Immunogenetics (2008) 60:677–688
DOI 10.1007/s00251-008-0323-1

ORIGINAL PAPER

Characterization of expressed class II MHC sequences in the banner-tailed kangaroo rat (*Dipodomys spectabilis*) reveals multiple *DRB* loci

Joseph D. Busch · Peter M. Waser · J. Andrew DeWoody

Received: 14 May 2008 / Accepted: 16 July 2008 / Published online: 3 October 2008

Abstract Genes of the major histocompatibility complex (MHC) are exceptionally polymorphic due to the combined effects of natural and sexual selection. Most research in wild populations has focused on the second exon of a single class II locus (*DRB*), but complete gene sequences can provide an illuminating backdrop for studies of intragenic selection, recombination, and organization. To this end, we characterized class II loci in the banner-tailed kangaroo rat (*Dipodomys spectabilis*). Seven *DRB*-like sequences (provisionally named *MhcDisp-DRB**01 through *07) were isolated from spleen cDNA and most likely comprise ≥5 loci; this multiformity is quite unlike the situation in muroid rodents such as *Mus*, *Rattus*, and *Peromyscus*. In silico translation revealed the presence of important structural residues for glycosylation sites, salt bonds, and CD4+ T-cell recognition. Amino-acid distances varied widely among the seven sequences (2–34%). Nuclear DNA sequences from the *Disp-DRB**07 locus (~10 kb) revealed a conventional exon/intron structure as well as a number of microsatellites and short interspersed nuclear elements (B4, *Alu*, and IDL-Geo subfamilies). Rates of nucleotide substitution at *Disp-DRB**07 are similar in both exons and introns (π=0.015 and 0.012, respectively), which suggests relaxed selection and may indicate that this locus is an expressed pseudogene. Finally, we performed BLASTn searches against *Dipodomys ordii* genomic sequences (unassembled reads) and find 90–97% nucleotide similarity between the two kangaroo rat species. Collectively, these data suggest that class II diversity in heteromyid rodents is based on polylocism and departs from the muroid architecture.

Keywords MHC architecture · SINEs · Gene cluster · cDNA · Hybridization enrichment

Introduction

Major histocompatibility complex (MHC) genes have long been of interest to immunologists but are now widely used by population biologists as well (Acevedo-Whitehouse and Cunningham 2006; Milinski 2006; Piertney and Olivier 2006). Immune system genes are an intuitive choice for addressing ecological and evolutionary questions because of their central role in clearing pathogenic infections (Klein 1986). MHC class I and class II genes typically exhibit an evolutionary signature of balancing selection and are among the most polymorphic loci in vertebrate genomes (Apanius et al. 1997; Hughes and Yeager 1998). A comparison of vertebrate MHC regions reveals the dynamic evolutionary history of this gene family in which loci are

Electronic supplementary material The online version of this article (doi:10.1007/s00251-008-0323-1) contains supplementary material, which is available to authorized users.

Nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession numbers EU817477–EU817485.

J. D. Busch (✉) · J. A. DeWoody
Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA
e-mail: jdbusch@purdue.edu

P. M. Waser
Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

duplicated, lost, and/or disabled (Kelley et al. 2005). The number of class I and class II loci, the allotment of polymorphism within each class, and the chromosomal position of each linkage group are evolutionarily labile (Klein 1986; Kindt and Singer 1987; Yuhki et al. 2003), but mammals in particular display a dramatic duplication of MHC genes (Kelley et al. 2005).

In the class II region, mammals have enhanced their antigen-presenting capability by duplicating loci, which encode the alpha and beta subunits of class II molecules (e.g., *DRB1*, *DRB2*, *DRB3* etc.). This has led to a complex assortment of class II configurations. Commonly, a small number of functional class II beta genes (*DRB*, *DQB*, *DPB*) are expressed, each with a high level of allelic diversity at the population level. For instance, humans maintain 542 *DRB1* alleles, 87 *DQB1* alleles, and 127 *DPB1* alleles worldwide (IGMT/HLA Database 2007; Robinson et al. 2003). Variants of this architecture may include expansion or contraction of polygenic clusters: Domestic cats (*Felis catus*) have three functioning *DRB* loci, no *DQB* genes, and a nonfunctional *DPB* locus (Yuhki et al. 2003), while mole rats (*Spalax ehrenbergi*) have discarded *DRB* genes entirely in favor of an expanded *DPB* cluster (Nizetic et al. 1987). A completely different class II configuration is employed by macaques (*Macaca* spp.) and the California sea lion (*Zalophus californicus*), which have a variable number of *DRB* loci per individual (two to eight) with limited polymorphism at any single locus (three to 40 alleles; Doxiadis et al. 2001; Bowen et al. 2004). The total level of *DRB* diversity in macaques and sea lions is driven by the expression of multiple genes within the *DRB* cluster rather than major allelism at one to two loci (de Groot et al. 2004).

Class II architectures have been discerned in >15 model organisms (Debenham et al. 2005; Kelley et al. 2005; Schaschle and Wegner 2007) but remain poorly described in nonmodel species because most studies focus on the second exon of single *DRB* and *DQB* genes (Bernatchez and Landry 2003; Nigenda-Morales et al. 2008). In rodents, nearly all previous MHC work has occurred within two muroid families, Cricetidae and Muridae (Figueroa et al. 1990; Richman et al. 2003; Smulders et al. 2003; Froeschke and Sommer 2005; Harf and Sommer 2005; Oliver and Piertney 2006; Babik and Radwan 2007). The class II region of model organisms such as *Mus*, *Rattus*, and *Peromyscus* has one to two *DQB* genes and a single copy of *DRB*. However, there are more than 30 nonmurid rodent families containing approximately 1,500 species (Wilson and Reeder 2005). The only full-length *DRB* sequences available outside of superfamily Muroidea come from the Abert's squirrel, *Sciurus aberti* (Wettstein and coworkers, unpublished), although exon 2 sequences have been sampled from single *DRB* loci in Castoridae (Babik et al. 2005) and Ctenomyidae (Cutrera and Lacey 2007). The MHC architectures of nonmurid rodents are unknown; because rodents diversified some 55–90 MYA (Huchon et al. 2007) and have rapid rates of molecular evolution (Li et al. 1996; Triant and DeWoody 2006), it remains to be determined if MHC organization will be conserved across families. Here, we characterize full-length class II MHC sequences in a heteromyid rodent, the banner-tailed kangaroo rat (*Dipodomys spectabilis*). This species has served as an ecological model for over two decades (Brown et al. 1986; Waser et al. 2006), and the genomic architecture of the loci we describe reveals an unusual configuration of *DRB* multiformity compared to most members of the superfamily Muroidea.

Materials and methods

Spleen complementary DNA (cDNA)

We collected a single adult female *D. spectabilis* in December, 2004, at the "Rucker" population (Busch et al. 2007) in southeastern Arizona, USA (Cochise County, 31°56′ N, 109°5′ W). Spleen and other tissues from the animal were immediately frozen in liquid nitrogen with subsequent storage at −80°C. Approximately 40 mg of spleen tissue was used to collect total RNA using the TRIzol® procedure (Invitrogen). The SMART™ PCR cDNA synthesis kit (ClonTech) was used to reverse transcribe spleen messenger RNA (mRNA) using PowerScript™ enzyme primed by a poly-T oligo. We replaced the *Rsa*I recognition sequence in all three SMART™ oligos with an *Nhe*I restriction site to facilitate downstream cloning (see Table S1). After stockpiling ten replicate first-strand cDNA synthesis reactions, we removed mRNA by adding 1 U of RNase H (Invitrogen) and incubating for 20 min at 37°C. Products were purified using a QIAGEN® PCR prep before proceeding to the next step.

