected in these two regions (see Results). Therefore, we designed a protocol where one motif-specific primer within exon 3 and one general primer in intron 3 were used in motif-specific PCRs (Figure 1). To preferentially

Table 1 Primer names and nucleotide sequences

Primer name	*Primer sequences in 5′–3′ direction*
HN11	AGCGCTGCTGAGATCACCA
HN22	CAAGATCAGCGTCCCGTGTT
HN30	GAATATTGGGATAGGAAC
HN31	CAGAATATTGGGATAGCCAG
HN34	CCATGGGTCTCTGTGGGTA
HN36	TCCCCACAGGTCTCCACACAGT
HN38	TCCCCACAGGTCTCCACACACG
HN40	CAGGTAATTCGTCCAATG
HN41	TCCAGGTAATTCGTCTGTCT
HN42	TTCAGGTAATTCGTCTTC
HN45	CCATGGAATTCCCACAGGAA
HN46	ATCCCAAATTCCCACCCACCTT
GC46	GCGGGCGGCGCGGGGCGCGGGCA GGGCGGCGGGGGCGGGCATCCCA AATTCCCACCCACCTT

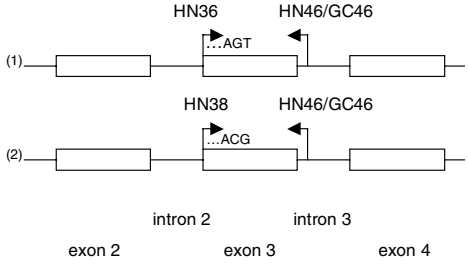

Figure 1 Motif-specific primer combinations that were designed from cDNA sequences so that transcribed sequences are promoted. The sequences that have been amplified with the primer combination (HN36-HN46) have the AGT sequence at positions 12–14 in exon 3 (1), while the sequences that have been amplified with the primer combination (HN38-HN46) have the ACG sequence at positions 12–14 in exon 3 (2).

Table 2 Verified genomic MHC class I exon 3 sequences (Acar-UA*), in nestling 692, its parents (H0-27, V9-79) and full-sibs (691, 693), are here denoted with dots when they exist as both genomic sequence and DGGE band, and with a plus sign when they only exist as genomic sequence

Sequences	*Transcribed*	*Motif*	*H0-27*	*V9-79*	*691*	*692*	*693*
Acar-UA*3	Yes[a]	AGT	•			•	
Acar-UA*20	Yes[a]	AGT	•			•	
Acar-UA*5	Yes[a]	AGT		•		•	
Acar-UA*4		AGT	•	•	•	•	•
Acar-UA*6		AGT		•	•		•
Acar-UA*7		ACG		•	•		•
Acar-new[c]		ACG	•			•	
Acar-UA*8		ACG		•	•		•
Acar-UA*1/UA*15	Yes[a,b]	ATG		+/+		+/+	
Acar-UA*2/UA*19	Yes[a,b]	ACG		+/+		+/+	
Acar-UA*9		ATG			+		+
Acar-UA*10	Yes[a]	ATG	+			+	
Acar-UA*11		ATG			+		+
Acar-UA*12		ATG	+			+	
Acar-UA*13		CGC	+	+	+	+	+
Acar-UA*14		ACT	+		+	+	+
Acar-UA*16 (ψ)		ATG				+	+
Acar-UA*17 (ψ)		ATG	+		+		
Acar-UA*18 (ψ)		ATG		+			
Acar-UA*21 (ψ)		ATG		+	+	+	+
Acar-UA*22 (ψ)		ATG	+				
Verified sequences			10	10	10	15	10
Pseudogene sequences			2	2	2	2	2
Sequenced clones			25	32	35	26	31

[a]Tanscribed sequences (Westerdahl *et al*, 1999).
[b]These cDNA sequences are identical in exon 3 but differ in other parts of the gene (Westerdahl *et al*, 1999).
[c]This DGGE allele has not been verified using cloning procedures and is not included in further analysis.
The motif at positions 12–14 in exon 3 is given. The primer HN36 contains the AGT motif and the primer HN38 contains the ACG motif. PCR products when using the GC46 primer and the HN36/HN38 primers are separated with DGGE. The number of verified sequences, putative pseudogene sequences (ψ) and total number of cloned sequences per individual are also reported.

amplify transcribed alleles, we used the cDNA sequences as references when designing exon 3 primers (HN36 and HN38) (Westerdahl *et al*, 1999). The primer HN36 has the AGT motif in its 3′-end and the primer HN38 has the ACG motif in its 3′-end, both motifs are found at positions 12–14 in exon 3. One additional ATG motif was found in three out of all eight cDNA sequences; however, we chose not to use it since it is also found in several pseudogene sequences (Table 2). The CGC and ACT motifs were found in exon 3 sequences of genomic DNA, but they were not further investigated since they were not polymorphic (Table 2). Both the primer combinations HN36-HN46 and HN38-HN46 were successful and amplified 260 bp from the variable exon 3 (primers not included, the complete exon 3 is 274 bp).

