Acta Theriologica 52 (3): 227–235, 2007.
PL ISSN 0001–7051

Sequence diversity of MHC class II DRB genes in the bank vole *Myodes glareolus*

Wiesław BABIK and Jacek RADWAN

Babik W. and Radwan J. 2007. Sequence diversity of MHC class II DRB genes in the bank vole *Myodes glareolus*. Acta Theriologica 52: 227–235.

In recent years, the bank vole *Myodes glareolus* (Schreber, 1780) has emerged as a model system for parasitological, behavioural and ecological studies and seems ideally suited to address questions concerning the importance of MHC variation at individual and population levels. Here, we provide the first extensive survey of sequence variation in the MHC class II DRB genes in this species. Among 34 analysed voles we found 15 unique sequences, representing most likely two loci, at least one of them expressed. Despite very high overall sequence divergence, particularly in the Antigen Binding Sites (ABS), we detected signatures of positive selection that has been acting on DRB in the bank vole. Phylogenetic analysis demonstrated that the bank vole DRB alleles do not form a monophyletic group but are intermingled with other rodent alleles that is consistent with long-term persistence of ancient allelic lineages maintained through balancing selection. Our sequence data will forward the design of efficient genotyping methods, which will permit testing hypotheses pertaining to the ecological causes and consequences of MHC variation in the bank vole.

Helmholtz Centre for Environmental Research – UFZ, Department of Community Ecology, Theodor-Lieser-Str. 4, D-06120 Halle/Saale, Germany, e-mail: wieslaw.babik@ufz.de (WB); Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland (JR, WB)

Key words: Major Histocompatibility Complex, polymorphism, rodents, selection, recombination

Introduction

The Major Histocompatibility Complex (MHC) contains the most polymorphic genes described in vertebrates (Garrigan and Hedrick 2003). Polymorphism is concentrated in genes coding for proteins that present antigens from intracellular (MHC I) and extracellular (MHC II) pathogens to lymphocytes (Janeway *et al.* 1999, Sommer 2005). High numbers of alleles are thought to be maintained in natural populations because of the selective pressure induced by fast evolving pathogens. Pathogens readily adapt to common host genotypes, triggering frequency-dependent selection which may prevent the loss of rare alleles (Snell 1968, Borghans *et al.* 2004). Heterozygote advantage, enabling heterozygous individuals to recognise a wider range of pathogen-derived antigens, is another mechanism that

can contribute to the maintenance of high polymorphism at MHC loci (Doherty and Zinkernagel 1975). The benefits of rare alleles and heterozygosity at MHC loci have been shown in recent research on humans and other species (eg Thursz *et al.* 1997, Paterson *et al.* 1998, Penn *et al.* 2002, Trachtenberg *et al.* 2003). Relative significance of mechanisms maintaining extremely high MHC variation remains a matter of debate (reviewed in Sommer 2005, Milinski 2006).

However, both mechanisms will provide an advantage to rare and new alleles, as well as favor non-synonymous mutations thus generating the signal of positive selection. Indeed, such signal is often detected in MHC sequences, especially in the antigen-binding sites (ABS) (reviewed in Garrigan and Hedrick 2003, Sommer 2005, Piertney and Oliver 2006). Selection on MHC genes seems to be able to maintain polymorphism for a very long time, resulting in trans-species polymorphism (Klein 1987, Edwards *et al.* 1997).

At the intra-individual level, MHC variation can be increased by duplication of MHC loci; for example, more than one DRB locus is present in several mammalian species, and haplotypes may differ in the number of DRB loci (Fraser and Bailey 1996, Khazand *et al.* 1999, Kennedy *et al.* 2002). While gene duplication may be expected to facilitate recognition of pathogen-derived antigens by expression of a wider range of MHC proteins, other processes, such as increased risk of auto-immunoaggression, may limit the optimal number of MHC proteins expressed by individuals (Borghans *et al.* 2004, Milinski 2006)

In recent years, the bank vole *Myodes glareolus* (Schreber, 1780), until recently known as *Clethrionomys glareolus*, has emerged as a model system for parasitological, behavioural and ecological studies (Bajer *et al.* 2001, Behnke *et al.* 2001, Koskela *et al.* 2004, Labocha *et al.* 2004, Radwan *et al.* 2004, Karbowiak *et al.* 2005). As the species is extensively studied both in nature and in the lab, correlational data collected in the field can be further scrutinized in controlled laboratory settings. Thus, the species seems ideally suited to address questions concerning the importance of MHC variation at individual and population levels. So far, only MHC DQA sequences have been described in the bank vole. Bryja *et al.* (2006) found 7 unique MHC DQA sequences, distributed among 2 expressed loci, in 7 individuals. Here, we provide the first extensive survey of sequence variation in another class II MHC gene, the DRB. It is the most polymorphic and most widely studied class II gene in mammals (Garrigan and Hedrick 2003, Piertney and Oliver 2006). Sequence variation was tested for signs of positive selection acting in the past and was used to determine the phylogenetic relationship with DRB genes of other species. Characterisation of the sequences will facilitate the design of genotyping methods permitting the interpretation of the extensive data on extracellular parasites that this species abundantly hosts (Bajer *et al.* 2001, Behnke *et al.* 2001, Karbowiak *et al.* 2005).

