Molecular Ecology (2008) **17**, 2339–2355 doi: 10.1111/j.1365-294X.2008.03757.x

Contrasting patterns of variation in MHC loci in the Alpine newt

W. BABIK,*‡ M. PABIJAN† and J. RADWAN‡
*Department of Community Ecology, Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Str. 4, 06120 Halle/Saale, Germany, †Department of Comparative Anatomy, Institute of Zoology, Jagiellonian University, Ingardena 6, 30-060 Kraków, Poland, ‡Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland

Abstract

Major histocompatibility complex (MHC) genes are essential in pathogen recognition and triggering an adaptive immune response. Although they are the most polymorphic genes in vertebrates, very little information on MHC variation and patterns of evolution are available for amphibians, a group known to be declining rapidly worldwide. As infectious diseases are invoked in the declines, information on MHC variation should contribute to devising appropriate conservation strategies. In this study, we examined MHC variation in 149 Alpine newts (*Mesotriton alpestris*) from three allopatric population groups in Poland at the northeastern margin of the distribution of this species. The genetic distinctiveness of the population groups has previously been shown by studies of skin graft rejection, allozymes and microsatellites. Two putative expressed MHC II loci with contrasting levels of variation and clear evidence of gene conversion/recombination between them were detected. The *Meal-DAB* locus is highly polymorphic (37 alleles), and shows evidence of historical positive selection for amino acid replacements and substantial geographical differentiation in allelic richness. On the contrary, the *Meal-DBB* locus exhibits low polymorphism (three alleles differing by up to two synonymous substitutions) and a uniform distribution of three alleles among geographical regions. The uniform frequencies of the presumptively neutral *Meal-DBB* alleles may be explained by linkage to *Meal-DAB*. We found differences in allelic richness in *Meal-DAB* between regions, consistent with the hypothesis that genetic drift prevails with increasing distance from glacial refugia. Pseudogene loci appear to have evolved neutrally. The level of *DAB* variation correlated with variation in microsatellite loci, implying that selection and drift interplayed to produce the pattern of MHC variation observed in marginal populations of the Alpine newt.

Keywords: linkage disequilibrium, *Mesotriton alpestris*, MHC pseudogenes, MHC variation, positive selection

Received 23 December 2007; revision accepted 25 February 2008

Introduction

Proteins encoded by genes of the major histocompatibility complex (MHC) play an essential role in the adaptive immune response of vertebrates (Klein 1986). Class I and II proteins are located on the cell surface and present antigens of intra- and extracellular origin to lymphocytes, triggering an adaptive immune response (Kumanovics *et al*. 2003). MHC genes are the most polymorphic genes in vertebrates (Bernatchez & Landry 2003). An excess of nonsynonymous substitutions, observed mainly in codons involved in antigen binding (antigen binding sites, ABS), is a commonly observed hallmark of positive selection acting on the MHC (Garrigan & Hedrick 2003; Sommer 2005; Piertney & Oliver 2006). Alleles are often very divergent, implying the retention of allelic lineages by balancing selection (Takahata & Nei 1990), which results from negative frequency-dependence (Snell 1968) or from overdominance (Doherty & Zinkernagel 1975). Whatever the exact mechanism, the adaptive significance of MHC variation is obvious in many species and there is clear evidence for an association between MHC

Correspondence: W. Babik, Fax: +48 12 664 69 12; E-mail: wieslaw.babik@uj.edu.pl

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

genotype and susceptibility to disease (Briles *et al*. 1977; Thursz *et al*. 1997; Langefors *et al*. 2001; Carrington & Bontrop 2002; Bonneaud *et al*. 2006). Therefore, it has been suggested that loss of MHC variation may seriously jeopardize the future prospects of species and populations and increase the risk of extinction by facilitating the spread of infectious diseases (O'Brien & Evermann 1988; Hughes 1991; Hedrick 2001; Siddle *et al*. 2007).

The role of selection in shaping patterns of MHC variation is emphasized by reports of species possessing relatively high MHC variability despite a lack of/low genetic variation in other parts of the genome, an indication of the selective retention of MHC variation (Aguilar *et al*. 2004; Jarvi *et al*. 2004). However, a link between the level of MHC variation and the condition of populations or species has not been convincingly demonstrated. There are examples of species with very limited MHC variation showing demographic and/or geographical expansion in the recent past but with no apparent increase in susceptibility to pathogens (Ellegren *et al*. 1993; Mikko *et al*. 1999; Babik *et al*. 2005). Moreover, there are numerous examples of correlations between levels of MHC and neutral variation, indicating that genetic drift may strongly affect variation in MHC, at least on a shorter evolutionary timescale (Hedrick *et al*. 2001; Landry & Bernatchez 2001; Campos *et al*. 2006). It is thus possible that relaxed selection on MHC related to rapid spatial and demographic expansions following, e.g. major climatic changes, as those observed in the temperate and boreal regions during the Pleistocene, may have led to a decrease in genetic variation as commonly observed for neutral markers (Hewitt 1999; Babik *et al*. 2005).

Populations in areas of postglacial expansion are not only of recent origin, but often occur at the margins of species' ranges. Peripheral populations may be more prone to extinction because of small population sizes and suboptimal environmental conditions (Lesica & Allendorf 1995). Small populations may also exhibit lower adaptive potential because of environmental stress and genetic problems (Willi *et al*. 2006). On the other hand, marginal populations may contribute to the evolutionary potential of a species by accumulating local adaptations to limiting environmental conditions or locally abundant pathogens (García-Ramos & Kirkpatrick 1997). These may be particularly important in the face of rapid climate change (Parmesan 2006). Local adaptations acquired at the range margin may spread as climatic conditions change and thus facilitate response to climate change in other portions of the range. In this context, studies of MHC and neutral variation in marginal populations inhabiting previously glaciated areas are of great interest, particularly in low-vagility poikilotherms such as urodele amphibians.

Studies of MHC variation in amphibians are also motivated due to the catastrophic decline of this group of vertebrates (Stuart *et al*. 2004; Lips *et al*. 2006; Pounds *et al*. 2006). Multiple, sometimes interrelated factors, such as habitat fragmentation, environmental pollutants, UV-B radiation, emerging infectious diseases, climate change or introduced species are involved in this decline; however, the relative importance of each of these factors may differ for various taxonomic groups and geographical areas (Beebee & Griffiths 2005). There is a substantial body of evidence showing that infectious diseases, particularly fungal chytridiomycosis, possibly aggravated by climate change, are a major factor contributing to declines in many areas (Berger *et al*. 1998; Docherty *et al*. 2003; Garner *et al*. 2005; Lips *et al*. 2006; Pounds *et al*. 2006; Morgan *et al*. 2007). Because of its role in fighting pathogen assault, characterization of the patterns of MHC variation in amphibians is an important and urgent task. Conservation strategies for declining amphibians would benefit from knowledge on the adaptive variation in species, reflected in part by MHC allelic diversity.

MHC genetic architecture and variation is well characterized in teleost fishes, birds and mammals, but the remaining groups of vertebrates have been poorly investigated (Bernatchez & Landry 2003; Kelley *et al*. 2005; Piertney & Oliver 2006). Among amphibians, the organization of the MHC complex is known in *Xenopus* due to the ongoing genome project (Ohta *et al*. 2006) and some data on variation in MHC I is also available (Bos & Waldman 2006). MHC II variation has been studied in a discoglossid toad *Bombina bombina* (Hauswaldt *et al*. 2007). In the Urodele amphibians, information on MHC variation is available only for two very closely related ambystomatid salamanders, the axolotl (*Ambystoma mexicanum*) and tiger salamander (*A. tigrinum*; Tournefier *et al*. 1998; Sammut *et al*. 1999; Laurens *et al*. 2001; Bos & DeWoody 2005; Richman *et al*. 2007). However, as no urodelan MHC sequences outside the *A. tigrinum/A. mexicanum* clade have been published or deposited in GenBank, assessing general trends and patterns in MHC variation in this ancient group of vertebrates has not been possible.