DGGE and DNA sequencing in the population

In order to separate the allelic variants in the sequence-specific amplification of exon 3 we used the DGGE (Denaturing Gradient Gel Electrophoresis) technique (Myers *et al*, 1987; Andersson *et al*, 1993; Lessa and Applebaum, 1993; Miller *et al*, 1999). To improve the separation of alleles, we replaced the primer HN46 with GC46 (Table 1), which differs only in a 5′GC-clamp (Sheffield *et al*, 1989). The DGGE gels contained 7% 19:1 acrylamide/bisacrylamide and 1 × TAE buffer and a denaturatuing gradient of urea and formamide (Myers *et al*, 1987). The DGGE gels were run at 60°C (C.B.C. Scientific Company, Inc.) in 1 × TAE buffer, and the conditions were optimised using gels with 0–80% denaturant. The best separation was found in 40–70% denaturant, and the optimal running time was found by loading samples at 1-h intervals. The PCR products from the primer combination HN36-GC46 were run in 40–65% denaturant while the PCR products from HN38-GC46 were run in 40–70% denaturant. Both combinations were run overnight at 190 V for 16.5 h. All gels included three copies of a marker (ie PCR fragments made from genomic DNA from two and three great reed warbler individuals, respectively, for the two primer combinations HN36-GC46 and HN38-GC46) to enable comparisons between gels. Gels were stained using SYBRgold following the manufacturer's protocol (Molecular Probes, Leiden, The Netherlands), and the DNA was visualised in a FluorImage SI (Molecular Dynamics Inc.).

The migration distance of the bands on the DGGE gels were identified relative to the marker bands. When the parents were known, we confirmed the allelic assignment by comparing the DGGE bands of the offspring with that of their parents (a large number of families, $N = 84$, were included among the screened great reed warblers hatched in Lake Kvismaren between 1983 and 1993). Note that the true paternity of these chicks was verified by multilocus minisatellite DNA fingerprinting (Hasselquist *et al*, 1995).

Two to four randomly picked DGGE bands were excised, using a toothpick, from each DGGE gel, plus the DGGE bands from nestling 692 and its family, and the bands were dissolved in 150 μl of ddH_2O. This solution was frozen (-80°C) and melted (4°C) repeatedly, diluted 1:50, reamplified with the original primers and then directly sequenced as described above. In all, we sequenced 91 DGGE bands, but because most sequences had a few unresolved base pairs, we decided to clone a certain number of the DGGE bands prior to sequencing. Eight different DGGE bands were cloned from two individuals and four clones were sequenced per individual. We name the DGGE bands 'DGGE alleles' and not 'alleles', to avoid confusion with cloned exon 3 sequences, mentioned above.

When running DGGE gels based on a PCR containing multiple alleles, heteroduplexes (a hybrid between strands from two different alleles) may be formed. We know that the two DGGE bands that migrated the shortest distance in the marker are heteroduplexes (bands 7 and 8 in the marker generate the two heteroduplex bands 1 and 2, Figure 4). In order to avoid screening heteroduplexes we only registered bands that migrated further than the marker's number 2 band in the marker when we screened 248 great reed warblers for their DGGE-alleles. However, we cannot eliminate the possibility that heteroduplexes between more similar alleles may be formed and migrate further. In the segregation analysis (see Results), one heteroduplex band was detected in one out of 15 families, thereby confirming our idea that heteroduplexes are only rarely migrating further than the second marker band.

Segregation analyses of DGGE bands were performed in families from 15 matings with three to eight chicks and included in total 28 parents. Three of these matings were 'enlarged broods' in which the same parents had bred twice. Linkage is usually evaluated using a χ^2 test, but this was not possible here due to the small sample size. Instead, we used pairwise simple correlations between DGGE alleles to estimate the degree of linkage among the parental DGGE haplotypes. DGGE alleles that are linked should have a correlation coefficient close to 1 while no linkage should give a correlation coefficient close to zero. This analysis included only DGGE alleles that were found in at least five parental haplotypes (nine DGGE alleles were analysed) and in total 36 pairwise correlations were performed.

Phylogeny- and statistical analysis

The computer package MEGA was used when measuring the evolutionary distance between alleles (ie exon 3 sequences) with the Neighbor-joining method and Kimura 2-parameter model. The number of synonymous (d*S*) and nonsynonymous substitutions (d*N*) were calculated using Nei and Gojobori's method (Kumar *et al*, 1993) and the codons corresponding to the PBR were superimposed as in Westerdahl *et al* (1999). The nucleotide diversity (π) under the Jukes and Cantor model was calculated within and between coding exon 3 sequences and noncoding exon 3 sequences (putative pseudoalleles) using the computer package DnaSP (Rozas and Rozas 1999). The *t*-tests were performed in Microsoft Excel in Office 2000 and the pairwise correlations and the Spearman correlation in SYSTAT 9.0 (Wilkinson, 1998).