Material and methods

We analysed variation in the DRB second exon in 34 bank voles. The samples came from three populations: Mazury Lake district (NE Poland, $n = 15$), Bieszczady Foothills (SE Poland, $n = 9$) and Puszcza Niepołomicka (S Poland, $n = 10$). Distances between populations ranged from ca 200 (Bieszczady – Puszcza Niepołomicka) to ca 600 km (Puszcza Niepołomicka – Mazury).

DNA was extracted using the Genomic Mini kit (A & A Biotechnology) from fresh, frozen, or alcohol-preserved tissues. We used the forward (5'GAGTGTCATTTCTACAAC GGGA3') and reverse (5'CTCTCCGCGGCACAAAGGAA3') primers described by Smulders *et al.* (2003), to amplify a 203 bp (excluding primers) fragment of the 2nd exon of the DRB gene. Amplifications were carried out in 20 μl volumes, the reaction mixtures consisted of: 2 μl of 10 × PCR buffer with $(NH_4)_2SO_4$ (Fermentas), 1.5 mM $MgCl_2$, 1 μM of each the forward and reverse primers, 0.2 mM of each dNTP and 1 U of *Taq* polymerase (Fermentas). The cycling scheme was as follows: 94°C for 3 min, followed by 35 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 30 s and a final extension step at 72°C for 3 min. PCR products were cloned using the pGEM T-Easy vector system (Promega). Recombinant clones were detected by blue/white screening and plasmid DNA minipreps were prepared with the Plasmid Mini kit (A & A Biotechnology). Multiple clones (7.3 per individual on average) containing inserts were sequenced using standard T7 or SP6 primers. Sequencing reactions were performed with the Big Dye 3.1 sequencing kit (ABI) and their products were separated on an ABI 3100 Genetic Analyser. Sequences were checked and aligned in SeqScape 2.1.1 (ABI). Some individuals with putative rare alleles were also cloned from a second, independent PCR. As PCR-

and cloning-generated (resulting from bacterial mismatch repair of cloned PCR-generated heteroduplexes) artefacts are common in analyses of polymorphic genes, particularly in the case of MHC genes (for a discussion see Jarvi *et al*. (2004)) we employed conservative criteria for allele identification. Throughout the text we use the term 'allele' to designate unique sequences, however it should be kept in mind that these unique sequences represent more than one locus. We treated sequences obtained in clones derived from at least two independent PCR reactions as confirmed alleles. Clones containing unconfirmed, according to the above rule, sequences differing by single substitutions from a known allele were regarded as representing this allele. Other sequence variants, which may have represented polymerase errors and PCR- or cloning-generated recombinants, were excluded from the analyses. In designating allele names we followed the nomenclatural rules set by Klein *et al.* (1990).

In order to confirm the expression of the assayed loci, mRNA was obtained from spleens of two voles. The samples of the spleen were preserved in RNAlater RNA stabilization reagent (Qiagen). RNA was extracted with the RNeasy kit (Qiagen) from about 20 mg of homogenized tissue sample. The samples were digested on columns with DNAse (Qiagen) for 20 minutes. cDNA was obtained using Omniscript Reverse Transcriptase (Qiagen) in 20 μl reaction containing 6 μl template RNA, 2 μl Oligo(dT) 12–18 primer (0.5 μg/μl, Invitorgen), 1 μl RNase inhibitor (10 u/μl) and 1 μl RT and 6 μl of RNase-free water. The reaction was incubated at 37°C for 60 min. DRB was amplified from cDNA as described above for genomic DNA with DNAase-treated RNA used as negative control. PCR products were cloned and sequenced as described above.

The average pairwise nucleotide distances (Kimura 2-parameter model, K2P), Poisson-corrected amino acid distances as well as the average rates of synonymous (dS) and nonsynonymous (dN) substitutions per site, using the Nei–Gojobori method with the Jukes–Cantor correction for multiple substitutions, were computed in MEGA3 (Kumar *et al.* 2004). Standard errors of the estimates were obtained through 1000 bootstrap replicates.