The object of our study, the Alpine newt *Mesotriton alpestris*, has an extensive and complex geographical distribution in Europe (Fig. 1). This species also shows deep phylogeographical structure as revealed by mitochondrial DNA (mtDNA; Sotiropoulos *et al*. 2007). Its geographical range in Central Europe is interrupted by lowlands. In Poland, the species reaches the northeastern limit of its distribution and forms three allopatric groups of populations confined to the Sudetes, Carpathians and Holy Cross Mountains, separated by lowland hiatuses of at least 100 km (Fig. 1). It is abundant in the western part of the Polish Carpathians, but becomes rare towards the east (Głowaciński & Rafiński 2003). Previous studies of Polish Alpine newt populations conducted with skin grafts approximately assessed the variation in MHC I and suggested that these allopatric groups are genetically differentiated (Rafiński 1970, 1974). Subsequent research involving allozymes (Pabijan *et al*.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

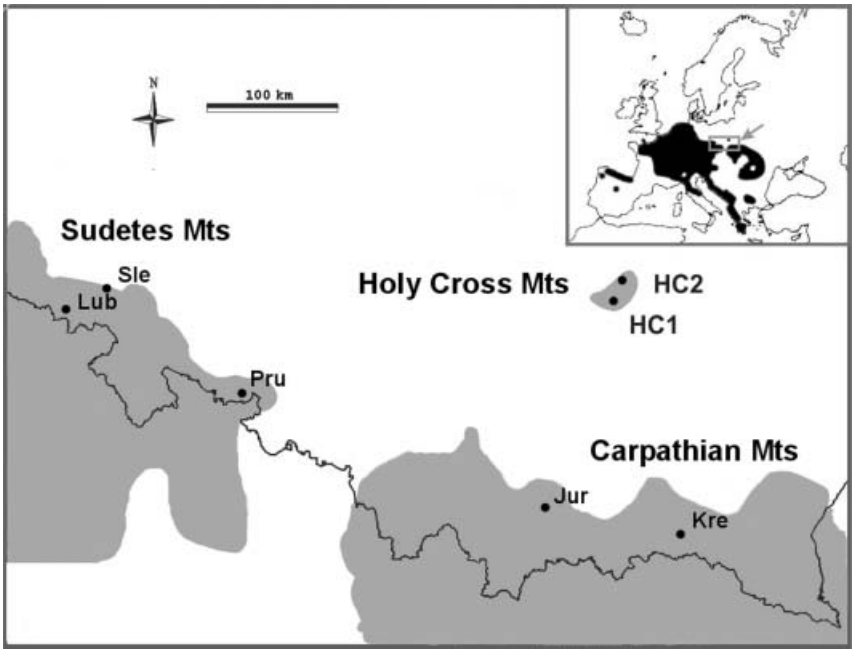

Fig. 1 Distribution of the Alpine newt *Mesotriton alpestris* in Poland and in Europe (shaded area). The locations of the studied populations are shown, abbreviations as in Pabijan & Babik (2006). The solid line represents Poland's southern political border.

2005) as well as microsatellites and mtDNA (Pabijan & Babik 2006) confirmed that these three groups of populations are genetically disparate and deserve the status of management units. However, mtDNA variation suggests that all north-central European populations are derived from a single glacial refugium (Pabijan & Babik 2006; Sotiropoulos *et al*. 2007). Microsatellite variation decreases eastwards in accordance with the inferred colonization route from a refugium located northeast of the Alps (Pabijan & Babik 2006; Sotiropoulos *et al*. 2007).

In the present study, we: (i) assess the levels of variation, number and expression pattern of the Alpine newt MHC II loci; (ii) test for signatures of historical positive selection and recombination in MHC II sequences; (iii) describe variation at the population level; (iv) estimate the degree of differentiation between populations and allopatric groups of populations; and (v) evaluate the relationship between MHC and microsatellite variation.

Materials and methods

Samples

Samples used in the present study were previously scored for microsatellite and mtDNA variation (Pabijan & Babik 2006). We analysed three populations from the Sudetes Mountains (SUD: Lub, Pru and Sle), two from the Carpathians (CAR: Jur and Kre) and two from the Holy Cross Mountains (HCM: HC 1 and HC 2) (Fig. 1). Details of sampling, DNA extraction, microsatellite and mtDNA typing are given in Pabijan & Babik (2006). Sample sizes used in the present study are given in Table 1.

Primer development

On the basis of a consensus of the *Ambystoma tigrinum* MHC II alleles (Bos & DeWoody 2005), we designed primers TrMHCII2F and TrMHCII1R to amplify partial exon 2 sequences from several individuals; sequences of all new primers used in this study are given in Table 2, and Fig. 2 shows location within the gene. We then used the consensus from these partial sequences to design specific primers within the exon, which led to successful amplification of full *Mesotriton alpestris* second exon sequences with adjacent parts of introns through vectorette polymerase chain reaction (PCR; see below). The primers were TrMHCII2F, TrMHCII4F and TrMHCII6F in the 3′ direction and TrMHCII3R, TrMHCII7R and TrMHCII8R in the 5′ direction. In the course of the study, we realized that not all MHC II alleles could be detected; therefore, additional vectorette PCR experiments were performed with the following primers: TrMHCII8F, TrMHCII9F, TrMHCII10F, TrMHCII11F, TrMHCII5R, TrMHCII9R and TrMHCII10R.

In the vectorette PCR approach, total genomic DNA is digested with a restriction enzyme (RE) producing sticky ends, then double-stranded adapters (vectorettes) matching the overhangs but showing some internal mismatch ('bubble') are ligated. Using one primer specific to the sequence in question and the other specific to the reverse complement of one of the vectorette strands (in the region of mismatch), it is possible to directionally amplify the genomic fragment located between the specific primer and the RE recognition site. Multiple REs are usually used to ensure a fragment of sufficient length is obtained. In the case of large, complex genomes (as in newts) a nested approach,

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Table 1 Sample sizes (N), allelic richness (R) and expected heterozygosities (H_E) for four putative MHC loci in seven peripheral populations of the Alpine newt. In the last column, mean allelic richness and expected heterozygosities for six microsatellite loci calculated for the same individuals as in the MHC analysis are presented. The data for the microsatellite analysis were taken from Pabijan & Babik (2006) and recomputed including only individuals scored for MHC variation

	N	*Meal-DAB*		*Meal-DBB*		*Meal-ψI*		*Meal-ψII*		Microsatellites	
Population/region		R	H_E	R	H_E	R	H_E	R	H_E	R	H_E
Kre	15	7.6	0.857	3.0	0.626	2.9	0.543	3.3	0.550	5.38	0.617
Jur	24	8.2	0.796	3.0	0.614	3.0	0.660	2.0	0.500	5.82	0.586
CAR	**39**	**13.7**		**3.0**		**3.0**		**3.9**		**5.6**	
HC1	13	3.0	0.564	3.0	0.564	3.0	0.526	2.0	0.500	4.51	0.543
HC2	24	3.0	0.560	3.0	0.619	2.5	0.531	2.0	0.382	4.98	0.573
HCM	**37**	**3.0**		**3.0**		**3.0**		**3.0**		**4.7**	
Pru	24	12.6	0.866	3.0	0.572	3.5	0.457	3.8	0.662	6.29	0.667
Lub	25	12.9	0.919	3.0	0.549	3.6	0.637	3.0	0.543	5.72	0.661
Sle	24	8.4	0.834	3.0	0.390	3.0	0.504	4.0	0.720	5.39	0.657
SUD	**73**	**22.9**		**3.0**		**3.8**		**3.9**		**5.8**	

CAR, Carpathians.
HCM, Holy Cross Mountains.
SUD, Sudetes.