Enrichment and cloning

The kilobase BINDER™ kit (Dynal®) was used to isolate full-length MHC transcripts. Thirty picomoles of biotinylated DRBEXON2probe (Table S1) was added to 1 μg of first-strand cDNA product in 100 μl (final volume) of medium-stringency SSPE hybridization solution (final concentration=675 mM NaCl, 45 mM NaH_2PO_4, 45 mM EDTA, and 0.1% SDS). This hybridization reaction was denatured at 95°C for 10 min and taken through a stepdown thermal profile as follows: 80°C for 1 min, 75°C for 30 s, followed by 60 decrement steps (from initial 75°C, −0.2°C decrement every 30 s to 55°C final), then 55°C for 1 min and an additional 20 decrement steps (from 55°C, −0.5°C decrement every 30 s to 45°C final). The hybridization was

immediately chilled to 4°C and added to 100 μl of Binding Solution (Dynal) containing 400 μg of M-280 streptavidin DynaBeads®. Conjugation of the biotinylated probe to the paramagnetic beads took place during 3 h of incubation at room temperature (RT) with rotation. A magnet was applied to the tube for 1 min to concentrate the DynaBeads, after which the supernatant was discarded. The pellet of DynaBeads was resuspended in 1× Wash Solution (10 mM Tris-HCl pH 8, 1 mM EDTA, and 2 M NaCl) and sent through six 2-min washes: two at RT, two at 50°C, and two at 60°C. After each wash, DynaBeads were concentrated with a magnet for 1 min, and the supernatant was discarded. Upon completion of the washing steps, the bead complex was resuspended in 100 μl TLE (10 mM Tris-HCl pH 8.5, low EDTA 0.1 mM) and denatured at 95°C for 10 min to release the single-stranded cDNA transcripts. The magnetic beads were pelleted one final time, and the supernatant was transferred to a fresh tube (storage at −20°C).

The supernatant was diluted 1/10, and 5 μl was used as template in a polymerase chain reaction (PCR). Final concentrations of reagents were as follows: 1× ASS buffer (Andrew's Secret Sauce: 10 mM Tris-HCl pH 8.9, 50 mM KCl, and 50 μg/ml BSA), 2 mM $MgCl_2$, 0.2 μM dNTPs, 0.4 μM oligo "5′ PCR primer II A" (Table S1), 1.25 U *Taq* DNA polymerase, and 0.025 U *Pfu* DNA polymerase in a total volume of 50 μl. The thermal profile began with a 95°C step of 1 min 30 s, followed by 32 amplification cycles (95°C for 30 s, T_m for 30 s, and 72°C for 1 min) and a finishing step of 72°C for 10 min. Amplification products were purified using a QIAGEN PCR prep and quantified in an ND-1000 spectrophotometer (Nanodrop). This purified template was subjected to a second round of probe-hybridization and DynaBeads enrichment as above. Products of the second enrichment were digested with *Nhe*I to produce 4 bp overhangs at the SMART oligo restriction sites and purified with a final QIAGEN PCR prep. *Nhe*I fragments were ligated into pBluescript SK+ (Stratagene) and electroporated into *Escherichia coli* DH5α using an Electroporator 2510 (Eppendorf) set to 1,400 V. Colonies were grown overnight on YT agar with 100 μg/ml ampicillin, 50 μg/ml X-gal, and 1% IPTG.

cDNA sequencing

A total of 230 white colonies were screened for MHC inserts using two separate PCRs: The first reaction used universal vector primers T7/T3 to size each insert, while in the second PCR, the T3 primer was paired with DRBEXON2probe. Successful amplification in this second PCR served to validate the presence of MHC class II sequences. Conditions for both PCR screens were as follows: 1× ASS buffer (see above), 2 mM $MgCl_2$, 0.2 μM dNTPs, 0.4 μM each primer, 1.25 U *Taq* DNA polymerase, and 0.025 U *Pfu* DNA polymerase in a total volume of 12.5 μl. The thermal profile began with a 95°C step of 2 min, followed by 32 amplification cycles (95°C for 30 s, T_m for 30 s, and 72°C for 1 min) and a finishing step of 72°C for 10 min. We chose 41 positive clones to sequence using BigDye 3.1 chemistry (Applied Biosystems) with primers T3, T7, DRBexon2-F2, CONS-R1, and *Disp*DRB_exon4_F1 (Table S1).

Sequences were edited and aligned manually in Sequencher 4.1 (GeneCodes). We checked for the presence of long open reading frames (ORFs) in each sequence and compared exon boundaries with human *HLA-DRB1* (GenBank# NP_002115). Any partial transcripts or suspected false alternative transcripts were removed before further analysis (Cocquet et al. 2006). As a final quality control measure, we only included sequences that were represented by at least two *E. coli* clones.

In order to amplify any class II loci missed by the hybridization-enrichment procedure and further validate the ones initially discovered, we designed four new forward primers at the 5′ UTR boundary and one conserved reverse primer in the 3′ UTR (Table S1 and Fig. S1). Each forward and reverse combination was used in a separate PCR that had the same reagents and thermal conditions as our colony screen. This was followed by QIAGEN purification, ligation with a pGEM™-T Easy System I kit (Promega), and electroporation into *E. coli* DH5α (described above). Full-coverage sequencing (2×) was accomplished using primers T7, SP6, DRBexon2-F2, and CONS-R1. Electropherograms were edited and aligned manually in Sequencher 4.1, and as before, sequences were only included if represented by two or more clones.

Genomic DNA (gDNA) intron sequencing

In addition to characterizing exon-derived polymorphism in class II genes, we were interested in estimating the level of intronic variation exhibited by *D. spectabilis*. Following the strategy of von Salome et al. (2007), we used a series of long PCR fragments to span the entire 10-kb nuclear gene (Fig. S2). Exon-anchored primers were designed to amplify three overlapping PCR fragments (Table S1). Each PCR used 50 ng of genomic DNA in the following conditions: 1× ASS buffer (see above), 2 mM $MgCl_2$, 0.2 μM dNTPs, 0.4 μM each primer, 1.25 U *Taq* DNA polymerase, and 0.025 U *Pfu* DNA polymerase in a total volume of 10 μl. The thermal profile began with a 95°C step of 2 min, followed by 32 amplification cycles (95°C for 30 s, T_m for 30 s, and 72°C for 3 min) and a finishing step of 72°C for 30 min. A second *D. spectabilis* individual was included to identify single nucleotide polymorphisms (SNPs) in this sample of four chromosomes. Each of the three PCR

fragments was cloned using a pGEM-T Easy System I kit and sequenced as described above using a suite of intron-specific primers (Table S1 and Fig. S2).