We tested for signs of recombination in our dataset with three methods. The software RDP (Martin and Rybicki 2000) was applied for the analysis using GENECONV (Padidam *et al.* 1999) and MaxChi2 (Maynard Smith 1992) methods. These two methods performed best in an assessment of fourteen recombination detection methods (Posada 2002). In the third method, Genetic Algorithm Recombination Detection (GARD) (Pond *et al.* 2006), a web-based routine (http://www.datamonkey.org/GARD/) was used for computations.

We tested if positive selection shaped the evolution of the assayed fragment of the DRB second exon using two approaches. First, the one-tailed *Z*-test of positive selection, as implemented in MEGA3, compared the rates of synonymous vs. nonsynonymous substitutions at all amino acid positions, antigen binding sites (ABS) and non-ABS. The location of the putative ABS and non-ABS was inferred from the human MHC II molecule structure (Brown *et al.* 1993). Additionally, we compared ABS vs. non-ABS sites with respect to rates of non-synonymous and synonymous substitutions using two-tailed *Z*-test.

Second, we tested for positive selection by comparing likelihoods of three codon-based models of sequence evolution available in PAML 3.15 (Yang 1997): (1) M0: one ω (dN/dS) ratio, (2) M1a: nearly neutral, p_0 sites under purifying selection ($\omega_0 < 1$) and the rest neutral ($\omega_1 = 1$), (3) M2a positive selection, p_0 sites under purifying selection ($\omega_0 < 1$), p_1 neutral ($\omega_1 = 1$) and the rest under positive selection ($\omega_2 > 1$). The best fitting model was chosen on the basis of the likelihood-ratio test (Yang 1997) and the lowest value of the Akaike Information Criterion (Posada and Buckley 2004). Positively selected codons under M2a model were identified through the Bayes empirical Bayes procedure (Zhang *et al.* 2005).

We constructed phylogenetic trees in order to elucidate the position of the bank vole DRB alleles among the alleles of other rodent species and to test if the transspecific mode of sequence evolution applies to the bank vole alleles. The most similar (closest BLAST matches) non rodent sequences, *Alces alces* and *Macaca mulatta* were used to root the trees. The best fitting model of sequence evolution was chosen using ModelTest 3.5 (Posada and Crandall 1998). The Akaike Information Criterion revealed the best fit of the Transversional Model, a simplified variant of the GTR model (see Tavaré 1986), with non-zero proportion of invariable sites and gamma-distributed rate variation (TVM + I + Γ) with the following parameter values: nucleotide frequencies 0.200 (A) 0.295 (C) 0.324 (G) 0.181 (T), proportion of invariable sites I = 0.238 and gamma shape parameter α = 1.060. A neighbor joining tree was constructed in PAUP from the matrix of distances computed according to the model. The robustness of the tree topology was tested with 1000 bootstrap replicates. Another tree was constructed under the Bayesian approach with MRBAYES 3.1 (Ronquist and Huelsenbeck 2003). The likelihood settings corresponded to the general time-reversible (GTR) model + I + Γ with the parameter values to be estimated from the data; priors were set to default values. Four Metropolis Coupled Monte Carlo Markov Chains (three of them "heated", temperature = 0.20) were run in two runs for $1*10^6$ generations each and sampled every 1000 generations. The first ten percent of trees were discarded as burn-in, resulting in 1800 sampled trees. To calculate the posterior probability of each bipartition, the majority-rule consensus tree was computed from these 1800 sampled trees.

Results

We detected fifteen DRB alleles among 34 assayed bank voles. Sequences of all alleles were deposited in GenBank (Accesion Numbers EF469837-51). Three unique DRB sequences per individual were found in nine individuals and four sequences were found in two voles. This indicates the presence of at least two DRB loci. Analysis of PCR products obtained from spleen mRNA confirmed that at least one locus is

expressed. In one individual, whose genomic DNA (11 clones sequenced) contained alleles *Mygl-DRB* 01*,**05*,**06* and **14*, in cDNA (6 clones) we found alleles **6* and **14*; in the other individual, we detected alleles **3*, **5*, **7* and **10* in genomic DNA (9 clones), but only allele **7* in cDNA (13 clones). Because neither individual genotypes nor phylogenetic relationships unequivocally assigned individual alleles to a given locus, in the following presentation we consider them as collectively representing the DRB loci in the bank vole.

Eighty seven of 203 (42.9%) nucleotide and forty two of 67 (62.7%) amino acid positions were variable (Table 1). No insertions/deletions or stop codons were detected. The number of nucleotide differences between pairs of alleles ranged from 1 (*Mygl-DRB*02* vs. *Mygl-DRB*04*) to 42 (*Mygl-DRB*01* vs. *Mygl-DRB*15*) and the number of amino acid differences ranged from 0 to 26 in the respective allele pairs. The average pairwise K2P distances and Poisson-corrected amino acid distances computed for all sites were 15.9 ± (SE) 1.8% and 28.1 ± 4.0% respectively. Both nucleotide and amino acid distances were much higher at ABS (27.6 ± 5.0 and 58.2 ± 12.1%) than at non-ABS (11.8 ± 1.8 and 18.9 ± 3.6%).