Table 2 Sequences of the primers used in the study of MHC II variation in the Alpine newt. Their location within the gene is shown in Fig. 2

Primer	Sequence (5′–3′)
TrMHCII1F	TCTCTCCGCAGTGGACTTCGTG
TrMHCII2F	GAGTGTCACTTCCTGAACGG
TrMHCII4F	GGGTACTTCGTGGCCGACAC
TrMHCII6F	TCCTGCACTTCGACAGCGAC
TrMHCII7F	CTCCTCAGATGATTTCGTG
TrMHCII8F	GGTCCGGTTCCTGCAGCGCTT
TrMHCII9F	GAGTGTCASTTCSTGAACGG
TrMHCII10F	GAGTGTCASTTCSTGAACGGCTCTCAG
TrMHCII11F	TTCSTGAACGGCTCTCAGCGGGTCCGG
TrMHCII1R	TTCCAAGATCCCGTAGTTG
TrMHCII3R	GTGTCGGCCACGAAGTACCC
TrMHCII4R	AGCCCTCGTAGATCCCGTAGTTG
TrMHCII5R	CTCACGCYTCCGSTGCTCCATG
TrMHCII7R	CTCCAGGAAGTCCTCGTCTAT
TrMHCII8R	ASGAACTGCTGCYGGTTGTA
TrMHCII9R	GTCGGCCACGAAATACCCGC
TrMHCII10R	GGCTTCAGCCWCCGSTGCTCCATGAA
TrMHCII11R	CTCGAAGGAACCGTAGTTG
TrMHCII12R	CTCCATGAAGCCCTGGAGGAT

using an internal specific primer in the second PCR, is essential in order to eliminate false positives (i.e. spurious bands not representing the region of interest). Nested PCR was used in all the vectorette PCR experiments.

We adopted the modified vectorette PCR protocol of Ko *et al*. (2003). Briefly, approximately 2.5 μg portions of genomic DNA were digested with 20 U of *Bsu*15I, *Eco*RI, *Mun*I and *Xap*I REs (Fermentas) at 37 °C overnight in 100 μL volumes.

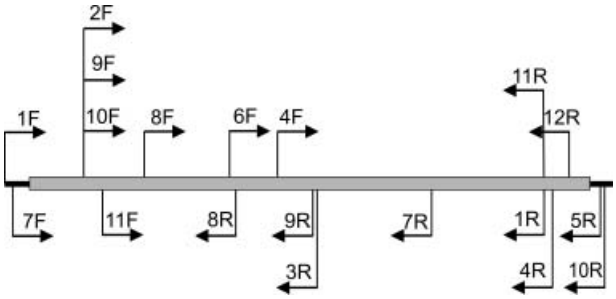

Fig. 2 Schematic representation of the position of various primers used in the study of MHC II variation in the Alpine newt. Grey box, MHC II 2nd exon; thick black line, adjacent parts of introns 1 and 2.

Double-stranded vectorette adapters consisting of vect53 and vect57TTAA (Ko *et al*. 2003) were ligated to digested DNA with 500 U of T4 DNA ligase (New England Biolabs). Products of the ligation reaction were either ethanol-precipitated or purified with Microcon 100 columns (Millipore) and suspended in 30 μL of 0.5× TE buffer. For the first vectorette PCR, we used a specific forward primer for amplification of the 3′ portion of the second MHC II exon and a reverse primer for the 5′ portion. The 20 μL PCRs contained 10 μL of 2× Hot Star Taq Master Mix (QIAGEN), 1 μM of the specific and C20 vectorette (Ko *et al*. 2003) primers and 1 μL of vectorette-ligated digested DNA. The touchdown PCR scheme was as follows: 15 min at 94 °C followed by five cycles of 30 s at 94 °C, 30 s at 65 °C and 60 s at 72 °C, five cycles with annealing at 62 °C and 20–25 cycles at 56 °C or 58 °C. In the second PCR, nested forward (for 3′) and reverse (for 5′ end amplification) primers were used together

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

with the B21 vectorette primer (Ko *et al*. 2003). As template, we used 0.5 μL of the product of the first PCR amplification. The second PCR cycling scheme consisted of: 15 min at 94 °C followed by 30 cycles of 30 s at 94 °C, 30 s at 58 °C and 60 s at 72 °C. The PCR product was run on a 1.5% agarose gel, clearly visible bands were excised, purified using the MinElute Gel Purification Kit (QIAGEN), and directly sequenced with the nested specific and vectorette D19 (Ko *et al*. 2003) primer using the BigDye Terminator (BDT) chemistry (ABI). Sequencing reaction products were electrophoresed on an ABI 3100 Genetic Analyser.

Three relatively divergent sequences were obtained with the vectorette PCR technique. We designed three primer pairs to specifically amplify each of these. The primer pairs were: (i) TrMHCII1F-4R, (ii) TrMHCII7F-5R, and (iii) TrMHCII1F-11R. In the course of the study, it was necessary to use additional primer combinations (iv) TrMHCII1F-12R and (v) TrMHCII1F-5R to amplify *DAB* alleles from cluster II (see Results).

Genotyping, allele identification and validation

We amplified *M. alpestris* MHC for all samples with each combination of primers (i–iii) in 10 μL reactions containing 0.5 U of HotStart *Taq* polymerase (Fermentas), 1.5 mM $MgCl_2$, 1 μM of FAM (forward) and HEX (reverse) labelled primers and 10–100 ng of genomic DNA. The cycling scheme was as follows: 3 min at 94 °C followed by 30–35 cycles of 30 s at 94 °C, 30 s at 58 °C and 30 s at 72 °C and the final extension step of 3 min at 72 °C.

Several approaches were adopted for genotyping and determining sequences of individual alleles. First, fluorescent single strand conformational polymorphism (SSCP) analysis was performed as follows. One microlitre of the PCR product was mixed with 0.5 μL of the GeneScan ROX 500 marker (ABI), 0.75 μL of 0.3 M sodium hydroxide and 12.75 μL formamide, denatured at 95 °C for 5 min, cooled on ice for 2 min and separated on an ABI PRISM 3100 Genetic Analyser. We used 5% GeneScan nondenaturing polymer (ABI) with 10% (w/v) glycerol and 1× running buffer with EDTA (ABI). Samples were injected for 15 s; electrophoresis was carried out at 15 kV and 20 °C. GENEMAPPER 3.7 software (ABI) was used for size calling and allele scoring.

Relatively few SSCP patterns were detected with primer combination (i). Individuals representing unique SSCP patterns (at least two per pattern where available) were amplified as above but with unlabelled primers. Three microlitres of the amplification product was mixed with 7 μL formamide containing 10 mM sodium hydroxide, and a grain of bromophenol blue. The mixture was denatured at 95 °C for 5 min, cooled on ice and run on SSCP GMA precast gels (Elchrom Scientific) for 16–20 h at 6 V/cm and 9 °C in 0.75× TAE buffer in the SEA2000 electrophoresis apparatus (Elchrom). After electrophoresis, gels were stained with SYBR Gold (Molecular Probes) and photographed under UV light. Then, disposable BandPicks (Elchrom) were used to excise small (*c*. 5 μL) pieces of individual SSCP bands. PCR reagents were added to the total volume of 30 μL and PCR was performed as described above for 30 cycles. PCR products were purified with ExoSAP-IT (USB) and sequenced with the TrMHCII1F primer. Amplicons showing a single sequence were sequenced in the other direction as well.

With primer combination (ii) most alleles as defined by their SSCP patterns occurred at least once as a homozygote (often due to nonamplification of a co-occurring allele, see Results). Such 'homozygous' individuals were amplified with unlabelled primers and sequenced directly. Sequencing of heterozygous individuals was used together with the respective SSCP patterns for cross-validation of some allele sequences. Sequences obtained from heterozygous individuals were usually of very high quality, allowing unambiguous detection of heterozygous peaks. In several individuals, cloning was used for determination of sequences of individual alleles. Comparisons of the cloned sequences with directly sequenced PCR products allowed the confirmation of allele sequences in two independent PCRs and controlled for PCR and cloning-generated artefacts. PCR products were cloned using the pGEM T-Easy vector system (Promega). Recombinant clones were detected by blue/white screening, colonies picked from plates were used directly as a template in PCR with M13F and M13R primers and PCR products of the expected length were purified and sequenced as described above.

Primer combination (iii) amplified the putative *MealDBB* locus exhibiting very limited polymorphism; multiple samples from each SSCP pattern ('genotype) were amplified with unlabelled primers and PCR products were sequenced directly. Each allele was sequenced from multiple homozygotes, and heterozygous genotypes detected by SSCP were confirmed by sequencing.

In the case of primer combinations (iv) and (v), fluorescent SSCP genotyping was followed by direct sequencing of unique SSCP patterns. All but one individual contained only a single allele amplified with either combination of primers; thus, genotyping and sequencing were straightforward.