Sequence analysis

Unique class II sequences from cDNA were evaluated for signatures of natural selection as estimated by the ratio of nonsynonymous to synonymous substitutions in DnaSP v4.2 (Rozas et al. 2003). In order to check for selection in individual exons, we analyzed d_N/d_S ratios in (1) exons 1, 2, and 3 separately; (2) the 24 peptide-binding site (PBS) codons of exon 2; and (3) the entire coding sequence (CDS) minus exon 2. The Jukes–Cantor correction for multiple substitutions was used for all calculations (Jukes and Cantor 1969; Nei and Gojobori 1986). Deduced amino-acid residues were aligned in BioEdit (Hall 1999) and used in a neighbor-joining (NJ) phylogenetic analysis that was based on a matrix of simple amino-acid distances in PAUP (Swofford 1989). Bootstrap support was obtained from 1,000 replicates of data resampling. A search for potential recombination breakpoints between class II sequences was performed using the MaxChi method (Maynard Smith 1992; Posada and Crandall 2001) implemented in RDP v2.3 (Martin et al. 2005) with the search window size set to 20 polymorphic nucleotides.

We assayed SNP variation in *Disp-DRB**07 gDNA sequences using DnaSP v4.2. Exon boundaries were determined empirically by aligning the cDNA sequence from *DRB**07; this allowed us to compare rates of evolution in coding versus noncoding regions. We checked all sequences in RepeatMasker to characterize any putative repetitive elements such as short interspersed nuclear elements (SINEs) and microsatellite repeats (A.F.A. Smit, R. Hubley, and P. Green; RepeatMasker at http://repeatmasker.org). Finally, we ran BLASTn (nucleotide–nucleotide) searches of *D. spectabilis* class II sequences against the unassembled Ord kangaroo rat (*Dipodomys ordii*) genome (taxid=10020; Baylor College of Medicine, unpublished). Any *D. ordii* sequence reads with high similarity scores were downloaded, trimmed, and assembled into candidate contigs.

Results

cDNA sequences

The hybridization-enrichment procedure yielded five unique *DRB*-like sequences from 27 clones. This same set of five was verified independently via PCR amplification of cDNA template. Two additional unique sequences were discovered in this PCR step that had not been detected using hybridization enrichment. Multiple lines of evidence indicate these seven unique sequences, which we have named *Disp-DRB**01–*Disp-DRB**07 according to Klein et al. (1990), encode class II beta-subunit molecules. First, BLASTp runs always found closest similarities with other mammalian MHC class II loci, especially *DRB* genes from dog, cattle, and various rodents (Table 1). Second, each

Table 1 Results of protein–protein BLAST (BLASTp) procedure in GenBank for seven putative class II *DRB* sequences in *D. spectabilis*

GenBank accession	*Dipodomys spectabilis* MHC class II	# aa	Species/gene/GenBank accession[a]	*e*-value	% identity
EU817479	*MhcDisp-DRB**01	266	*Canis lupus familiaris DLA-DRB1* (NP_001014768)	2*e*-112	72
			Sciurus aberti DRB (AAA42356)	3*e*-100	73
EU817480	*MhcDisp-DRB**02	266	*C. lupus familiaris*	2*e*-116	72
			Sigmodon hispidus H2-Eb (AAN87895)	1*e*-110	71
EU817481	*MhcDisp-DRB**03	259[b]	*C. lupus familiaris*	4*e*-109	71
			S. hispidus	1*e*-105	70
EU817482	*MhcDisp-DRB**04	266	*C. lupus familiaris*	1*e*-115	73
			S. hispidus	2*e*-112	72
EU817483	*MhcDisp-DRB**05	266	*Bos taurus BLA-DRB3* (BAA08217)	1*e*-112	71
			Rattus norvegicus R2-DRB (AAV40633)	7*e*-101	70
EU817484	*MhcDisp-DRB**06	265	*C. lupus familiaris*	8*e*-99	66
			S. hispidus	2*e*-98	66
EU817485	*MhcDisp-DRB**07	258[b]	*B. taurus BLA-DRB3* (CAA77092)	2*e*-95	63
			Mus musculus H2-Eb (AAB17688)	4*e*-85	63
			*Callithrix jacchus Caja-DRB1*0305* (CAJ34981)[b]	5*e*-92	66

The length of deduced amino-acid sequences is provided. The *e*-value and percentage identity (BLOSM62) are estimated in BLASTp.
[a] Best overall GenBank hit (above); best rodent hit (below). Entries with no accession numbers are identical to preceding entries.
[b] Transcript is missing exon 5

sequence has an ORF that encodes a full-length (266-aa) or nearly full-length (258-aa) transcript. Assuming that the exon and domain boundaries are consistent with *HLA-DRB1* (GenBank# NP_002115), the full-length *D. spectabilis* sequences are predicted to encode a 29-aa signal peptide, a 95-aa β1 domain, a 103-aa β2 domain, a 23-aa transmembrane domain (TM), and a 16-aa cytoplasmic tail (CT). The two shorter sequences, *Disp-DRB**03 and *DRB**07, are missing eight amino acids from the cytoplasmic tail domain and have a stop codon at the 3′ end of exon 4, suggesting that they are pseudogenes or nonclassical class II molecules. (Domain boundaries do not precisely match exon boundaries in *DRB* loci; therefore, these sequences retain a truncated CT.) BLAST searches of the β1, β2, and TM domains (each run separately for all seven sequences) in the IMGT/HLA database consistently predicted homology to mammalian *DRB* (Robinson et al. 2003). Finally, these sequences have retained all of the key amino-acid positions known to be conserved in classically expressed class II loci (Fig. 1), including residues that form disulfide bridges and salt bonds (β1–15, 79, 72, 76), glycosylation sites (β1–19, 21), and locations that interact with the Ig-like domain (β1–33, 93; Brown et al. 1993; Kaufman et al. 1994). Exceptions were found in *Disp-DRB**06, which has a replacement at β1–93 (Arg to Gln), and *Disp-DRB**07, which deviates at β1–45 (Gly to Arg) and β1–72 (Arg to Ala). These various lines of evidence convince us these seven sequences all code for class II beta-subunit molecules.

A neighbor-joining reconstruction (Fig. 2) predicts that *D. spectabilis* sequences are more related to each other than to class II sequences from cricetid (*Sigmodon hispidus*) and murid (*Mus musculus* and *Rattus norvegicus*) rodents. The node leading to *D. spectabilis* received 100% bootstrap support, suggesting that these seven sequences diverged after an ancestral split between Heteromyidae and superfamily Muroidea some 65–75 MYA (Huchon et al. 2007). The amino-acid divergences within the *D. spectabilis* branch ranged from 2% to 34%.

Class II loci often exhibit $d_N/d_S>1$ in exon 2, which is a signature of past or present selection pressure (Hughes and Yeager 1998). Surprisingly, the estimate for exon 2 (mean 0.74, min–max 0.25–1.37) was about the same as the remainder of the CDS (0.82, 0.32–1.64). This changed when only the 24 PBS codons were included, however (1.29, 0.74–2.71). A moderate number of substitutions are found elsewhere in the CDS, except in exon 6. The d_N/d_S estimates in exon 1 retain signatures consistent with neutrality, with a mean value of 0.91 (0.49–1.28). Exon 3, however, shows evidence for balancing selection with a mean d_N/d_S estimate of 1.28 (0–4.69). The upper d_N/d_S values for exon 3 are driven mainly by the divergent *Disp-DRB**06 and *DRB**07 sequences. These results must be interpreted with caution because the sample size is small; however, they provide some insight into the patterns of substitution that occur outside of exon 2.