The frequencies of individuals carrying particular alleles in the three sampled populations are shown in Table 2. Differentiation in allelic composition between populations was notable and manifested in the varying frequencies of individuals carrying the three most common alleles (χ^2 = 11.5, df = 4, p = 0.021) and in a relatively high number of "private" rare alleles detected only in a single population (Table 2).

There was no strict agreement among methods regarding the effect of recombination on the evolution of DRB sequences in our dataset. Whereas the GENECONV and GARD methods did not detect any recombination events, MaxChi2 suggested recombination in case of alleles *Mygl-DRB*03* (p = 0.045) *Mygl-DRB*15* (p = 0.013). However, in both cases only one parental sequence could be identified in the dataset.

The rate of nonsynonymous to synonymous substitutions (dN/dS), as revealed by Z tests of positive selection, was not significantly elevated either overall, or at the ABS sites (Table 3). However, it should be noted that dN was over three times higher at ABS sites than at non-ABS

Table 1. Amino acid sequences of MHC DRB alleles detected in the bank vole. Dots indicate identity to the reference sequence. Putative Antigen Binding Sites (ABS) (Brown *et al.* 1993) are shaded. Amino acid positions identified as being under positive selection, as identified by the Bayes empirical Bayes procedure are marked with asterisks (* – posterior probability (PP) > 95, ** – PP > 99%).

```
                        10        20        30        40        50        60
                ....|....|....|....|....|....|....|....|....|....|....|....|....|..
                                                                    *           *
                                                                                *
Mygl-DRB*01     QRVRFLDRYFYNREEYARFDSDVGEYRVVNELGLPDSEYWNSQEELLEQKRSLVDTYCRYNYQVAEG
Mygl-DRB*02     R...........Q..............A.....R.................................
Mygl-DRB*03     ....Y.Y.DI..Q..FV..........A.T...RR.A.V....K...DNR.AY......H..G.G.S
Mygl-DRB*04     R...........Q..............A.....R.................................
Mygl-DRB*05     ...........................A.K.....................................
Mygl-DRB*06     .H..Y........K..V..........A.T...R..AK.....K.......AK......H..G.G.S
Mygl-DRB*07     .H.L.VE.FI......V..........AL....R..A.....RK.I.DNR.AA.........RIM.S
Mygl-DRB*08     ............Q..............A.K.....................................
Mygl-DRB*09     .H..............V..........A.T...R..AK....RK.I.....AA..AF.....G.V.S
Mygl-DRB*10     ...............LV.....I..H.A.....RR.A......KDI..R..AY...F..H..G.G.S
Mygl-DRB*11     ....Y.Y.DI..Q..FV.........HA.T...RR.A.V....K...DNR.AY......H..G.G.S
Mygl-DRB*12     ..........H.Q..............A.......................................
Mygl-DRB*13     R...........Q..............A.....R......I..........................
Mygl-DRB*14     ...Q.VE..IS.Q..NV..........A.T...R..A......K.I..DA.AA.........G.V.S
Mygl-DRB*15     ....Y.F.DI..Q..HV........F.A.T...R..A......KDI..RT.AET..V..H..GIV.S
```

Table 2. Frequencies of individuals with particular DRB alleles detected in the three populations sampled. Maz – Mazury Lake District, Bie – Bieszczady Foothills, Nie – Puszcza Niepołomicka.

Allele	Total sample (n = 34)	Maz (n = 15)	Bie (n = 9)	Nie (n = 10)
*Mygl-DRB*01*	0.19	0.23	0.06	0.26
*Mygl-DRB*02*	0.16	0.19	0.33	0.00
*Mygl-DRB*03*	0.15	0.15	0.06	0.22
*Mygl-DRB*04*	0.08	0.08	0.00	0.13
*Mygl-DRB*05*	0.06	0.08	0.00	0.09
*Mygl-DRB*06*	0.04	0.00	0.00	0.13
*Mygl-DRB*07*	0.04	0.04	0.06	0.04
*Mygl-DRB*08*	0.04	0.00	0.17	0.00
*Mygl-DRB*09*	0.04	0.12	0.00	0.00
*Mygl-DRB*10*	0.04	0.00	0.11	0.04
*Mygl-DRB*11*	0.04	0.04	0.11	0.00
*Mygl-DRB*12*	0.03	0.04	0.06	0.00
*Mygl-DRB*13*	0.03	0.04	0.00	0.04
*Mygl-DRB*14*	0.02	0.00	0.00	0.04
*Mygl-DRB*15*	0.02	0.00	0.06	0.00

Table 3. The average rates of nonsynonymous substitutions per nonsynonymous site (d_N), synonymous substitutions per synonymous sites (d_S), with standard errors obtained through 1000 bootstrap replicates in parentheses, and the results of the Z-test of positive selection. d_S and d_N values are given as percentages per site.