We regarded only sequences obtained in at least two independent PCRs as confirmed alleles, regardless if they occurred in the same or different individuals. These include sequences obtained from re-amplification of SSCP bands, from direct sequencing and cloning. In designating allele names we followed the nomenclature rules set by (Klein *et al*. 1990).

Analysis of MHC expression

In order to validate if the assayed Alpine newt loci are expressed, we obtained RNA from the spleen of an individual

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

from the Carpathians. Immediately after excision, the tissues were preserved in the RNAlater stabilization reagent (QIAGEN).

RNA was extracted with the RNeasy kit (QIAGEN) from about 10 mg of the homogenized tissue sample including the DNAse treatment step. cDNA was obtained using Omniscript Reverse Transcriptase (QIAGEN) in 20 μL reactions containing 6 μL template RNA, 2 μL Oligo$(dT)_{12–18}$ primer (0.5 μg/μL, Invitrogen), 1 μL RNase inhibitor (10 μ/μL) and 1 μL RT and 6 μL of RNase-free water. The reaction was incubated at 37 °C for 60 min. The same five primer combinations as described above were used for amplification. PCR products from cDNA were run on SCCP GMA gels alongside those from genomic DNA; SSCP patterns were visually compared and all clearly visible bands were excised, re-amplified and sequenced.

Statistical, phylogenetic and population genetic analyses

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, using the Nei & Gojobori (1986) method with the Jukes-Cantor correction for multiple substitutions, were computed in MEGA 4 (Tamura *et al*. 2007). Standard errors of the estimates were obtained through 1000 bootstrap replicates.

We checked for signatures of recombination in our dataset using three methods. Two of these, GENECONV (Padidam *et al*. 1999) and MAXCHI 2 (Maynard Smith 1992) performed very well in an assessment of 14 recombination detection methods (Posada 2002). They are implemented in the RDP 2 software (Martin *et al*. 2005), used for computations. Additionally, a new method, genetic algorithm recombination detection (GARD) (Pond *et al*. 2006), was applied, through a web-based routine (http://www.datamonkey.org/GARD/).

We tested if positive selection shaped the evolution of the assayed MHC II sequences using two approaches. First, the one-tailed *Z*-test, as implemented in MEGA 4, compared the rates of synonymous vs. nonsynonymous substitutions at all codons, antigen binding sites (ABS) and non-ABS. The locations of the putative ABS and non-ABS were inferred from the human MHC II molecule structure (Brown *et al*. 1993). Second, we tested for positive selection by comparing the likelihoods of three codon-based models of sequence evolution available in PAML 3.15 (Yang 1997): (i) M0: one ω (dN/dS ratio), (ii) M7: nearly neutral ($\omega \leq 1$) with the beta distribution approximating ω variation, (iii) M8: positive selection (a proportion of sites evolving with $\omega > 1$) with the beta distribution approximating ω variation. The best-fitting models were chosen on the basis of the value of the Akaike information criterion (AIC) (Posada & Buckley 2004; Sullivan & Joyce 2005). Positively selected codons were identified through the Bayes empirical Bayes procedure (Zhang *et al*. 2005). Analyses could not be performed for the *Meal-DBB* locus because of the lack of nonsynonymous variation. Analyses for the *Meal-DAB* locus were performed for all alleles as well as for the alleles from clusters I and II (see Results) separately. All putative pseudogene sequences were analysed together after removing positions with indels.

Phylogenetic relationships among the Alpine newt MHC II sequences were reconstructed using an *A. tigrinum* allele as an outgroup. Two *Triturus cristatus* sequences representing a putative expressed locus and a pseudogene were also included.

The best-fitting model of sequence evolution chosen on the basis of the AIC criterion using MODELTEST 3.7 was the Kimura model with unequal base frequencies and gamma-distributed rate variation (K81uf + Γ). According to the model, nucleotide frequencies were as follows: 0.234 (A), 0.285 (C), 0.286 (G), 0.195 (T) and the gamma shape parameter $\alpha = 0.598$. A neighbour-joining tree was constructed from the matrix of distances computed according to the model. The robustness of the obtained tree topology was tested with 1000 bootstrap replicates. Another tree was constructed under the Bayesian approach with MRBAYES 3.1 (Ronquist & Huelsenbeck 2003). The likelihood settings corresponded to the GTR + Γ model with the parameter values estimated from the data; priors were set to default values. Two independent runs of four Metropolis coupled Markov chain Monte Carlo simulations (three of them 'heated', temperature = 0.20) were each run for 1.1×10^6 generations and sampled every 1000 generations. The first 100 trees were discarded as burn-in, resulting in 2000 sampled trees. To calculate the posterior probability of each bipartition, the majority-rule consensus tree was computed from the 2000 sampled trees.

We tested conformance of the allele frequencies with Hardy–Weinberg expectations for all putative loci in all populations by applying exact tests (Guo & Thompson 1992) in GENEPOP 3.3 (Raymond & Rousset 1995). Linkage disequilibria within populations were tested for each pair of loci using the Fisher exact test or its Markov chain Monte Carlo approximations in GENEPOP. A locus-by-locus analysis of molecular variance (AMOVA) approach as implemented in ARLEQUIN (Excoffier *et al*. 2005) was used to partition the total variance in allele frequency data into among group, among populations within groups and within population components.

The standardized genetic differentiation measure of Hedrick (2005) was implemented in order to compare the overall level of genetic differentiation for the different loci. This measure compares the observed differentiation in relation to the maximum possible given the allele frequencies and within-population heterozygosity.

Differences in allelic richness between the three allopatric groups of populations for each putative MHC locus were assessed through randomization tests (10 000 randomizations)

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

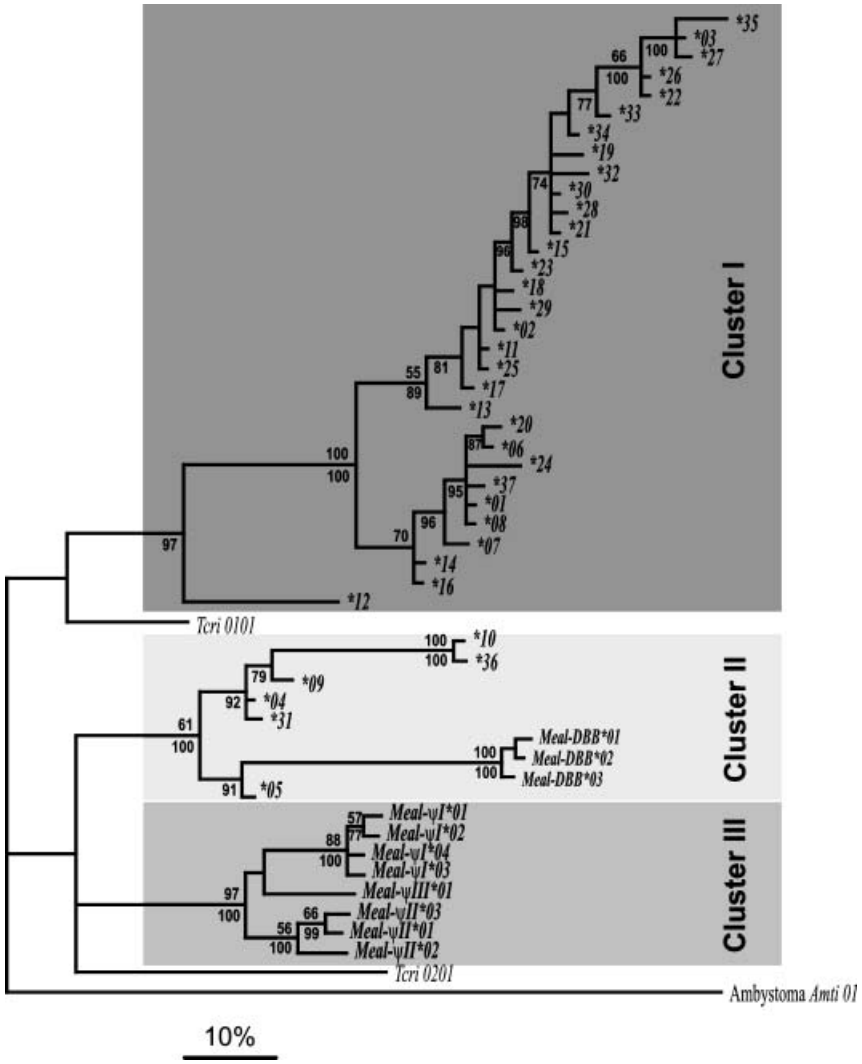

Fig. 3 Phylogenetic relationships of the Alpine newt MHC II alleles. The 50% majority rule tree from the Bayesian analysis is shown. Numbers **01–*37* refer to alleles *Meal-DAB*01–37*. Bootstrap values above 50% (from the neighbour-joining analysis) are shown above respective branches; Bayesian posterior probabilities above 70% are shown below branches. *Triturus cristatus* alleles from expressed (*Tcri0101*) and pseudogene (*Tcri0201*) loci were included for comparison. Trees were rooted using *Ambystoma tigrinum* sequence.

using the Excel add-in POPTOOLS (Hood 2006). The sequential Bonferroni procedure was applied where appropriate to keep the type I error level at $\alpha \leq 0.05$ (Rice 1989).