Shared substitutions provide the information needed to predict the locations of recombination breakpoints using the MaxChi method (Maynard Smith 1992; Posada and Crandall 2001). Because prediction of recombination breakpoints can be challenging, we report only the most likely events. An association between *Disp-DRB**06 and *DRB**07 received the strongest support (P value$=1.5\times10^{-5}$). These sequences share many nucleotide changes in the first two exons, but this pattern drops sharply near the 3′ end of exon 2 (position 277 onward; Fig. S1). An interlocus breakpoint is inferred in this area, either in exon 2 or intron 2, which is consistent with previous studies of gene conversion (Andersson et al. 1991; Reusch and Langefors 2005). A second recombination event is predicted wherein *Disp-DRB**02 and *DRB**07 exchange sequences to form *DRB**03 (P value$=6.14\times10^{-3}$). *DRB**03 is identical to *DRB**07 from the 3′ end of exon 4 onward. The absence of exon 5 in both sequences provides further evidence of this interlocus exchange.

gDNA sequences

A DNA sequence alignment of the three overlapping PCR fragments predicts that the nuclear *Disp-DRB**07 sequence is about 10.1 kb in length (Fig. 3 and Fig. S2). We recovered a single sequence in the original kangaroo rat that matched the *Disp-DRB**07 sequence determined from cDNA. The second kangaroo rat was a heterozygote, with one allele that matched the original *Disp-DRB**07 sequence and a second undescribed allele. Both sequences have been deposited in GenBank (#EU817477 and EU817478). Comparison of the cDNA and gDNA sequences confirmed that exon boundaries in this locus were identical to human *HLA-DRB1* (NP_002115), except for exon 5. Introns varied in size from 1.1 to 3.6 kb.

A pairwise alignment revealed 112 SNPs between these alleles, with a mean density of one SNP per 83.7 bp. The overwhelming majority of SNPs were found in introns (100), though one occurred in exon 2 and 11 more were found in exon 3 (Fig. 3). When scaled according to the total number of nucleotides per category (π) the rates of substitution appear similar for coding and noncoding regions ($\pi=0.015$ and 0.012, respectively). However, the substitution rate in exon 3 alone ($\pi=0.041$) is an order of magnitude greater than in exon 2 ($\pi=0.0038$). This reflects an increase in SNP density found in the region of 6.5–8.5 kb, where exon 3 is found.

The second PCR fragment (Fig. S2) encompassed a region that was not amenable to sequencing (noted in Fig. 3). Despite successful amplification of a smaller PCR fragment that spanned this region, all sequences abruptly

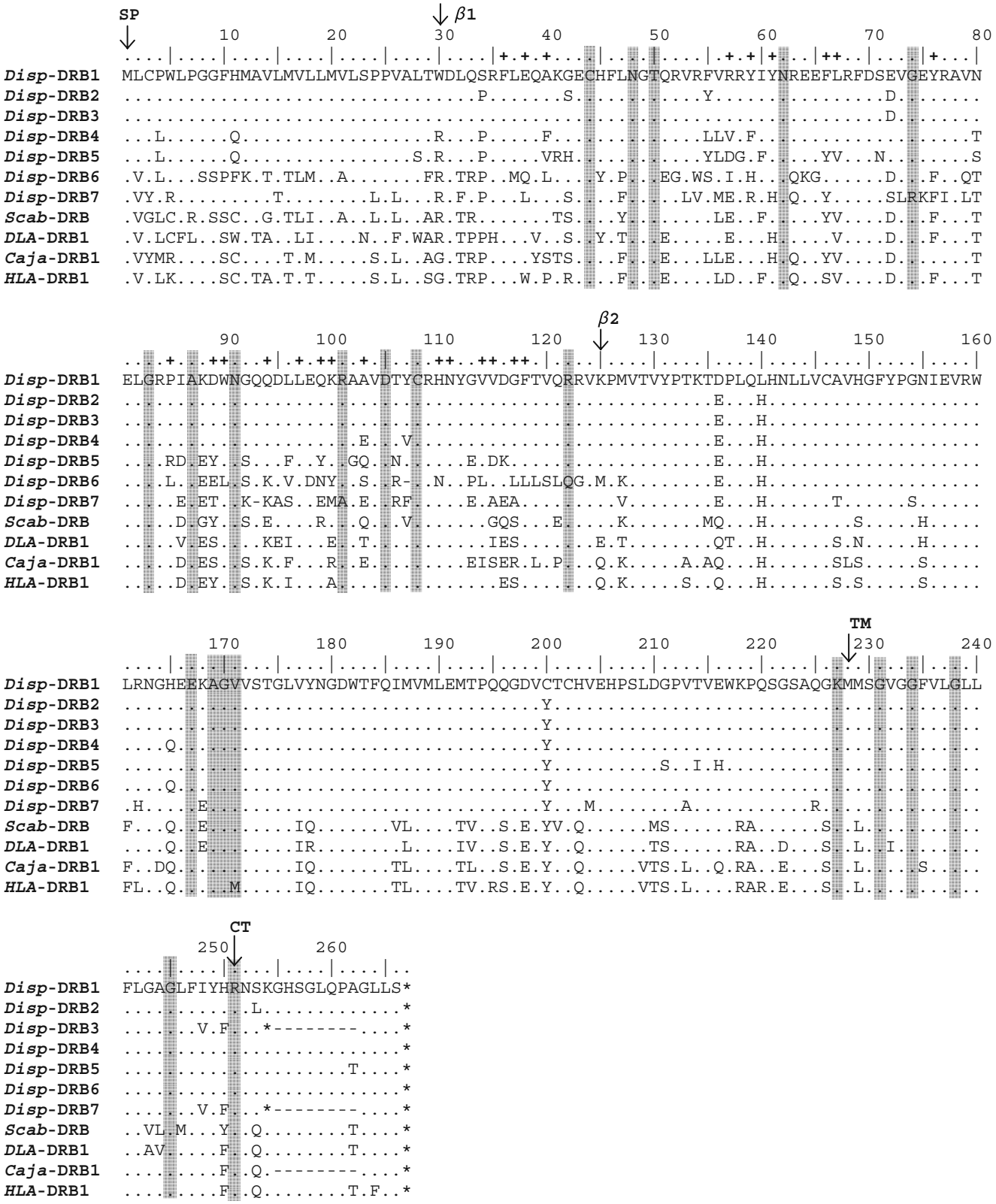

Fig. 1 Deduced amino-acid alignment of *MhcDisp-DRB* sequences (one-letter code). Domain boundaries are based on the human *HLA-DRB1* molecule (NP_002115). *Dots* represent identity to the *Disp-DRB**01 (*top*) amino-acid sequence; *dashes* represent deletions [including missing positions in *DRB**06 (#107) and *DRB**07 (#93)]; *asterisks* denote stop codons; *shaded boxes* show conserved structural residues (Kaufman et al. 1994); *plus signs* identify putative PBS based on human (Brown et al. 1993)