Sites	d_N	d_S	Z	p
All	14.3 (2.3)	19.9 (4.6)	1.162	1.000
ABS	30.1 (7.0)	20.0 (9.3)	1.869	1.000
non-ABS	8.9 (1.9)	20.1 (5.4)	1.038	0.151

sites (Z = 2.92, p = 0.004), with dS being very similar (Z = 0.01, p = 0.99). Also, the average K2P nucleotide distances and Poisson-corrected amino acid distances between alleles were about three times higher at ABS than at non-ABS sites. The model of codon evolution assuming that a fraction of codons has been affected by positive selection fitted the data substantially better than the uniform ω and nearly neutral models (χ^2 = 11.4, df = 2 , p = 0.0033 and δAIC = 7.4 in comparison between models M1a and M2a). However, only a small fraction of codons were identified as evolving under positive selection. Accordingly, the Bayes empirical procedure identified only two codons as positively selected (posterior probability > 95%), both at ABS sites.

Both the NJ and Bayesian trees were poorly resolved, with only several well supported groupings. General relationships among rodent alleles remain unresolved with several notable exceptions (Fig. 1). Three well supported groups were distinguished among the bank vole alleles, one, clustering seven similar alleles, received 100% bootstrap support (BS) and Bayesian posterior probability (PP), the others grouped two bank vole alleles each: *Mygl-DRB*03* and *Mygl-DRB*11* with 100% BS and PP, and *Mygl-DRB*07* and *Mygl-DRB*14* with 86% BS and 99% PP. It is evident from the phylogenetic analysis that the bank vole alleles do not form a monophyletic group, but are intermingled with other rodent alleles.