We adopted the approach of Beaumont & Nichols (1996) to test for loci showing a disproportionately high or low level of differentiation among populations, which may indicate the action of selection. Overall F_{ST} was estimated for six microsatellite and four MHC loci with the program DATACAL. This F_{ST} estimate served as a basis for coalescent simulations aimed at determining the expected variance of the F_{ST} estimates. The simulations involved 100 demes of which seven were sampled, with 46 genes drawn from each deme. F_{ST} and expected heterozygosity were recorded for each of the 100 000 simulations. Simulations using the infinite allele model (IAM) were performed with the program FDIST 2. We tested for F_{ST} outliers in our data set by assessing if F_{ST} values of individual loci fall inside the heterozygosity-scaled 95% F_{ST} quantiles obtained from simulations.

Results

Phylogenetic analysis

In the studied populations of the Alpine newt, we detected 48 unique sequences showing homology to the MHC II 2nd exon (GenBank Accession nos EU512166–213). Despite the relatively short and highly variable sequences, phylogenetic analysis distinguished three well-supported clusters (Fig. 3). Cluster I contains 31 diverse alleles and has 0.97 posterior probability (PP) but bootstrap support (BS) below 0.50; however, when the rather divergent allele *Meal-DAB*12* is excluded, the remaining 30 alleles receive PP and BS 1.00. Cluster II groups nine alleles with PP 1.00 and BS 0.61. Within this cluster group, *Meal-DBB*01–03* received 1.00 PP and BS. Nine sequences group together with 1.00 PP and 0.97 BS, forming cluster III. Two subgroups within this cluster are apparent: one, with PP 1.00 and BS 0.88, comprises alleles

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

*Meal-ψI*01–04*, and the other consists of *Meal-ψI*01–03* and receives 1.00 PP and 0.56 BS. The positions of two *T. cristatus* alleles, representing an expressed gene and a putative pseudogene (W. Babik, M. Pabijan, J. Radwan, unpublished data), are not well resolved, but it should be noted that they do not form a well-supported group to the exclusion of the Alpine newt alleles.

Analysis of expression and pseudogene detection

PCR amplification using a cDNA template prepared from DNAse-treated spleen RNA revealed that several alleles detected in genomic DNA could not be amplified from cDNA. Alleles *DAB*02* and *DBB*02* were amplified successfully but alleles *ψI*01, ψII*01* and **02* detected in the genomic DNA, were not amplified from cDNA. This suggests that these alleles represent pseudogenes. Interestingly, primers TrMHCII1F–TrMHCII4R amplified allele *DAB*04* from cDNA but not from genomic DNA. The presence of allele *DAB*04* in genomic DNA was confirmed using the allele-specific primer TrMHCII12R. All sequences in the putatively expressed clusters I and II exhibit a number of amino acid residues, including cysteines (C) in positions 7 and 71 and glycosylation signals (N-11, F-32), that are highly conserved across vertebrates (Kaufman *et al*. 1994).

The hypothesis that alleles *ψI*01, ψII*01* and **02* represent nonfunctional pseudogenes is also supported by sequence analysis. These alleles form a well-defined group with several other variants on the phylogenetic tree (cluster III, Fig. 3). Two members of this group (*ψII*01* and *03*) exhibit two insertions (2 bp and 1 bp long) causing a frame-shift compared to the reference *Ambystoma* and *Triturus cristatus* sequences. In all three *ψI* alleles, there is a tryptophan (W) instead of a cysteine (C) in position 71, and thus, the formation of the disulphide bond essential for the proper functioning of the MHC II molecule is not possible.

Thus, from the expression analysis, it is clear that only a subset of MHC II sequences are expressed in the Alpine newt.

The number of loci and assignment of alleles

We were able to assign alleles to putative loci by combining genomic DNA, cDNA, phylogenetic and population genetic analyses. Up to nine unique sequences were detected per individual, implying the existence of at least five loci. Relationships of sequences revealed by phylogenetic analysis, the maximum number of sequences from particular clusters observed per individual and patterns of co-occurrence of sequences in genotypes of individual newts enabled several conclusions to be drawn regarding the assignment of alleles to loci. These conclusions, further supported by tests of Hardy–Weinberg (HW) proportions, were as follows:

1. Alleles *Meal-DBB*01* and **02* (confirmed expression) and *Meal-DBB*03* of cluster II clearly belong to the same locus which we named *Meal-DBB*; these sequences are closely related (PP and BS 1.00) and differ by at most two synonymous substitutions; never were more than two sequences detected per individual and under the assumption that these alleles belong to the same locus, genotype frequencies in all populations conform to HW expectations. This should not be the case if alleles from other clusters belong to this locus, as their exclusion would result in an excess of homozygotes. Thus, we conclude that only these three alleles are present in appreciable frequencies at the *DBB* locus in Polish populations of the Alpine newt.
2. The remaining alleles from cluster II and all alleles from cluster I most likely belong to the other expressed locus which we designate *Meal-DAB*. Assignment of these alleles to a single locus is not obvious from inspection of the tree, but several lines of evidence support such a conclusion. First, we never detected more than two sequences from these clusters (excluding sequences assigned to the *DBB* locus; see above) in a single individual. Second, we did not detect any cluster II alleles in individuals heterozygous for alleles from cluster I. Moreover, among three individuals which consistently did not amplify with primers MHCTr1F–MHCTr1R (a combination specific for cluster I alleles), two turned out to be heterozygous for alleles from cluster II. Third, before alleles from cluster II were detected almost all populations showed a deficit of heterozygotes at the putative *DAB* locus. After inclusion of these alleles only one population (Lub) showed an excess of homozygotes in *DAB (*see below). These observations provide strong support for the hypothesis that both sequences from cluster I and II belong to the single *DAB* locus.
3. From two to five unique sequences from cluster III were detected per individual. Sequences form two well-supported groups which we interpret as representing pseudogene loci *Meal-ψI* and *Meal-ψII*. In addition to phylogenetic evidence, this interpretation is supported by population level analyses, because we never detected more than two sequences from either group per individual and genotypes conform to HW proportions in all populations for locus *ψI* and in four of seven populations for *ψII (*before Bonferroni correction). Allele *Meal-ψIII*01*, detected only in minority of individuals and not in all populations, probably belongs to another locus, from which not all alleles could be amplified with our primers. We included allele *Meal-ψIII*01* in the phylogenetic analysis but do not consider it further in the population-level analyses.