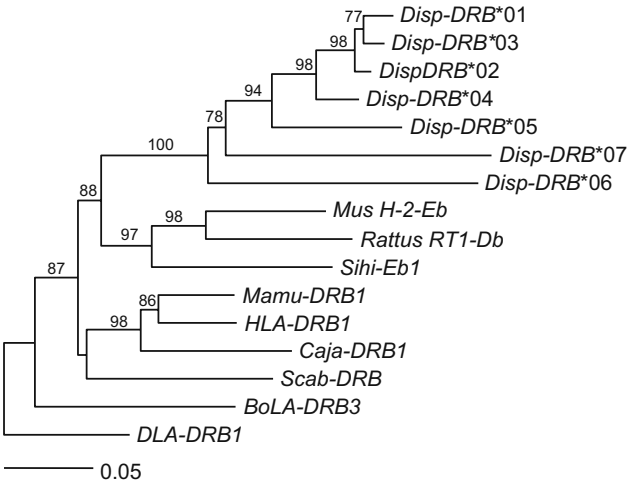

Fig. 2 Neighbor-joining tree constructed from deduced amino-acid alignments of class II sequences from *D. spectabilis* and other mammals. Amino-acid differences were calculated using simple distance (uncorrected). Dog was used as an outgroup and bootstrap values (1,000 replicates) are reported for nodes that received >50% support. Abbreviations and GenBank accession numbers of comparison sequences were: *DLA-DRB1* dog, NP_001014768; *BoLA-DRB3* cattle, BAA08217; *HLA-DRB1* human, NP_002115; *Mamu-DRB1* rhesus macaque, AAB63307; *Caja-DRB1* common marmoset, CAJ34981; *Scab-DRB* Abert's squirrel, AAA42356; *Sihi-Eb1* cotton rat, AAN87895; *Rattus RT1-Db* rat, AAV40633; *Mus H-2-Eb* mouse, AAB17688

terminated in the same area, which was flanked by multiple SINE elements (see below). These repetitive elements may interact to form secondary structure that is impassable with conventional sequencing methods. Therefore, we omitted the middle 742 bp of PCR fragment-2 (positions 3996–4737 in Fig. 3 and Fig. S3), leaving 9,378 nucleotide positions for our survey of molecular variation.

While SNPs provided most of the interallelic variation, 13 indels of 1–5 bp were observed throughout this 10-kb nuclear region, including a single codon deletion in exon 3 found in the second allele. This codon is located in the CD4+ T-cell binding region (amino-acid positions β2 134–147; Kaufman et al. 1994). Although not located on one of the four critical sites, its absence may have consequences for T-cell communication.

Numerous repetitive elements were predicted to exist in the introns of *Disp-DRB**07 (Fig. 3). The repeatmasker search recovered 11 locations that were consistent with SINE-like elements (e.g., IDL, B4, and Alu), representing 12.3% of the total sequence. A B4 SINE and two partial Alu elements formed a close-knit cluster upstream of the 742-bp sequencing gap, with an additional Alu element flanking the downstream side. Introns 1 and 3 had short elements (180 and 191 bp, respectively), which bore high similarity to the IDL-Geo SINE subfamily. IDL-Geo SINEs are found only in the rodent families Geomyidae and Heteromyidae (Gogolevsky and Kramerov 2006).

Lower complexity repeats also exist in *Disp-DRB**07, the most notable of which are two $(GT)_n$ microsatellite repeats inside intron 2. In mammals, this intron often houses one to two microsatellite loci that are tightly linked with exon 2 (Worley et al. 2006). Therefore, we developed two primer sets (Table 2) to test the level of polymorphism

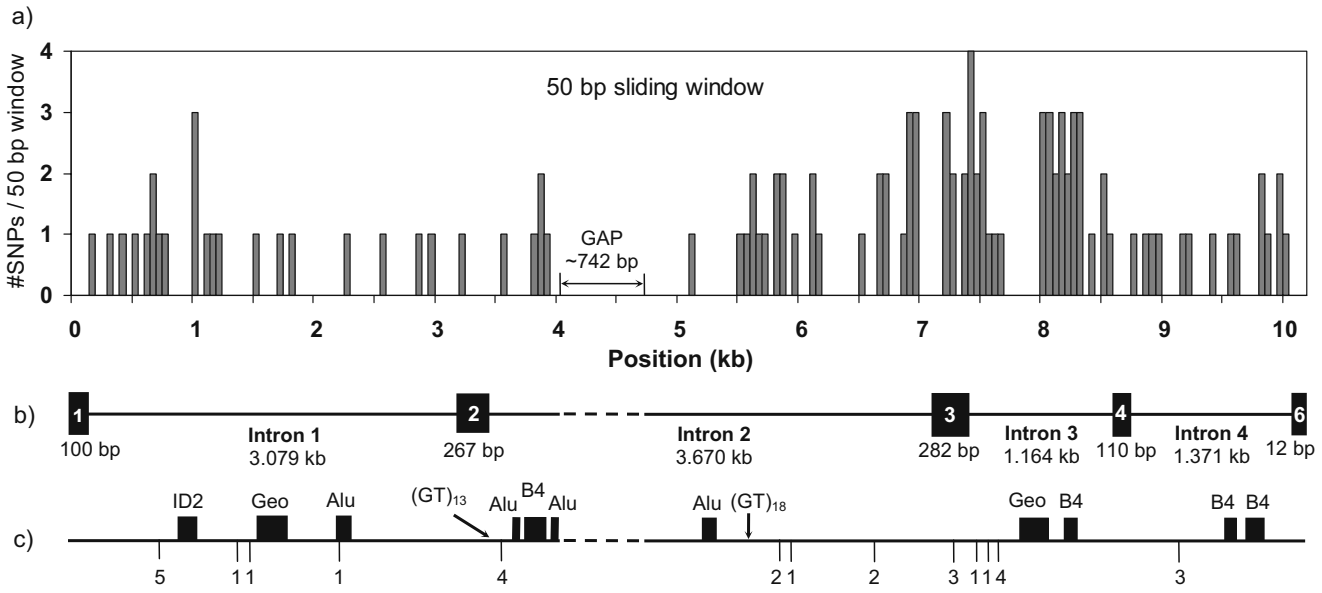

Fig. 3 Features of the gDNA sequence from two *Disp-DRB**07 alleles (GenBank# EU817477 and EU817478). **a** A 50-bp sliding window graph of SNPs from comparison of two sequences. **b** Arrangement of the entire *Disp-DRB**07 locus; *black boxes* are exon boundaries (*labeled in white font*), sizes given below; introns and their sizes are provided below the line; *dashed line* symbolizes the 742 bp gap. **c** Various features of the sequence; names of repetitive elements are given *above* the line; indels and their sizes are given below the line

Table 2 Microsatellite primers from the second intron of *Disp-DRB**07

Name	Location	T_m	Sequence (5′–3′)	Notes
INT2msat1-F2	Intron 2 (5′ end)	64	GCCTGGGAGAGAGGAGAAAGTG	292 bp PCR product
INT2msat1-R1		64	CAGTCCAGGTTAGGCATAAAGTGAGAT	Monomorphic $(GT)_{13}$
INT2msat2-F1	Intron 2 (middle)	65	TTGTTCCTGAAAGTCAGAGCTCCATATTTAT	142–158 bp PCR product
INT2msat2-R1		65	TTGAGCTGAAGAGATCAAGGACAGTG	9 alleles $(GT)_{10}$–$(GT)_{18}$

in a sample of 30 kangaroo rats from the Rucker population, using the PCR conditions of Waser et al. (2006). The $(GT)_{13}$ repeat immediately downstream of exon 2 was monomorphic, while the $(GT)_{18}$ microsatellite in the middle of intron 2 had nine alleles and an observed heterozygosity of 0.85.