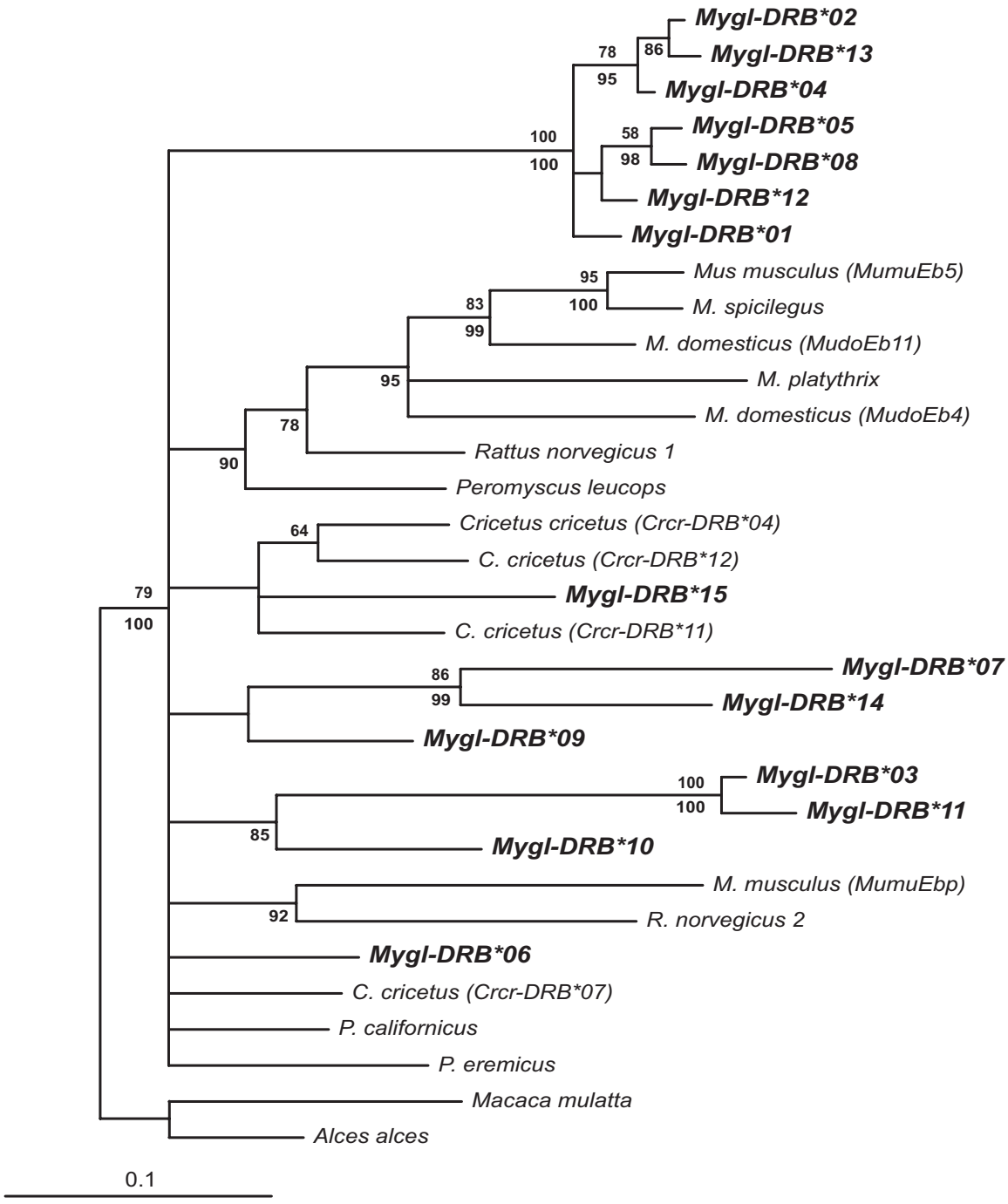

Fig 1. Phylogenetic relationships of the bank vole MHC II DRB alleles (bold) with a representative set of other rodent alleles. 50% majority rule tree from the Bayesian analysis is shown. Bootstrap values above 50% are shown above respective branches, Bayesian posterior probability above 70% are shown below branches. Other rodent sequences included, with Gen Bank Accession Numbers in parentheses: *Cricetus cricetus*: *Crcr-DRB*04* (AJ490314), *Crcr-DRB*07* (AJ490317), *Crcr-DRB*11* (AJ490321), *Crcr-DRB*12* (AJ490322); *Mus musculus musculus*: *MumuEb5* (U88933) *MumuEbp* (AY740480); *M. m. domestcius*: *MudoEb4* (U88913) *MudoEb11* (U88920); *Mus platythrix* (U88949); *Mus spicilegus* (AJ634272); *Peromyscus californicus* (AY219808), *Peromyscus eremicus* (AY219815), *Peromyscus leucopus* (AY219804), *Rattus norvegicus*: *1* (AJ003232) *2* (M57819). Tree was rooted using *Macaca mulatta* (AY487258) and *Alces alces* (X82398) DRB sequences.

Discussion

Our primers amplified alleles from at least two loci, as shown by the detection of three or four alleles in several individuals. The presence of more than one DRB locus is a common phenomenon among mammals (eg Khazand *et al.* 1999, Kennedy *et al.* 2002, Harf and Sommer 2005). However, in the only other investigated member of subfamily Arvicolinae, the water vole *Arvicola terrestris* there seems to be only a single DRB locus (Oliver and Piertney 2006). We have found that at least one of DRB loci is expressed in the bank vole. The lack of indels or stop codons in any of the sequences suggests that we have not amplified a pseudogene, although one of the loci may be expressed at a lower level and thus be difficult to detect in a limited number of clones.

Inconclusive outcomes of the tests of recombination indicate that, even if present, recombination has not been a major force in the evolution of the assayed DRB fragment in the bank vole. A similar conclusion was reached by Bryja *et al.* (2006) for DQA.

We have found evidence that positive selection has contributed to the evolution of the bank vole DRB genes, and explicitly identified two codons, located at putative ABS, as being under positive selection. This indicates that the polymorphism of DRB genes in the bank vole has been driven by the peptide-binding properties of their products, and thus, most likely, by their ability to bind pathogen-derived antigens. However, unlike a number of other species (reviewed in Bernatchez and Landry 2003, Sommer 2005, Piertney and Oliver 2006), we have not found a significant excess of non-synonymous substitutions at ABS sites. This is an exception rather than a rule (Bernatchez and Landry 2003), but could be partly accounted for by the high divergence of allelic sequences and saturation of non-synonymous sites (Richman *et al.* 2001). Indeed, whereas synonymous substitution rates were high and similar at ABS and non-ABS sites, the non-synonymous substitution rate was about three times higher at ABS sites. Additionally, we cannot exclude the possibility of differences in ABS location between rodent and human MHC II molecules. Overall, our results indicate that codons associated with the specificity of binding antigens have been under positive selection.

The high sequence divergence is consistent with the long persistence of allelic lineages, as supported by trans-species polymorphisms observed in our phylogenetic analysis. Trans-species polymophisms were found in other rodents as well (Edwards *et al.* 1997, Smulders *et al.* 2003, Musolf *et al.* 2004). Bank voles also share DQA allelic lineages with other vole species (Bryja *et al.* 2006). Relatively short sequence length together with very high sequence variation obviously affected robustness of the trees obtained in our phylogenetic analyses as evidenced by generally low bootstrap and posterior values. This is however the result of a "natural" limitation – the length of the DRB 2nd exon.