Thus, we conclude that most likely two expressed and three pseudogene loci are present in the Alpine newt.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Table 3 Amino acid alignment of the Alpine newt MHC II alleles from the expressed loci. Position 1 of the alignment corresponds to the fourth amino acid position of the second exon. Length differences among alleles result from the use of various reverse primers (see text). Shaded columns represent putative antigen binding sites (ABS) (Brown *et al*. 1993); signs 'x', '*' and '**' denote amino acid positions under positive selection with posterior probability > 50, 95 and 99%, respectively, as identified by the Bayes empirical Bayes procedure (Zhang et al. 2005); positive selection analysis based on all cluster I (see text and Fig. 3) alleles except divergent *Meal-DAB*12*

```
                                10        20        30        40        50        60        70        80
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
                        x x x  x        x x       xx       x     x       x      xx  x x       x     x
                           *             *        **       *     *       *       *  *          *     *
                           *                      **                               *  *          *     *
                                                                                                  DAB locus
Meal-DAB*01   HQMKAECQFLNGSQRVRLLERWVYNRQQYAHFDSDVGYFVADTEAGEPSAKHWNSDKALLEQMRGEVDSFCRHNYGILEG
Meal-DAB*02   L...G..H.........F..........IL.......H.....G.......Q......F...A.A.............
Meal-DAB*03   L.K.S..H.........Y.A........FV.....................R......FM..E.A.......T....F..
Meal-DAB*04   AEC.S..H.V.......F.Q.YS.....FL.....R.........F...D..Y...NE.YM..L.....R...T...
Meal-DAB*05   AEC.S..H.V.......F.Q.YS.....L......L..Y.....F...D..Y...NE.YM..L.....R...T...
Meal-DAB*06   ....G.....................................................................F..
Meal-DAB*07   ........................L.....................................................F..
Meal-DAB*08   .........................................................................F..
Meal-DAB*09   AEC.S..H.V.......Y.R.FS.....FL.....R.........F...D..Y...NE.YM..L.....R...T...
Meal-DAB*10   AEL.R..H.V.......Y.Q.LS.....FL.....R..Y.....F...DV.K....E.YM..L.....R...Y..E
Meal-DAB*11   L...G..H.....................IL.......H.....G.......Q......F...A.A.............
Meal-DAB*12   YEW.S..H.V...........HS.....FL.....L..Y.....F......Q..N...FM..E.AA..R...Y...
Meal-DAB*13   ...............F..........IL.......H.....G.......Q......FM..E.A.............
Meal-DAB*14   ........................IL.......H.....................................F..
Meal-DAB*15   L...G..H.........Y.A........IL.......H.....G.......Q......F...A.A.............
Meal-DAB*16   ........................IL.....................................................F..
Meal-DAB*17   ....G..H.....................IL.......H.....G.......Q......F...A.A.............F..
Meal-DAB*18   L...G..H.........F..........IL.......H.....G.............F...A.A.............
Meal-DAB*19   L...G..H.........Y.A........IL.......H.....G.......Q......F...E.A.......T....F..
Meal-DAB*20   ....G..H.....................................................................F..
Meal-DAB*21   L...G..H.........Y.A........FV.......H.....G.......Q......F...A.A.............F..
Meal-DAB*22   L...G..H.........Y.A........FV.....................R......FM..E.A.......T....F..
Meal-DAB*23   L...G..H.........F.A........IL.......H.....G.......Q......F...A.A.............
Meal-DAB*24   ..............................H.....G.......Q.....................................
Meal-DAB*25   L...G..H.....................IL.......H.....G.......Q......F...A.A.............F..
Meal-DAB*26   L...G..H.........Y.A........FV.....................R......FM..E.A.......T.......
Meal-DAB*27   L.K.S..H.........Y.A........FV.....................R......FM..A.A.......T....F..
Meal-DAB*28   L...G..H.........Y.A........FV.......H.....G.......Q......F...A.A.......T.......
Meal-DAB*29   L...G..H.........F..........IL.......H.............Q......F...A.A.............
Meal-DAB*30   L...G..H.........Y.A........FV.......H.....G.......Q......F...A.A.............
Meal-DAB*31   AEC.S..H.V.......F.Q.YS.....IL.....R.........F...D..Y...NE.YM..L.....R...T...
Meal-DAB*32   L.K.S..H.........Y.A........FV.......H.....G.......Q......F...A.A.............
Meal-DAB*33   L...G..H.........Y.A........FV.....................R......F...A.A.............
Meal-DAB*34   L...G..H.........Y.A........FV.............G.......Q......F...A.A.............
Meal-DAB*35   L.K.S..H.........Y.A........FV.....................R.....................T....F..
Meal-DAB*36   AEL.G..H.V.......Y.Q.LS.....FL.....R..Y.....F...DV.K....E.YM..L.....R...Y..E
Meal-DAB*37   ................G.....................................................F..
                                                                                                  DBB locus
Meal-DBB*01   AEC.S..H.V.......F.Q.YS.....L......L..Y.....L..ADV.D....EDVIRRK.AQ..AY..
Meal-DBB*02   AEC.S..H.V.......F.Q.YS.....L......L..Y.....L..ADV.D....EDVIRRK.AQ..AY..
Meal-DBB*03   AEC.S..H.V.......F.Q.YS.....L......L..Y.....L..ADV.D....EDVIRRK.AQ..AY..
```

Sequence diversity, selection and recombination

Allelic diversity and sequence divergence among alleles differed strikingly among loci (Tables 1 and 3). The most variable and diverse was locus *DAB* in which 37 alleles were detected. Nucleotide and amino acid distances among alleles were highest at this locus as well. Even when *DAB* sequences from clusters I and II were considered separately, sequence divergence among alleles was high (Table 4). Both nucleotide and amino acid distances were much higher in the putative antigen binding sites (ABS) than in non-ABS sites. The three alleles of the *DBB* locus were very similar, differing by at most two synonymous substitutions (0.9%). Sequence divergence among alleles from all three putative pseudogene loci was moderate, and there was no difference between ABS and non-ABS (Table 4).

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

K2P nucleotide distance			Poisson-corrected amino acid distance		
All sites	ABS	Non-ABS	All sites	ABS	non-ABS
Meal-DAB all alleles					
9.6 (1.2)	22.0 (4.4)	5.8 (1.1)	19.2 (3.7)	47.6 (12.7)	11.2 (2.8)
Meal-DAB cluster I alleles					
5.6 (1.0)	13.7 (3.2)	3.0 (0.9)	12.8 (2.8)	34.1 (9.8)	6.5 (2.2)
Meal-DAB cluster II alleles					
6.3 (1.2)	16.5 (4.2)	3.1 (1.0)	10.3 (2.6)	28.5 (9.6)	4.7 (2.1)
Pseudogenes					
6.6 (1.2)	7.7 (2.7)	6.4 (1.3)	11.0 (2.5)	13.3 (6.3)	10.3 (2.7)

Table 4 The average nucleotide and amino acid distances among the Alpine newt MHC II alleles from the expressed putative *Meal-DAB* locus and alleles from putative pseudogenes *Meal-ψI* and *Meal-ψII*. Nucleotide distances are corrected for multiple substitutions with the Kimura 2-parameter model (K2P), amino acid distance is corrected using expectations from the Poisson distribution. ABS, putative antigen binding sites as determined by Brown *et al*. (1993). Distances are given as percentages per site with standard errors based on 1000 bootstrap replicates are given in parentheses

Table 5 The average rates of nonsynonymous substitutions per nonsynonymous site (dN) and synonymous substitutions per synonymous sites (dS) computed according to the Nei–Gojobori (1986) method, with standard errors obtained through 1000 bootstrap replicates in parentheses, and the results of the Z test of positive selection. ABS, putative antigen binding sites as determined by Brown *et al*. (1993). dS and dN values are given as percentages per site

Sites	dN	dS	*P*	Z
Meal-DAB all alleles				
All	10.9 (2.1)	5.1 (1.7)	0.014	2.221
ABS	26.8 (7.9)	5.3 (3.2)	0.004	2.677
non-ABS	6.0 (1.6)	5.0 (2.0)	0.355	0.374
Meal-DAB cluster I alleles				
All	6.6 (1.6)	2.3 (1.4)	0.022	2.042
ABS	16.8 (4.9)	1.4 (1.0)	0.001	3.181
Non-ABS	3.2 (1.1)	2.5 (1.7)	0.379	0.309
Meal-DAB cluster II alleles				
All	6.1 (1.7)	7.0 (2.8)	1.000	–0.278
ABS	19.5 (7.3)	7.0 (6.8)	0.120	1.179
Non-ABS	2.0 (0.9)	7.0 (3.3)	1.000	–1.585
Pseudogenes				
All	4.9 (1.1)	13.1 (3.9)	1.000	–2.141
ABS	5.6 (2.5)	14.2 (9.7)	1.000	–0.745
Non-ABS	4.6 (1.3)	12.8 (4.4)	1.000	–1.805

Clear signals of historical positive selection for amino acid replacements in the codons involved in antigen binding were detected at the *DAB* locus. The ratio of nonsynonymous to synonymous substitutions (dN/dS) was significantly higher than one in the ABS but not in non-ABS (Table 5). When alleles from the *DAB* locus belonging to clusters I and II were analysed separately, the *Z*-test of positive selection was significant only for alleles from cluster I. However, dN was almost three times higher than dS in ABS for alleles from cluster II contrasting strongly with results for non-ABS, where dS was much higher than dN (Table 5). The *Z*-test of positive selection was not significant for pseudogene alleles, in either ABS or non-ABS positions.