Results

Genomic exon 3 sequences in the family

Genomic MHC class I exon 3 sequences were cloned and sequenced from nestling 692, its full-sibs (691, 693) and parents (H0-27, V9-79) (Table 2). In total, 149 genomic exon 3 sequences were sequenced and from these we verified 20 alleles (all these sequences were found twice in independent PCRs and cloning events). Five of these 20 verified exon 3 sequences were putative pseudogenes since they contained a 5-bp deletion (Table 2). Six of the 20 exon 3 sequences have previously been found in cDNA clones (Westerdahl *et al*, 1999).

A phylogenetic tree was constructed from the 20 exon 3 sequences using the Neighbor-joining method. The alleles did not group in significantly supported clusters as would have been expected if the exon 3 sequences had been phylogenetically related within loci (Figure 2). There was a lower nucleotide diversity within the five putative pseudoalleles ($\pi = 0.013 \pm 0.004$) than found in the 15 coding alleles ($\pi = 0.112 \pm 0.013$). The ratio between nonsynonymous substitutions (d*N*) and synonymous substitutions (d*S*) was calculated using Nei and Gojobori's method of pairwise comparisons. The substitution rate was measured separately for the PBR ($dS = 0.343 \pm 0.129$ [mean $\pm$ se], $dN = 0.376 \pm 0.066$) and the non-PBR ($dS = 0.130 \pm 0.024$, $dN = 0.062 \pm 0.010$). In the PBR the dN/dS ratio was 1.1, hence d*N* was only slightly, and not significantly, higher than d*S*.

Introns in nestling 692

Since the exon 3 sequences did not permit us to identify a specific locus, we amplified intron 2 and intron 3 sequences from nestling 692. Four different intron 2 sequences of different lengths (329–662 bp) were detected and three of them could be linked to verified exon 3 sequences in nestling 692 (Figure 3). Seven intron 3 sequences of variable lengths (136–341 bp) were detected (Figure 3). There was high sequence similarity between introns, despite the variation in length, when comparing introns 2 and 3 separately. The length difference was mainly caused by a varying number of repeated blocks, each block containing an 11-bp core sequence (5′-GGGGCTGGGAT-3′) in intron 2 and a 9-bp core sequence (5′-GGAATTCCA-3′) in intron 3 (Figure 3). The variable number of repeats suggests that the introns in passerine MHC genes may be unstable.

Motif-specific amplification and DGGE polymorphism in the population

Neither the genomic exon 3 sequences (Figure 2) nor the introns (Figure 3) allowed locus identification, hence we could not amplify any locus specifically. To reduce the redundancy in the PCR, we decided to amplify exon 3 sequences with specific motifs in the primer sites. The primer combinations HN36-GC46 and HN38-GC46 both amplify 260 bp from exon 3 but they amplify two different sets of sequences since the primers HN36 and HN38 have different motifs in the 3′-end of the primer (AGT and ACG, respectively; Figure 1, Table 2). Both primer combinations showed positive PCRs in all individuals tested ($N = 248$).

We screened 248 individuals hatched in Lake Kvismaren between 1983 and 1993 for their MHC class I DGGE