BLASTn with *D. ordii*

As a means of complementing our BLASTp searches in GenBank, we made use of the trace archives for the unassembled (and thus unannotated) *D. ordii* genome. Our searches included 9.4 kb of nuclear sequence from *Disp-DRB**07 and 798 bp of cDNA sequence from *Disp-DRB**01, *DRB**05, and *DRB**06. These comparisons found a much higher nucleotide similarity between the *D. spectabilis* and *D. ordii* sequences (97% for *Disp-DRB**07 and 90–93% for the others) than with any of the closest assembled GenBank sequences (77%). A pairwise comparison of the 7.879 kb *D. ordii* contig with the two *Disp-DRB**07 alleles yielded a mean of 240 interspecies differences (π=0.03). These changes were distributed proportionately across the coding and noncoding regions of *DRB**07.

Discussion

The evolution of MHC genes in nonmodel species is generally inferred from the second exon of MHC class II loci (e.g., Bos et al. 2008) because it displays high polymorphism and plays an important role in pathogen defense (Hughes and Yeager 1998). However, comparisons across related clusters of loci can provide a deeper understanding of MHC evolution (Bondinas et al. 2007). Towards this end, we describe seven full-length *DRB* sequences from the banner-tailed kangaroo rat (*D. spectabilis*).

cDNA Discovery

The hybridization enrichment of cDNA from spleen was a key strategy for isolating full-length coding sequences. The protocol we used (ClonTech) incorporates synthetic oligos onto the first-strand cDNA ends in a *cis* orientation so that subsequent amplifications are primed from these linkers and no exonic sequence information is lost. Acquiring the 5′ and 3′ UTRs allowed us to design sequence-specific primers anchored outside of the coding region. The resulting amplifications identified two additional class II sequences, which may have been expressed at levels too low to be detected with a hybridization procedure (e.g., *Disp-DRB**03) or had diverged to the point where hybridization was inefficient (e.g., *Disp-DRB**07). As seen in this and other studies, discovery of multiple MHC loci is enhanced when both strategies are used (Bowen et al. 2004; Strand et al. 2007).

Five of the *D. spectabilis* sequences have full-length ORFs (266-aa) and bear the hallmarks of functional class II loci, including the conservation of amino-acid residues that form structural motifs. Two sequences (*Disp-DRB**03, *DRB**07) have a premature stop codon at the end of exon 4, which would terminate translation at the fourth position of the CT domain. Both are also missing exon 5 but retain exon 6 and the normal stop codon. It is possible that these transcripts are encoded by pseudogenes, which we discuss in greater detail below. Estimates of d_N/d_S ratios across all sequences suggest a weak signature of historic balancing selection on exon 3 and the 24 putative PBS codons in exon 2. An elevated d_N/d_S ratio for exon 3 is unexpected because the β2 domain it encodes is fairly well conserved (IMGT/HLA database; Robinson et al. 2003).

BLAST searches of all seven sequences in the GenBank and IMGT/HLA databases suggest homology to mammalian *DRB* loci. While we expected the greatest similarity would be shared with other rodents, we found slightly higher *e*-values for dog and cattle (Table 1). This is probably because full-length coding sequences are unavailable for other heteromyids. A neighbor-joining reconstruction suggests these *D. spectabilis* sequences are more closely related to each other than to the most similar class II gene in GenBank (Fig. 2). Although trans-species persistence of class II alleles is well known (Klein 1987; Babik and Radwan 2007), it is clear from this tree that heteromyids do not share ancient alleles with Muridae or Cricetidae. This pattern is consistent with other rodent phylogenies based on several nuclear loci and the estimated time (55–65 MYA) since these three families shared a common ancestor (Huchon et al. 2002, 2007).

Assigning sequences into discrete loci is difficult in the absence of chromosome-level information for the class II region. However, the amino-acid dissimilarities (Fig. 1) and the NJ topology (Fig. 2) strongly suggest that we sampled multiple *DRB* genes. The *Disp-DRB**06 and *Disp-DRB**07 sequences likely represent two different loci, based on their large amino-acid divergence from each other (34%) and from the other five *Disp-DRB* sequences (20–28%). Furthermore, only one or two alleles per individual were amplified from the highly polymorphic $(GT)_{18}$ microsatellite, suggesting that *Disp-DRB**07 has not undergone recent duplication.

*Disp-DRB**05 shows an intermediate level of divergence from the four terminal sequences (12–17%) and may be a third locus. Amino-acid distances among the *Disp-DRB**01–*DRB**04 sequences range from 2% to 9%, and their allelic relationships are unclear (Fig. 2). We use the logic of Axtner and Sommer (2007) and conservatively assume that these four sequences (*DRB**01–*DRB**04) represent two loci, as would be the case for a diploid heterozygote, but we cannot determine which sequences are allelic with these data. This suggests that we have described at least five putative class II *DRB* genes, though the assignment of sequences to loci is not trivial (Klein et al. 1990) and will require comparison with the *D. ordii* genome once it is assembled and annotated. Future exploration of these kangaroo rat loci will reveal whether allelic diversity is evenly distributed across all loci or sequestered in one to two genes that exhibit major allelism.

This elevated level of *DRB* duplication is unusual for a rodent; the presence of 1–2 *DRB* loci is more typical (Smulders et al. 2003; Babik et al. 2005; Froeschke and Sommer 2005; Harf and Sommer 2005; Meyer-Lucht and Sommer 2005; Oliver and Piertney 2006; Babik and Radwan 2007; Cutrera and Lacey 2007). Polylocism in other species (e.g., *Peromyscus*) is usually confined to the *DQB* cluster (Richman et al. 2002). Only the bank vole (*Myodes glareolus*) displays *DRB* multiformity at a similar level to *D. spectabilis* (≥4 loci; Axtner and Sommer 2007). In this regard, the nonmurid rodents *D. spectabilis* and *M. glareolus* are more similar to humans, cats, or dogs than to *Mus* and *Rattus*. Our understanding of MHC evolution and architecture in rodents will no doubt be refined as the phylogenetic sample broadens across lineages.

A common source of MHC variation is recombination (both intra- and interlocus), which shuffles diversity among alleles via breakage at nonspecific breakpoints (Yauk et al. 2003; Bos and Waldman 2006; Schaschl et al. 2006). We found reasonable support for two interlocus recombination events in these *D. spectabilis* class II sequences. At least one event has a breakpoint that is consistent with a known recombination hotspot in the second intron of mammalian *DRB* genes (Carrington 1999; Yauk et al. 2003). The other event may have a breakpoint anywhere between this hotspot and exon 4. The predicted involvement of *Disp-DRB**07 in these recombination events necessarily implies that both are interlocus exchanges. With our data, we are unable to discern unequal crossing over from gene conversion between different class II loci (Doxiadis et al. 2006; Reusch and Langefors 2005).

gDNA gene sequence

We investigated a 10.1-kb block of nuclear sequence from the *Disp-DRB**07 locus in order to determine the patterns of intronic variation in the *D. spectabilis* MHC. Alignment against the known cDNA sequences for *Disp-DRB**07 allowed us to determine exon–intron boundaries and compare substitution rates in coding and noncoding regions. Nucleotide variation was low in all coding regions except exon 3. When combined, the average diversity of coding regions (π=0.015) was nearly the same as for noncoding regions (π=0.012). The kangaroo rat intron diversity is similar to related human *DRB1* introns (d= 0.001–0.057; von Salome et al. 2007). We expected a higher rate of neutral substitution in rodents versus primates (Li et al. 1996), but our sample was restricted to just two individuals.