Similar results were obtained from comparisons of various models of codon evolution (Table 6). A model incorporating positive selection explained the evolution of *DAB* sequences best; this was also the best model when only alleles from cluster I were included in the analysis, but when only sequences from cluster II were included, a model with one dN/dS ratio adequately explained data (Table 6). When all *DAB* sequences were analysed, the Bayes empirical Bayes (BEB) procedure identified four codons (PP > 0.9) as evolving under positive selection, all of which were in ABS ($P = 0.0038$, Fisher's exact test). The same four codons were identified as positively selected when only sequences from cluster I were analysed. Interestingly, removing the divergent *Meal-DAB*12* allele from cluster I sequences increased the number of codons with PP of positive selection > 0.9 to 11, eight of which are located in ABS, which is significantly more than expected by chance ($P = 0.0015$, Fisher's exact test).

The evolution of the *DAB* cluster II sequences is adequately approximated by the nearly neutral model; ΔAIC between this and the model incorporating positive selection was only 1.4 and only a single codon had PP of being under positive selection > 0.9. For the putative pseudogene sequences, AIC values were virtually the same for all models, indicating that one dN/dS ratio model ($\omega = 0.175$) adequately explains the evolution of these sequences (Table 6).

Visual inspection of the alignment revealed an unusual pattern: the proximal 85–120 nucleotides of alleles *Meal-DAB*04, 05, 09* and *31* are identical to the corresponding portion of alleles *Meal-DBB*01–03* (Table 3). This is a suggestive example of interlocus gene conversion or recombination, which was confirmed by the GENECONV, CHI 2 and GARD methods (sequential Bonferroni adjusted $P < 0.05$).

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Model	ln *L*	ΔAIC	Parameters
	DAB all alleles		
M0 – one ω	–1059.1	173.4	$\omega = 0.673$
M7 – nearly neutral with beta	–992.2	41.6	
M8 – positive selection with Beta ($\omega_0 \leq 1, \omega_2 > 1$)	–969.4	Best	$p_0 = 0.947, p_2 = 0.053, \omega_2 = 7.667$
	DAB cluster I alleles		
M0 – one ω	–795.9	121.4	$\omega = 1.238$
M7 – nearly neutral with beta	–753.1	37.8	
M8 – positive selection with Beta ($\omega_0 \leq 1, \omega_2 > 1$)	–732.2	Best	$p_0 = 0.947, p_2 = 0.053, \omega_2 = 11.267$
	DAB cluster II alleles		
M0 – one ω	–436.5	20.8	$\omega = 0.266$
M7 – nearly neutral with beta	–564.2	Best	
M8 – positive selection with Beta ($\omega_0 \leq 1, \omega_2 > 1$)	–564.1	1.2	$p_0 = 0.830, p_2 = 0.170, \omega_2 = 2.755$
	Pseudogenes		
M0 – one ω	–484.1		$\omega = 0.175$
M7 – nearly neutral with beta	–484.2		
M8 – positive selection with Beta ($\omega_0 \leq 1, \omega_2 > 1$)	–484.2		$p_0 = 1.000$

Table 6 Evaluation of the goodness of fit for different models of codon evolution and estimated parameter values. ω, dN/dS; nearly neutral with beta, for all sites $\omega \leq 1$ and the beta distribution approximates ω variation; positive selection, a proportion of sites evolves with $\omega > 1$; p_0, proportion of sites with $\omega \leq 1$; p_2, proportion of positively selected sites ($\omega > 1$); ω_2, estimated value of ω for sites under positive selection

Level of AMOVA structure	*Meal-DAB*	*Meal-DBB*	*Meal-ψI*	*Meal-ψ2*
Among regions	4.7	–1.9	14.5*	8.6***
Among populations within groups	5.0***	8.7***	4.8**	4.5*
Within populations	90.3***	93.2***	80.7***	86.8***

Table 7 Percentage of total variation, explained by various levels of population structure in MHC loci of the Alpine newt. $^*P < 0.05$, $^{**}P < 0.001$ $^{***}P < 0.0001$

No other inter- or intralocus recombination event detected by all three methods was observed.

Genotype proportions, null alleles and linkage disequilibria

Meal-DAB genotype frequencies conformed to HW expectations in all but one population (Lub, $P < 0.0001$, unadjusted *P*). Loci *Meal-DBB* and *Meal-ψI* did not depart significantly from HW proportions in any populations. However, significant departures from HW proportions were detected for pseudogene locus *Meal-ψII* in three populations (Jur, $P = 0.0012$, Pru, $P = 0.0036$, Sle, $P = 0.0032$).

If departures from HW equilibrium were caused by the presence of null alleles in *DAB* and *ψII*, then their frequencies can be estimated. Because no null homozygotes were observed in *DAB* in Lub, the frequency of null allele(s) was probably < 0.2. Based on the count of apparent null homozygotes, the frequencies of *ψII* null(s) in Jur, Pru and Sle were estimated by the EM algorithm as 0.24, 0.21 and 0.24.

Significant linkage disequilibria were detected only between loci *DAB* and *DBB* in four populations (Kre, $P = 0.0053$; Jur $P = 0.0009$; HC1 and HC2 $P < 0.0001$).

Geographic differentiation and comparison with microsatellites

Allelic richness and expected heterozygosity (gene diversity) are given in Table 1.

According to the standardized G_{ST} values of Hedrick (2005), the putative loci achieved the following percentages of maximum among population differentiation: *Meal-DAB*, 58.1%; *Meal-DBB*, 22.1%; *Meal-ψI*, 49.1%; and *Meal-ψII*, 30.9%.

AMOVA showed that for all loci, over 80% of the genetic variation was partitioned within populations (Table 7), a common finding for highly polymorphic loci. Less than 9% of the variation was partitioned among populations within groups. Among-group differentiation ranged from –1.9% in *DBB*, and 4.7% in *DAB*, both values not significant, indicating a lack of structure among regions in expressed genes, to a relatively high value of 14.5% in *Meal-ψI* (Table 7).

Striking differences were observed in the levels of variation between regions for the *Meal-DAB* locus: 31 alleles (23 private) were detected in SUD, 14 (five private) in CAR and only three alleles (**01*, **02*, **04*, none of them private) in HCM. The differences in allelic richness were significant in

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

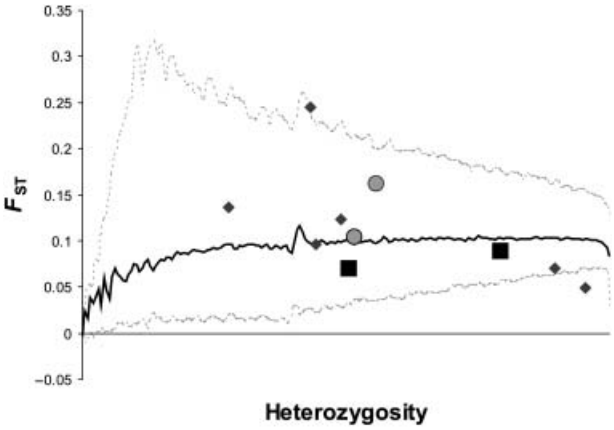

Fig. 4 F_{ST} outlier analysis. Solid line indicates median value and dashed lines the upper and lower 95% quantiles based on 100 000 simulations. Diamonds are observed values for microsatellite loci, squares for expressed MHC loci and circles for putative MHC pseudogenes.

all three pairwise comparisons (CAR vs. SUD $P < 0.05$, CAR vs. HCM $P < 0.001$, HCM vs. SUD $P < 0.001$, all P values Bonferroni corrected). The less polymorphic loci did not show such remarkable differences in allelic richness. The three *Meal-DBB* alleles were detected in all populations. Both pseudogene loci showed modest differences in the number of alleles, ranging from three (CAR and HCM) to four (SUD) per region in *Meal-ψI* and from two (HCM) to four (CAR and SUD) in *Meal-ψII*. Allelic richness for *Meal-ψI* was significantly higher in SUD than in CAR ($P < 0.05$), and for *Meal-ψII*, it was higher in SUD than in HCM ($P < 0.001$).