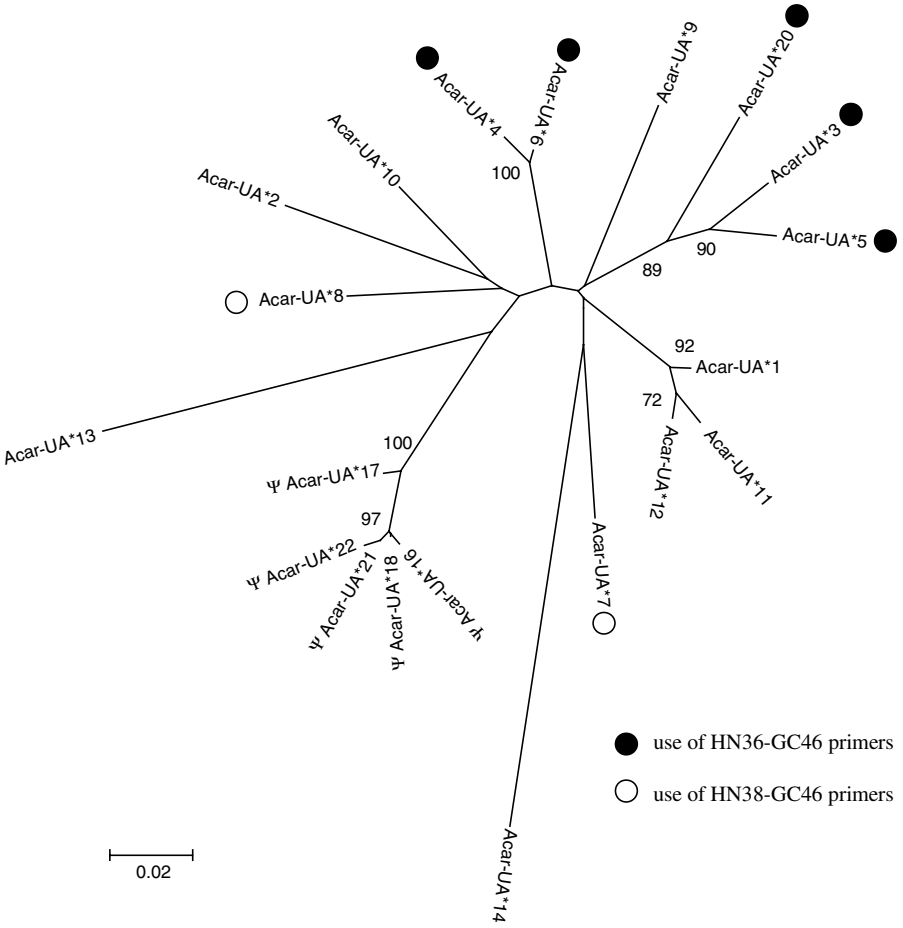

Figure 2 Unrooted tree constructed with the Neighbour-joining method using Kimura 2- parameter model from 20 genomic great reed warbler exon 3 sequences (bootstrap values <50 have been excluded). The sequences that have been amplified with the DGGE primer sets are denoted with circles. The scale bar indicates distance in units of nucleotide substitutions per site.

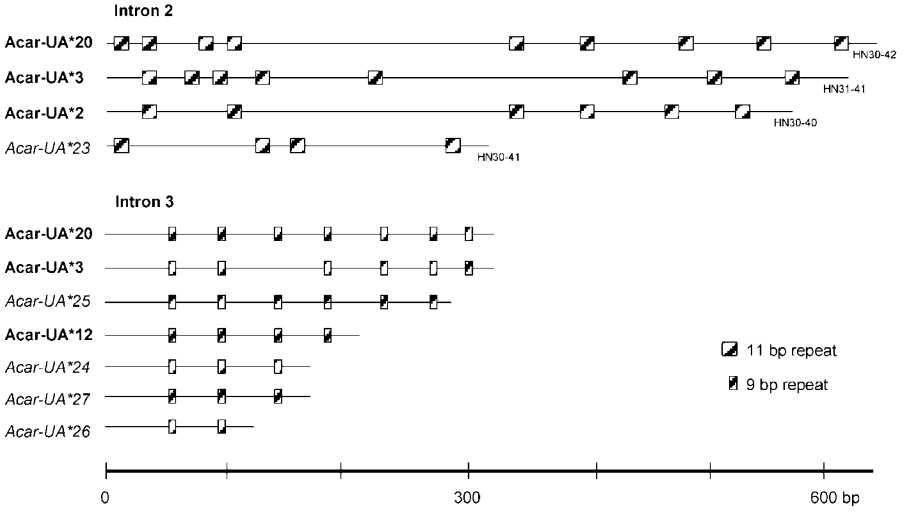

Figure 3 Four intron 2 sequences with the core repeated regions (11 bp) indicated by striped squares and seven intron 3 sequences with the core repeated regions (9 bp) indicated by striped rectangles. All intron sequences are from nestling 692. Introns linked to verified exon 3 sequences in nestling 692 are written in bold (Table 2), while introns linked to unknown exon 3 sequences are written in italics.

genotypes. Each individual had between two and 12 bands (6.4 ± 1.8 (SD)), and therefore we amplify alleles from at least six different loci (if each locus was heterozygous in the 12-banded case). We screened 67 different DGGE bands in total (33 DGGE bands from the primer combination HN36-GC46 and 34 DGGE bands from the primer combination HN38-GC46) and there were 234 DGGE genotypes detected in 248 individuals.

In total, 51% of the screened DGGE alleles were rare, being present in less than 6% of the individuals, while the most frequent DGGE allele was detected in 93% of the individuals. There was no correlation between the number of amplified DGGE bands (DGGE alleles) per individual when comparing the two primer combinations HN36-GC46 and HN38-GC46 (Spearman correlation, $r_S = -0.076$, $p > 0.05$).

DNA sequencing of DGGE bands in the family

Sequencing of all DGGE bands in nestling 692 and its family showed that the separation of DGGE alleles with the DGGE method worked as expected; sequences that are similar in base composition migrated a similar distance in the DGGE gel. The similar sequences, Acar-UA*4 and Acar-UA*6 (which differ in 3 bp; Figure 2), migrated an equal distance and were found in the lower part of the gel (Figure 4). The more diverged sequences Acar-UA*7 and Acar-UA*8 (which differ in 25 bp; Figure 2) were found distant apart in the DGGE gel (Figure 4). The three sequences Acar-UA*3, Acar-UA*5 and Acar-UA*20 are found in the middle of the gel (Figure 4). However, it was not possible to separate Acar-UA*3 and Acar-UA*20 satisfactorily in this DGGE gradient, so the sequences Acar-UA*3 and 20 appeared as a single DGGE band (Figure 4).