The comparison of exon and intron sequences suggests a pattern of substitution consistent with selective neutrality. *Disp-DRB**07 and *Disp-DRB**03 are both missing exon 5 (CT domain), a feature they share with a marmoset pseudogene (Doxiadis et al. 2006). Pseudogenes are often expressed at a low level, and although their functions remain largely unknown, they could play a role in thymic education/selection or as soluble MHC molecules if premature stop codons occur downstream of the β2 domain (de Groot et al. 2004; Schaschle and Wegner 2007). Some pseudogenes in primates are known to persist over remarkably long evolutionary time spans via genetic hitchhiking or interlocus conversion events that reactivate their functionality (Doxiadis et al. 2006). This creates interesting possibilities for pseudogenes to act as a reservoir of exonic variation that related genes can access (de Groot et al. 2004). Continued surveys of MHC genes will probably uncover a variety of class II sequences whose functions are yet undescribed.

We searched the unassembled *D. ordii* genome for sequences similar to *D. spectabilis* class II loci. The *D. ordii* contigs that we assembled shared 90–97% identity with *D. spectabilis*. The nucleotide divergence among the kangaroo rat species (240 substitutions) was more than twice as large as the polymorphism found within our sample of two *D. spectabilis* alleles (112 SNPs). Apparently, the similarities between *D. spectabilis* and *D. ordii* extend to the unsequenced gap in intron 2 because we were unable to

find long *D. ordii* sequence reads that extended into this region. SINE elements flanked this gap in both kangaroo rat species.

SINEs and other repetitive elements made up 12.3% of the total nuclear sequence of *Disp-DRB**07, which is similar to that found in human *HLA-DRB1* introns (12%; von Salome et al. 2007). The two longest repetitive elements were IDL-Geo SINEs of the ID family (Fig. 3). These occur only in the geomyoid rodents and have been described previously in one other kangaroo rat, *Dipodomys deserti* (Gogolevsky and Kramerov 2006). We also discovered a polymorphic microsatellite repeat $(GT)_{18}$ in *Disp-DRB**07. The type and location of such repetitive elements will be valuable for future delineation of MHC haplotype blocks (von Salome et al. 2007) and for determining homology of class II genes in heteromyid rodents.

Conclusions

To date, evaluations of MHC architecture in rodents have been restricted to the superfamily Muroidea and would suggest that *DRB* multiformity is rare. However, our extensive sequence data demonstrate that polylocism of the *DRB* cluster has been employed by kangaroo rats of the family Heteromyidae. As there are more than 30 rodent families, it remains to be determined if multiple *DRB* loci are common in other species. Given the expansion of the mammalian MHC relative to other vertebrates (Kelley et al. 2005), comparisons of diverse lineages will help crystallize our understanding of the modes and mechanisms of MHC evolution.

Acknowledgements We thank the members of the DeWoody lab, D. Bos, and J. Patton for their insight. J. McCreight provided assistance in DNA sequencing. Funding to J.D.B. was provided by the AMNH Theodore Roosevelt Memorial Fund, the American Society of Mammalogists, and Purdue University. J.A.D.'s lab is funded primarily by the National Science Foundation. This manuscript is ARP#2008-18322 from Purdue University.

References

Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? Trends Ecol Evol 21:433–438 doi:10.1016/j.tree.2006.05.010

Andersson L, Gustafsson K, Jonsson A-K, Rask L (1991) Concerted evolution in a segment of the first domain exon of polymorphic MHC class II b loci. Immunogenetics 33:235–242

Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997) The nature of selection on the major histocompatibility complex. Crit Rev Immunol 17:179–224

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 doi:10.1007/s00251-007-0205-y

Babik W, Durka W, Radwan J (2005) Sequence diversity of the MHC DRB gene in the Eurasian beaver (*Castor fiber*). Mol Ecol 14:4249–4257 doi:10.1111/j.1365-294X.2005.02605.x

Babik W, Radwan J (2007) Sequence diversity of MHC class II *DRB* genes in the bank vole *Myodes glareolus*. Acta Theriol (Warsz) 52:227–235

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 doi:10.1046/j.1420-9101.2003.00531.x

Bondinas GP, Moustakas AK, Papadopoulos GK (2007) The spectrum of *HLA-DQ* and *HLA-DR* alleles, 2006: a listing correlating sequence and structure with function. Immunogenetics 59:539–553 doi:10.1007/s00251-007-0224-8

Bos DH, Waldman B (2006) Evolution by recombination and transspecies polymorphism in the MHC class I gene of *Xenopus laevis*. Mol Biol Evol 23:137–143 doi:10.1093/molbev/msj016

Bos DH, Gopurenko D, Williams RN, DeWoody JA (2008) Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. Evolution Int J Org Evolution 62:1458–1468 doi:10.1111/j.1558-5646.2008.00364.x

Bowen L, Aldridge BM, Gulland F, Van Bonn W, Delong R, Melin S et al (2004) Class II multiformity generated by variable MHC-*DRB* region configurations in the California sea lion (*Zalophus californianus*). Immunogenetics 56:12–27 doi:10.1007/s00251-004-0655-4

Brown JH, Davidson DW, Munger JC, Inouye RS (1986) Experimental community ecology: the desert granivore system. Harper & Row, New York

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL et al (1993) Three-dimensional structure of the human class II histocompatibility antigen *HLA-DR1*. Nature 364:33–39 doi:10.1038/364033a0

Busch JD, Waser PM, DeWoody JA (2007) Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats. Mol Ecol 16:2450–2462 doi:10.1111/j.1365-294X.2007.03283.x

Carrington M (1999) Recombination within the human MHC. Immunol Rev 167:245–256 doi:10.1111/j.1600-065X.1999.tb01397.x

Cocquet J, Chong A, Zhang GL, Veitia RA (2006) Reverse transcriptase template switching and false alternative transcripts. Genomics 88:127–131 doi:10.1016/j.ygeno.2005.12.013

Cutrera AP, Lacey EA (2007) Trans-species polymorphism and evidence of selection on class II MHC loci in tuco-tucos (Rodentia: Ctenomyidae). Immunogenetics 59:937–948 doi:10.1007/s00251-007-0261-3

de Groot N, Doxiadis GGM, de Groot NG, Otting N, Heijmans C, Rouweler AJ et al (2004) Genetic makeup of the *DR* region in rhesus macaques: gene content, transcripts, and pseudogenes. J Immunol 172:6152–6157

Debenham SL, Hart EA, Ashurst JL, Howe KL, Quail MA, Ollier WER et al (2005) Genomic sequence of the class II region of the canine MHC: comparison with the MHC of other mammalian species. Genomics 85:48–59 doi:10.1016/j.ygeno.2004.09.009

Doxiadis GGM, Otting N, De Groot NG, Bontrop RE (2001) Differential evolutionary MHC class II strategies in humans and rhesus macaques: relevance for biomedical studies. Immunol Rev 183:76–85 doi:10.1034/j.1600-065x.2001.1830106.x

Doxiadis GGM, Van Der Wiel MKH, Brok HPM, Groot NG, Otting N, 't Hart BA et al (2006) Reactivation by exon shuffling of a conserved *HLA-DR3*-like pseudogene segment in a new world primate species. Proc Natl Acad Sci U S A 103:5864–5868 doi:10.1073/pnas.0600643103