The frequencies of even the three most common alleles differed substantially between assayed populations. Some rarer alleles were detected in only a single population (*Mygl-DRB*06,* 08,* 09,*14,*15*). This differentiation may reflect sampling error, differences in local selective pressures or may have historical reasons. The present-day area of Poland was colonized from at least two refugial areas after the last glacial maximum (Deffontaine *et al.* 2005, Kotlik *et al.* 2006). The differences in MHC allelic composition that may have emerged during period(s) of isolation in refugial populations, probably contribute both to the differentiation between populations and to overall high DRB diversity in bank voles from Poland.

Our sequence data will forward the design of efficient genotyping methods, such as Single Strand Conformational Polymorphism (SSCP) or Reference Strand Conformational Analysis (RSCA), which can first be tested using clones containing MHC DRB inserts of known sequences. This will permit testing hypotheses pertaining to the ecological causes and consequences of MHC variation in this thoroughly studied species. For example, it has been found that bank vole populations inhabiting neighbouring locations differ in composition of para-

sites (Bajer *et al.* 2001, Behnke *et al.* 2001). It would now be possible to examine if this differences are reflected in MHC variation and allelic composition.

Acknowledgements: We thank A. Kloch and Z. Boratyński for samples, B. Matuła for technical assistance and M. Pabijan for his comments. This work was supported by a grant from Polish Ministry of Science and Higher Education (0494/P04/2005/28) and European Community Centre of Excellence IBAES.

References

Bajer A., Pawełczyk A., Behnke J. M., Gilbert F. S. and Siński E. 2001. Factors affecting the component community structure of haemoparasites in bank voles (*Clethrionomys glareolus*) from the Mazury Lake District region of Poland. Parasitology 122: 43–54.

Behnke J. M., Bajer A., Siński E. and Wakelin D. 2001. Interactions involving intestinal nematodes of rodents in experimental and field studies. Parasitology 122: S39–S49.

Bernatchez L. and Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology 16: 363–377.

Borghans J. A. M., Beltman J. B. and De Boer R. J. 2004. MHC polymorphism under host-pathogen coevolution. Immunogenetics 55: 732–739.

Brown J. H., Jardetzky T. S., Gorga J. C., Stern L. J., Urban R. G., Strominger J. L. and Wiley D. C. 1993. 3-Dimensional structure of the human class-II histocompatibility antigen HLA-DR1. Nature 364: 33–39.

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

Deffontaine V., Libois R., Kotlik P., Sommer R., Nieberding C., Paradis E., Searle J. B. and Michaux J. 2005. Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). Molecular Ecology 14: 1727–1739.

Doherty P. C. and Zinkernagel R. M. 1975. Enhanced immunological surveillance in mice heterozygous at H-2 gene complex. Nature 256: 50–52.

Edwards S. V., Chesnut K., Satta Y. and Wakeland E. K. 1997. Ancestral polymorphism of MHC class II genes in mice: Implications for balancing selection and the mammalian molecular clock. Genetics 146: 655–668.

Fraser D. G. and Bailey E. 1996. Demonstration of three DRB loci in a domestic horse family. Immunogenetics 44: 441.

Garrigan D. and Hedrick P. W. 2003. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. Evolution 57: 1707–1722.

Harf R. and 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. Molecular Ecology 14: 85–91.

Janeway C. A., Travers P., Walport D. and Capra J. D. 1999. Immunobiology: The Immune System in Health and Disease. Current Biology Publications, London: 1–635.

Jarvi S. I., Tarr C. L., McIntosh C. E., Atkinson C. T. and Fleischer R. C. 2004. Natural selection of the major histocompatibility complex (MHC) in Hawaiian honeycreepers (Drepanidinae). Molecular Ecology 13: 2157–2168.

Karbowiak G., Rychlik L., Nowakowski W. and Wita I. 2005. Natural infections with blood parasites on the borderland of boreal and temperate forest zones. Acta Theriologica 50: 31–42.

Kennedy L. J., Ryvar R., Gaskell R. M., Addie D. D., Willoughby K., Carter S. D., Thomson W., Ollier W. E. R. and Radford A. D. 2002. Sequence analysis of MHC DRB alleles in domestic cats from the United Kingdom. Immunogenetics 54: 348–352.

Khazand M., Peiberg C., Nagy M. and Sauermann U. 1999. MHC-DQ-DRB haplotype analysis in the rhesus macaque: evidence for a number of different haplotypes displaying a low allelic polymorphism. Tissue Antigens 54: 615–624.

Klein J. 1987. Origin of Major Histocompatibility Complex polymorphism – the transspecies hypothesis. Human Immunology 19: 155–162.