We also investigated the correlation between sequenced DGGE bands and previously known exon 3 sequences, which we expected to amplify with the primer combinations HN36-GC46 and HN38-GC 46 in nestling 692 and its full-sibs. There was a perfect match between sequenced DGGE bands and previously known exon 3 sequences in the two full-sibs 691 and 693 (Figure 4). However, one transcribed exon 3 sequence (Acar-UA*2/UA*19) that we expected to amplify in nestling 692 was not detected and, interestingly, a new unknown sequence (Acar-new) was sequenced from a DGGE band (Figure 4, Table 2).

DNA sequencing of DGGE bands in the population

We cloned and sequenced eight DGGE bands in two different individuals to confirm that a particular DGGE allele corresponds to the same sequence. In one case out of eight, a single DGGE band was generated from two exon 3 sequences that differed by 5 bp. The separation of these eight exon 3 sequences on the DGGE gel, in comparison with the separation in the phylogenetic tree (not shown), again showed that similar sequences migrate a similar distance while more diverged sequences are found further apart in the gel.

Segregation of and linkage between DGGE alleles

The segregation of DGGE alleles was studied in 15 different matings. Two matings and presence/absence of 15 DGGE alleles are shown in Figure 5. None of the parents had more than two segregating MHC haplotypes (or haplogroups as we do not know the order of the genes) since the offspring only inherited two different sets of DGGE alleles in each mating. This result indicates that the MHC alleles are situated on the same chromosome and that they are tightly linked. On one occasion a DGGE allele that was not found in any of the parents

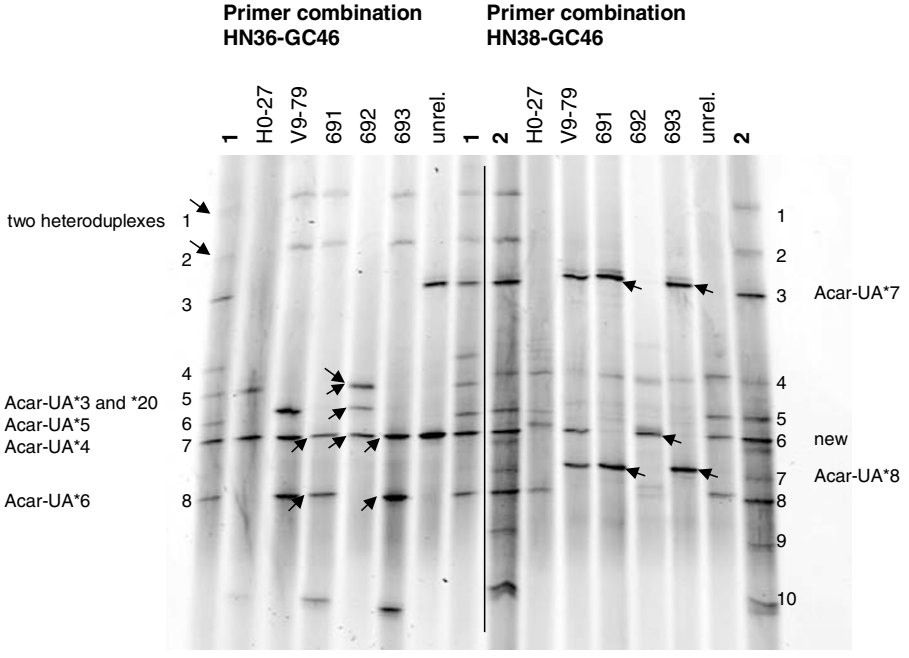

Figure 4 DGGE gel where exon 3 sequences from nestling 692, its full-sibs (691 and 693) and parents (H0-27 and V9-79) and an unrelated individual (unrel.) have been separated using the DGGE method (Table 2). Two different primer combinations were used, HN36-GC46 and HN38-GC46, respectively, and two different size markers were used (markers 1 and 2). Three different DGGE bands were separated in nestling 692 using the primer-combination HN36-GC46, represented by four sequences (Acar-UA*3, Acar-UA*4, Acar-UA*5 and Acar-UA*20 that are denoted with arrows. Using the primer combination HN38-GC46 we found a single sequence (Acar-new) in nestling 692 (denoted with an arrow). Two different DGGE bands were separated in nestling 691 and 693 using primer combination HN36-GC46, Acar-UA*4 and Acar-UA*6 (denoted with arrows), and another two sequences were separated using the primer combination HN38-GC46, Acar-UA*7 and Acar-UA*8 (denoted with arrows). Band one and two in marker 1 are heteroduplexes of bands seven and eight, while band one and two in marker 2 are heteroduplexes of bands six and eight.