Figueroa F, Gutknecht J, Tichy H, Klein J (1990) Class II MHC genes in rodent evolution. Immunol Rev 113:27–46 doi:10.1111/j.1600-065X.1990.tb00035.x

Froeschke G, Sommer S (2005) MHC class II *DRB* variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the southern Kalahari. Mol Biol Evol 22:1254–1259 doi:10.1093/molbev/msi112

Gogolevsky KP, Kramerov DA (2006) Short interspersed elements (SINEs) of the Geomyoidea superfamily rodents. Gene 373:67–74 doi:10.1016/j.gene.2006.01.007

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nuc Acid Sym Ser 41:95–98

Harf R, Sommer S (2005) Association between major histocompatibility complex class II *DRB* alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paeba*, in the southern Kalahari. Mol Ecol 14:85–91 doi:10.1111/j.1365-294X.2004.02402.x

Huchon D, Madsen O, Sibbald M, Ament K, Stanhope MJ, Catzeflis F et al (2002) Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. Mol Biol Evol 19:1053–1065

Huchon D, Chevret P, Jordan U, Kilpatrick CW, Ranwez V, Jenkins PD et al (2007) Multiple molecular evidences for a living mammalian fossil. Proc Natl Acad Sci U S A 104:7495–7499 doi:10.1073/pnas.0701289104

Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 32:415–435 doi:10.1146/annurev.genet.32.1.415

IMGT/HLA Database. release 2.19.0 (October 2007). April 15, 2008

Jukes TH, Cantor CR (1969) Evolution of protein molecules. Academic, New York

Kaufman J, Salomonsen J, Flajnik MF (1994) Evolutionary conservation of MHC class I and class II molecules—different yet the same. Semin Immunol 6:411–424 doi:10.1006/smim.1994.1050

Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. Immunogenetics 56:683–695 doi:10.1007/s00251-004-0717-7

Kindt TJ, Singer DS (1987) Class I major histocombatibility complex genes in vertebrate species: what is the common denominator? Immunol Res 6:57–66 doi:10.1007/BF02918104

Klein J (1986) Natural history of the major histocompatibility complex. Wiley, New York

Klein J (1987) Origin of major histocompatibility complex polymorphism—the transspecies hypothesis. Hum Immunol 19:155–162 doi:10.1016/0198-8859(87)90066-8

Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER et al (1990) Nomenclature for the major histocompatibility complexes of different species—a proposal. Immunogenetics 31:217–219

Li WH, Ellsworth DL, Krushkal J, Chang BHJ, Hewett-Emmett D (1996) Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. Mol Phyl Evol 5:182–187 doi:10.1006/mpev.1996.0012

Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. Bioinformatics 21:260–262 doi:10.1093/bioinformatics/bth490

Maynard Smith J (1992) Analyzing the mosaic structure of genes. J Mol Evol 34:126–129

Meyer-Lucht Y, Sommer S (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). Mol Ecol 14:2233–2243 doi:10.1111/j.1365-294X.2005.02557.x

Milinski M (2006) The major histocompatibility complex, sexual selection, and mate choice. Annu Rev Ecol Evol Syst 37:159–186 doi:10.1146/annurev.ecolsys.37.091305.110242

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

Nigenda-Morales S, Flores-Ramirez S, Urban-R J, Vazquez-Juarez R (2008) MHC *DQB1* polymorphism in the Gulf of California fin whale (*Balaenoptera physalus*) population. J Hered 99:14–21 doi:10.1093/jhered/esm087

Nizetic D, Figueroa F, Dembic Z, Nevo E, Klein J (1987) Major histocompatibility complex gene organization in the mole rat *Spalax ehrenbergi*—evidence for transfer of function between class II genes. Proc Natl Acad Sci U S A 84:5828–5832 doi:10.1073/pnas.84.16.5828

Oliver MK, Piertney SB (2006) Isolation and characterization of a MHC class II *DRB* locus in the European water vole (*Arvicola terrestris*). Immunogenetics 58:390–395 doi:10.1007/s00251-006-0121-6

Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. Heredity 96:7–21

Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. Syst Biol 50:580–601 doi:10.1080/106351501750435121

Reusch TBG, Langefors A (2005) Inter- and intralocus recombination drive MHC class II*B* gene diversification in a teleost, the threespined stickleback *Gasterosteus aculeatus*. J Mol Evol 61:531–541 doi:10.1007/s00239-004-0340-0

Richman AD, Herrera LG, Nash D (2002) Characterization of *Peromyscus* MHC class II beta sequences by ligation-anchored RT-PCR and denaturing gradient gel electrophoresis. Eur J Immunogenet 29:213–217 doi:10.1046/j.1365-2370.2002.00290.x

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 doi:10.1017/S0016672303006347

Robinson J, Waller MJ, Parham P, De Groot N, Bontrop R, Kennedy LJ et al (2003) IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res 31:311–314 doi:10.1093/nar/gkg070

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DNA Sequence Polymorphism v.4.20.2. Free software distributed by author (http://www.ub.es/dnasp)

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 doi:10.1038/sj.hdy.6800892

Schaschl H, Wegner KM (2007) Contrasting mode of evolution between the MHC class I genomic region and class II region in the three-spined stickleback (*Gasterosteus aculeatus* L.; Gasterosteidae: Teleostei). Immunogenetics 59:295–304 doi:10.1007/s00251-007-0192-z

Smulders MJM, Snoek LB, Booy G, Vosman B (2003) Complete loss of MHC genetic diversity in the common hamster (*Cricetus cricetus*) population in The Netherlands: consequences for conservation strategies. Con Gen 4:441–451 doi:10.1023/A:1024767114707

Strand T, Westerdahl H, Hoeglund 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 doi:10.1007/s00251-007-0234-6

Swofford DL (1998) PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer, Sunderland

Triant DA, DeWoody JA (2006) Accelerated molecular evolution in *Microtus* (Rodentia) as assessed via complete mitochondrial genome sequences. Genetica 128:95–108 doi:10.1007/s10709-005-5538-6

von Salome J, Gyllensten U, Bergström TF (2007) Full-length analysis of the *HLA-DRB1* locus suggests a recent origin of alleles. Immunogenetics 59:261–271 doi:10.1007/s00251-007-0196-8

Waser PM, Busch JD, McCormick CR, DeWoody JA (2006) Parentage analysis detects cryptic pre-capture dispersal in a philopatric rodent. Mol Ecol 15:1929–1937 doi:10.1111/j.1365-294X.2006.02893.x

Wilson and Reeder (2005) (eds) Mammal species of the world: a taxonomic and geographic reference, 3rd edn. The Johns Hopkins University Press, Baltimore

Worley K, Carey J, Veitch A, Coltman DW (2006) Detecting the signature of selection on immune genes in highly structured populations of wild sheep (*Ovis dalli*). Mol Ecol 15:623–637 doi:10.1111/j.1365-294X.2006.02829.x

Yauk CL, Bois PR, Jeffreys AJ (2003) High-resolution sperm typing of meiotic recombination in the mouse MHC *E*-beta gene. EMBO J 22:1389–1397 doi:10.1093/emboj/cdg136

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. Genet Res 13:1169–1179 doi:10.1101/gr.976103