Klein J., Bontrop R. E., Dawkins R. L., Erlich H. A., Gyllensten U. B., Heise E. R., Jones P. P., Parham P., Wakeland E. K. and Watkins D. I. 1990. Nomenclature for the Major Histocompatibility Complexes of different species – a proposal. Immunogenetics 31: 217–219.

Koskela E., Huitu O., Koivula M., Korpimaki E. and Mappes T. 2004. Sex-biased maternal investment in voles: importance of environmental conditions. Proceedings of the Royal Society of London Series B-Biological Sciences 271: 1385–1391.

Kotlik P., Deffontaine V., Mascheretti S., Zima J., Michaux J. R. and Searle J. B. 2006. A northern glacial refugium for bank voles (*Clethrionomys glareolus*). Proceedings of the National Academy of Sciences of the United States of America 103: 14860–14864.

Kumar S., Tamura K. and Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150–163.

Labocha M. K., Sadowska E. T., Baliga K., Semer A. K. and Koteja P. 2004. Individual variation and repeatability of basal metabolism in the bank vole, *Clethrionomys glareolus*. Proceedings of the Royal Society of London Series B-Biological Sciences 271: 367–372.

Martin D. and Rybicki E. 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16: 562–563.

Maynard Smith J. 1992. Analyzing the Mosaic Structure of Genes. Journal of Molecular Evolution 34: 126–129.

Milinski M. 2006. The major histocompatibility complex, sexual selection, and mate choice. Annual Review of Ecology Evolution and Systematics 37: 159–186.

Musolf K., Meyer-Lucht Y. and Sommer S. 2004. Evolution of MHC-DRB class II polymorphism in the genus *Apodemus* and a comparison of DRB sequences within the family Muridae (Mammalia: Rodentia). Immunogenetics 56: 420–426.

Oliver M. K. and Piertney S. B. 2006. Isolation and characterization of a MHC class II DRB locus in the European water vole (*Arvicola terrestris*). Immunogenetics 58: 390–395.

Padidam M., Sawyer S. and Fauquet C. M. 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265: 218–225.

Paterson S., Wilson K. and Pemberton J. M. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries* L.). Proceedings of the National Academy of Sciences of the United States of America 95: 3714–3719.

Penn D. J., Damjanovich K. and Potts W. K. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. Proceedings of the National Academy of Sciences of the United States of America 99: 11260–11264.

Piertney S. B. and Oliver M. K. 2006. The evolutionary ecology of the major histocompatibility complex. Heredity 96: 7–21.

Pond S. L. K., Posada D., Gravenor M. B., Woelk C. H. and Frost S. D. W. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Molecular Biology and Evolution 23: 1891–1901.

Posada D. 2002. Evaluation of methods for detecting recombination from DNA sequences: Empirical data. Molecular Biology and Evolution 19: 708–717.

Posada D. and Buckley T. R. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology 53: 793–808.

Posada D. and Crandall K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817–818.

Radwan J., Kruczek M., Labocha M. K., Grabiec K. and Koteja P. 2004. Contest winning and metabolic competence in male bank voles *Clethrionomys glareolus*. Behaviour 141: 343–354.

Richman A. D., Herrera L. G. and Nash D. 2001. MHC class II beta sequence diversity in the deer mouse (*Peromyscus maniculatus*): implications for models of balancing selection. Molecular Ecology 10: 2765–2773.

Ronquist F. and Huelsenbeck J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Smulders M. J. M., Snoek L. B., Booy G. and Vosman B. 2003. Complete loss of MHC genetic diversity in the common hamster (*Cricetus cricetus*) population in the Netherlands. Consequences for conservation strategies. Conservation Genetics 4: 441–451.

Snell G. D. 1968. The H-2 locus of the mouse: observations and speculations concerning its comparative genetics and its polymorphism. Folia Biologica (Prague) 14: 335–358.

Sommer S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology 2: 16.

Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. [In: Some mathematical questions in biology – DNA sequence analysis. R. M. Miura, ed]. American Mathematical Society., Providence: 57–86.

Thursz M. R., Thomas H. C., Greenwood B. M. and Hill A. V. S. 1997. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. Nature Genetics 17: 11–12.

Trachtenberg E., Korber B., Sollars C., Kepler T. B., Hraber P. T., Hayes E., Funkhouser R., Fugate M., Theiler J., Hsu Y. S., Kunstman K., Wu S., Phair J., Erlich H. and Wolinsky S. 2003. Advantage of rare HLA supertype in HIV disease progression. Nature Medicine 9: 928–935.

Yang Z. H. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Computer Applications in the Biosciences 13: 555–556.

Zhang J. Z., Nielsen R. and Yang Z. H. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution 22: 2472–2479.

Received 10 January 2007, accepted 26 March 2007.

Associate editors were Maarit Jaarola and Krzysztof Schmidt.