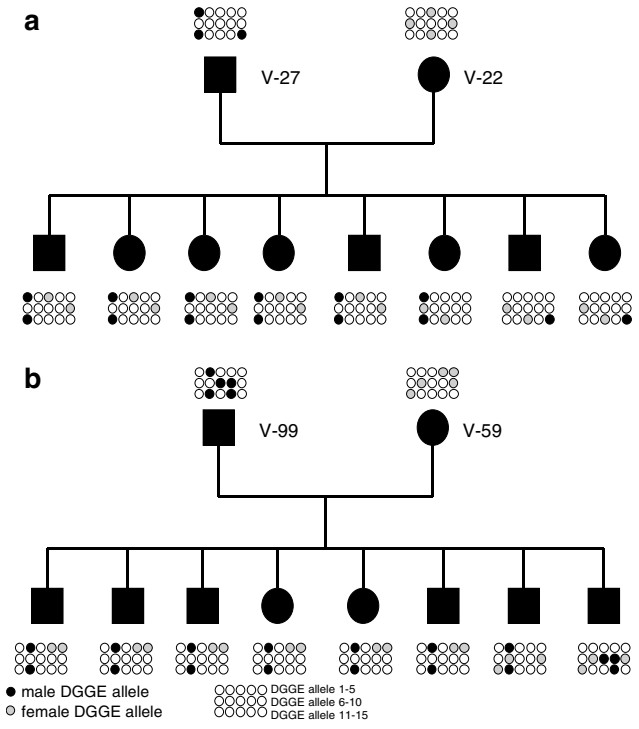

Figure 5 Segregation of 15 parental DGGE alleles (circles) from two matings (**a, b**) are shown. The presence of an allele is indicated with a filled circle, the paternal alleles are black and the maternal are grey. (**a**) The male V-27 has the two haplotypes [DGGE allele 1 and 11] and [DGGE allele 15], and the female V-22 has the two haplotypes [3 and 10] and [6 and 13]. (**b**) The male V-99 has the two haplotypes [2 and 12] and [8, 9 and 14], and the female V-22 has the two haplotypes [3 and 10] and [6 and 13].

appeared in all the chicks. This DGGE allele (which terminated migration in the middle of the gel) is probably a heteroduplex consisting of a unique combination of DGGE alleles only found in the chicks and not in the parents.

The linkage between different DGGE alleles was investigated in 46 DGGE haplotypes (ie all DGGE haplotypes found in the 28 parents mentioned above). Nine DGGE alleles were correlated pairwise and the distribution of the obtained correlations coefficients is illustrated in intervals of 0.1 against the number of correlations (Figure 6). Most correlation coefficients are between ± 0.1 and ± 0.2 and the distribution is around zero. This demonstrates that the MHC alleles are mainly unlinked in the population as a whole, despite segregating as two haplotypes in the families.

Discussion

Multiple class I alleles in a single great reed warbler

We have verified the occurrence of no less than 15 exon 3 MHC class I sequences in one great reed warbler nestling. This is consistent with the observation that the passerine MHC seems to be more variable and complex than the chicken MHC, both in MHC class I (Wittzell *et al*, 1999a; Westerdahl *et al*, 1999) and MHC class II (Edwards *et al*, 1998; Wittzell *et al*, 1999a; Hess *et al*, 2000; Gasper *et al*, 2001). In the domestic chicken there is a single, dominantly expressed, class I gene (two class I genes in all) (Kaufman *et al*, 1995, 1999) and in the great reed warbler there are at least eight class I genes (15 exon 3 sequences), hence there are major differences in the class I organisation among birds.

In the present study we were not able to assign locus identity to the different exon 3, intron 2 and intron 3 sequences (Figures 2 and 3). Alleles at the MHC class I loci of the great reed warbler are indeed very similar to one another. However, we cannot rule out that it is possible to differentiate between loci based on intron 2 and 3 lengths or on the untranslated 3′-region. This possibility can only be resolved by sequencing the entire class I genes, for example by using a cosmid library. Meanwhile, we have developed a screening method based on exon 3 sequences in transcribed MHC class I

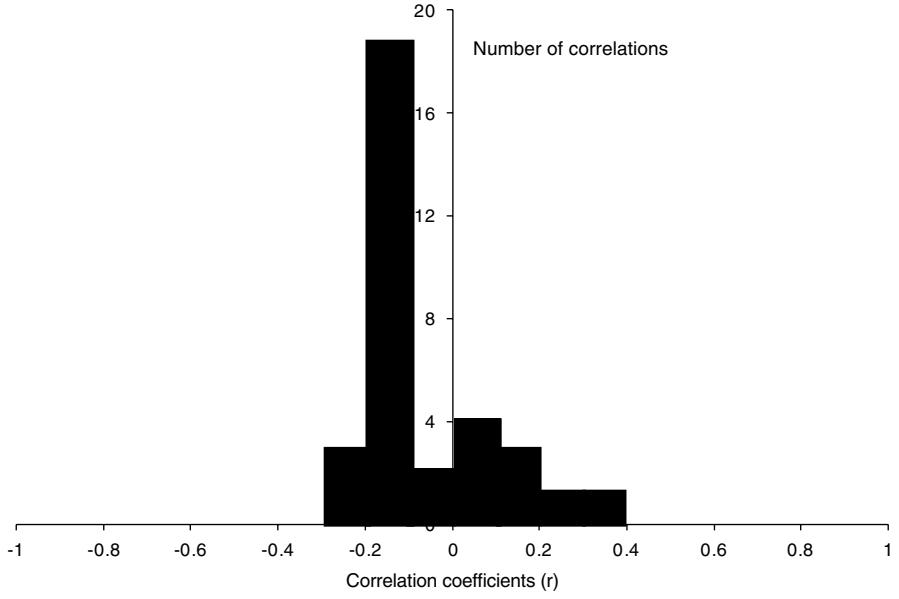

Figure 6 Distribution of pairwise correlation coefficients in intervals of 0.1 (black bars) between each of the nine DGGE alleles found in 46 DGGE parental haplotypes (36 correlations). Tightly linked DGGE alleles would have given a correlation coefficient (*r*) close to one.

alleles (Westerdahl *et al*, 1999). Using this screening method we will be able to investigate how specific MHC alleles occurrence varies over time, (Westerdahl *et al*, 2004) as we have done with microsatellite diversity (Hansson *et al*, 2000), in our study population. We will also be able to investigate whether there is disassortative mating based on MHC genes in great reed warblers, as has been found in humans and mice (reviewed in Penn and Potts 1999).

MHC polymorphism

Most species of passerine birds (great reed warblers (class I and II), bluethroats *Luscinia suecica* (class I), starlings *Sturnus vulgaris* (class I), willow warblers *Phylloscopus trochilus* (class I and II), red-winged blackbirds (class II), scrub jays *Aphelocoma coerulescens* (class II)) in which MHC polymorphisms have been investigated with the RFLP method have a complex RFLP pattern with a large number of bands per individual (Edwards *et al*, 1999; Wittzell *et al*, 1999a; Westerdahl *et al*, 2000; H Westerdahl, unpublished data). To evaluate RFLP band length differences across gels could be difficult, and the method requires much larger quantities of DNA than a PCR-based screening method. The amount of available DNA is often limiting when studying natural populations. The RFLP method is usually thought of as a rather crude method, although we have detected a very high level of genetic variation both in MHC class I and II using this approach in great reed warblers (Westerdahl *et al*, 1999, 2000).

In the present study we focused our screening on exon 3 of MHC class I, which includes the PBR of the MHC molecule, and investigated the polymorphism within this exon using the DGGE method. Miller *et al* (1999) had a similar approach when studying genetic variation in MHC class I genes of chinook salmon *Oncorhynchus tshawytscha*.

Exon 3 of MHC class I is a genetic marker under selection and an important factor in reactions of the immune defence, hence particularly interesting to study from an ecological and evolutionary perspective. We have tested the motif-specific amplification of exon 3 sequences in a number of songbirds using the sequence-specific primer combinations (HN36-GC46 and HN38-GC46) designed for the great reed warbler and the DGGE method. In all these species (*Phylloscopus trochilus, A. gracilirostris, A. stentoreus, A. griseldis, A. seychellensis*) there were positive PCRs (Richardson and Westerdahl, 2003; H Westerdahl, unpublished data). Using motif-specific amplification we have not managed to strictly differentiate between DGGE alleles and heteroduplexes, although we know that heteroduplexes often migrate a shorter distance in the gel. However, the segregation analyses showed that heteroduplexes probably are rare since a new DGGE band, not found in the parents, only appeared in the chicks in a single mating out of 15 different matings. Furthermore, excised DGGE bands often result in a DNA sequence with a few unresolved base pairs. This problem is most easily overcome by cloning every new DGGE allele. Occasionally two or three exon 3 sequences generate an identical DGGE band. This could cause an underestimation of the genetic diversity; however, from our own experience these sequences often have a high sequence similarity.

The segregation analysis of DGGE bands showed that two DGGE haplotypes segregate from each parental DGGE genotype in every investigated mating (Figure 5). This indicates that MHC class I loci investigated here are situated on a single chromosome, confirming an earlier segregation analysis using RFLP (Westerdahl *et al*, 2000), and that the MHC alleles within each haplotype are tightly linked. However, we did not find evidence of strong linkage between the DGGE alleles in different DGGE haplotypes among unrelated individuals (Figure 6). This suggests that recombination has occurred between haplotypes and/or that identical exon 3 sequences are found in different class I genes due to other processes such as concerted evolution.

The great reed warbler population of Lake Kvismaren

The presence of 234 DGGE genotypes in 248 individuals indicates a huge amount of variation in the MHC class I genes in the great reed warbler population of Lake Kvismaren. Data from neutral markers such as multilocus minisatellite DNA fingerprinting and microsatellites, show that the genetic variation in the breeding population at Lake Kvismaren has increased between 1987 and 1993 (Bensch *et al*, 1994; Hansson *et al*, 2000). In addition, when studying polymorphism at the MHC between the same years using either the DGGE or the RFLP method we find that nearly all individuals have unique DGGE or RFLP genotypes (Westerdahl *et al*, 1999). Genetic diversity at the selected MHC genes seems to have been present already when the great reed warbler population was newly founded and small in 1987.

The present screening method using sequence motif-specific amplification and DGGE is simple and PCR-based. This method is particularly useful in species with a low level of genetic variation since it is nearly as sensitive as DNA sequencing in detecting variation. It is also suitable when studying species with a large number of MHC alleles that are similar to one another due to concerted evolution and where locus designation is difficult to interpret.

Acknowledgements

We are grateful to Camille Bonneaud and David Richardson for improving the English and to Åsa Langefors for constructive comments on earlier drafts of the manuscript. The present study was financed by grants from the Swedish Research Council for Forestry and Agriculture to H Wittzell (HW), the Swedish Natural Science Research Council to T von Schantz (TvS), the Swedish Council for Planning and Coordination of Research to TvS and HW, Carl Trygger's Foundation for Scientific Research to TvS and Carl Tesdorf's Foundation to TvS.

References

Andersson B, Jia-Hsu Y, Lewis DE, Gibbs RA (1993). Rapid characterization of HIV-1 sequence diversity using denaturating gradient gel electrophoresis and automated DNA sequencing. *PCR Methods Appl* **2**: 293–300.

Bensch S, Hasselquist D (1991). Territory infidelity in the polygynous great reed warbler *Acrocephalus arundinaceus*: the

effect of variation in territory attractiveness. *J Anim Ecol* **60**: 857–871.

Bensch S, Hasselquist D, von Schantz T (1994). Genetic similarity between parents predicts hatching failure, non-insestuous inbreeding in the great reed warbler? *Evolution* **48**: 317–326.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL *et al* (1993). Three-dimentional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**: 33–39.

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, 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, Hess CM, Gasper J, Garrigan D (1999). Toward an evolutionary genomics of the avian Mhc. *Immunol Rev* **167**: 119–132.

Edwards SV, Wakeland EK, Potts WK (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.

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.

Gu X, Nei M (1999). Locus specificity of polymorphic alleles and evolution by a birth-and-death process in mammalian MHC genes. *Mol Biol Evol* **16**: 147–156.

Hansson B, Bensch S, Hasselquist D, Lillandt B-G, Wennerberg L, von Schantz T (2000). Increase of genetic variation over time in a recently founded population of great reed warblers (*Acrocephalus arundinaceus*) revealed by microsatellites and DNA fingerprinting. *Mol Ecol* **9**: 1529–1538.

Hasselquist D, Bensch S, von Schantz T (1995). Low frequency of extrapair paternity in the polygynous great reed warbler. *Acrocephalus arundinaceus Behav Ecol* **6**: 27–38.

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.

Hughes AL, Nei M (1988). Patterns 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 the major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA* **86**: 958–962.

Hughes AL, Hughes MK (1995). Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics* **42**: 233–243.

Kaufman J, Jacob J, Shaw I, Walker B, Milne S, Beck S *et al* (1999). Gene organisation determines evolution of function in the chicken MHC. *Immunol Rev* **167**: 101–117.

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

Kumar S, Tamura K, Nei M (1993). *MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01*. The Pennsylvania State University: University Park PA.

Lessa EP, Applebaum G (1993). Screening techniques for detecting allelic variation in DNA sequences. *Mol Ecol* **2**: 119–129.

Miller KM, Ming TJ, Schulze AD, Withler RE (1999). Denaturing gradient gel electrophoresis (DGGE): a rapid and sensitive technique to screen nucleotide sequence variation in populations. *Biotechniques* **27**: 1020–1024.

Myers RM, Maniatis T, Lerman LS (1987). Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol* **155**: 501–527.

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 USA* **94**: 7799–7806.

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

Potts WK, Wakeland EK (1990). Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol* **5**: 181–187.

Richardson DS, Westerdahl H (2003). MHC divesity in two *Acrocephalus* Species: the outbred Great reed Warbler and the inbred Seychelles Warbler. *Mol Ecol* **12**: 3523–3529.

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

Satta Y (1993). Balancing selection at HLA loci. In: Takahata N, Clark AG (eds) *Mechanisms of Molecular Evolution*. Sinauer: Sunderland, pp 129–149.

Sheffield VC, Cox DR, Lerman LS, Myers RM (1989). Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc Natl Acad Sci USA* **86**: 232–236.

Shiina T, Oka A, Imanishi T, Hanzawa K, Gojobori T, Watanabe S *et al* (1999). Multiple class I loci expressed by the quail. *Mhc Immunogenet* **49**: 456–460.

Takahata N, Satta Y, Klein J (1992). Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* **130**: 925–938.

von Schantz T, Wittzell H, Goransson G, Grahn M, Persson K (1996). MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proc Roy Soc London B* **263**: 265–271.

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

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

Westerdahl H, Hansson B, Bensch S, Hasselquist D (2004). Between-year variation of MHC allele frequencies in great reed Warblers: selection or drift? *JEB* **17**: 485–492.

Wilkinson L (1998). *SYSTAT 9.0 for Windows, SPSS Inc*. Evanston, IL.

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

Wittzell H, Madsen T, Westerdahl H, Shine R, von Schantz T (1999b). MHC variation in birds and reptiles. *Genetica* **104**: 301–309.