We used microsatellite loci genotypes from Pabijan & Babik (2006) in order to assess if the levels of neutral and adaptive genetic variation in populations of the Alpine newt are associated. Allelic richness for six microsatellite loci (Table 1) was calculated for the same individuals as in the MHC analysis. The number of alleles per locus ranged from 4 (*Ta3Caga2*), 5 (*Ta1Ca1*), 6 (*Ta4Ca4U*) and 8 (*Ta3Ca8*), to 46 (*Ta3Caga1*) and 51 (*Ta1Caga4*). There was a significant correlation between microsatellite and *Meal-DAB* allelic richness ($r = 0.88$, d.f. = 5, $P = 0.008$) as well as between allelic richness in *Meal-DAB* and the putative pseudogene *Meal-ψI* ($r = 0.87$, d.f. = 5, $P = 0.010$).

A weak correlation was observed between pairwise F_{ST} for microsatellites and *Meal-DAB* ($r_M = 0.22$, $P = 0.031$, Mantel test) and *Meal-ψI* ($r_M = 0.23$, $P = 0.031$); these correlations were not significant after Bonferroni adjustment. There was no correlation of pairwise F_{ST} between microsatellites and *Meal-DBB* or *Meal-ψII*.

Neither expressed nor putative pseudogene MHC loci were significant in the F_{ST} outlier analysis. However, two microsatellite loci were significant, one of which (*Ta4Ca4U*) showed a marginally higher than expected (one-sided $P = 0.022$) and the other (*Ta3Caga1*) lower than expected by chance differentiation among populations (one-sided $P = 0.00012$) (Fig. 4).

Discussion

Two putative expressed MHC II loci, differing strikingly in levels of variation, were detected in populations of the Alpine newt sampled at the northeastern periphery of its range. The *Meal-DAB* locus shows very high polymorphism, evidence of historical positive selection for amino acid replacements, and substantial geographical differentiation in allelic richness, despite relatively small geographical distances between sampled populations. On the contrary, the *Meal-DBB* locus exhibits low polymorphism and a uniform distribution of three alleles among populations and geographical regions. We compare these patterns with presumably neutral MHC pseudogenes and microsatellites and conclude that selection and drift interplayed to produce the pattern of MHC variation observed in marginal populations of the Alpine newt.

The number of loci and assignment of alleles

We detected multiple Alpine newt sequences showing homology to the second MHC exon and assigned them to two putative expressed and three nonexpressed, pseudogene loci. Alleles assigned to individual loci mostly clustered together on the phylogeny. We attribute the discrepancy observed for some alleles of the *Meal-DAB* locus, which formed a cluster with *Meal-DBB* alleles, to interlocus gene conversion or recombination, detected by several statistical methods. The assignment of alleles to loci is desirable because it enables more in-depth analyses and gives access to a wide repertoire of analytical tools devised for codominant markers.

This is the first report of more than one expressed MHC II locus in a urodele amphibian. The only other urodeles studied to date, *Ambystoma tigrinum* and *A. mexicanum*, possess a single expressed locus (Laurens *et al*. 2001; Bos & DeWoody 2005; Richman *et al*. 2007). However, multiple pseudogenes, containing the beta 1 domain of MHC II are probably present in *A. mexicanum* (Tournefier *et al*. 1998; Laurens *et al*. 2001).

The situation detected in the Alpine newt resembles that typical of mammals where alleles of various MHC II genes form monophyletic groups, sharing alleles among loci is relatively rare (Nei & Hughes 1992) and usually restricted to cases of relatively recent duplications (e.g. Babik & Radwan 2007). Despite the clear instance of interlocus gene conversion/recombination, the pattern observed in the Alpine newt is inconsistent with the common pattern in nonmammalian vertebrates in which extensive interlocus

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

allele sharing is common, making assignment of alleles to loci often difficult, problematic or even impossible (Wittzell *et al*. 1999; Miller & Lambert 2004; Reusch & Langefors 2005; van Oosterhout *et al*. 2006). Further studies on urodele MHC are obviously needed for a better understanding of MHC architecture and patterns of evolution in this ancient vertebrate group.

Patterns of historical selection

We detected clear signatures of positive selection at the *Meal-DAB* locus. However, the results of the analysis depended to some degree on which *DAB* sequences were included (Table 6). The relatively high overall sequence divergence in *DAB* may have saturated both the synonymous and nonsynonymous sites (Richman *et al*. 2001) and compromised the detection of positive selection in analyses involving all sequences. Accordingly, the strongest signature of positive selection, with a significantly elevated proportion of positively selected codons at ABS sites, was obtained for cluster I sequences after removal of the divergent *Meal-DAB*12* allele. For *DAB* cluster II sequences, a model allowing for positive selection offered only a marginally better fit to the data than the single ω model. Here, the recombination we have detected may have affected the tests of positive selection (Anisimova *et al*. 2003; Shriner *et al*. 2003); specifically, gene conversion/recombination has apparently homogenized the 5′ part of the second exon of several alleles, thus increasing similarity among the *DAB* cluster II sequences, which therefore may have decreased the power of the tests of selection. Unfortunately, we could not explicitly account for recombination in the analysis because the data clearly did not meet the assumption of the available method (Wilson & McVean 2006), which requires that sequences are sampled from a single panmictic population.

Our data and the results of Bos & DeWoody (2005) and Richman *et al*. (2007) from *Ambystoma* contradict the opinion that low MHC variation detected previously in *A. mexicanum* (Tournefier *et al*. 1998) may reflect a lack of selection for maintenance of MHC variation, suggested to be a general characteristic of urodeles explaining their weak immune response.

Signatures of positive selection were not detected in pseudogene sequences of the Alpine newt. This is contrary to what is expected if nonfunctional pseudogenes originate from MHC genes through the birth-and-death process (Nei & Rooney 2005), because patterns indicative of positive selection may persist in nucleotide sequences for long periods of evolutionary time (Garrigan & Hedrick 2003). Thus, the MHC II pseudogenes present in the Alpine newt have either evolved neutrally or under purifying selection, uniformly across all nucleotide positions. A similar conclusion was drawn from a study of MHC pseudogene variation in the little greenbul (Aguilar *et al*. 2005).

Levels of variation, geographical differentiation and comparison with neutral markers

The two putative expressed loci differ markedly in level of variation in populations of the Alpine newt at the northeastern margin of its range: a total of 37 and three alleles were detected in *Meal-DAB* and *Meal-DBB*, respectively. The *DAB* alleles showed extensive divergence and elevated dN/dS ratio in ABS, viewed as compelling evidence for positive selection acting on an evolutionary timescale. High levels of allelic diversity in marginal populations of postglacial origin imply that *DAB* variation has been maintained by selection on the ecological timescale as well.

On the contrary, surprisingly little variation of apparently recent origin (only a few synonymous substitutions), was detected at the other expressed locus, *DBB*. Three hypotheses may explain this low variation. First, *DBB* may represent a nonclassical MHC locus and low variation may result from functional constraint. Second, a lack of selection for maintaining variation at *DBB* could have resulted in the loss of variation via random drift. Third, variation previously present at this locus could have been lost when an advantageous allele had swept to fixation, whereas new mutations may have accumulated since then. Interestingly, the *DBB* alleles are present in high and similar frequencies in all populations, a situation typical for genes under balancing selection. However, it seems unlikely that *DBB* itself is under balancing selection because only synonymous variation was found at this locus. An assumption of nonsynonymous substitutions in parts of the gene which were not assayed would have to be made in order for balancing selection to be invoked. However, linkage disequilibrium (LD) with *DAB* could explain similar frequencies of *DBB* alleles even if they represent recent synonymous variants of the gene under stabilizing selection. Strong LD was observed between the *DAB* and *DBB* loci in the HCM and CAR. This indicates that the loci may be physically linked; linkage can be maintained over substantial chromosomal distances, at least in mammalian MHC (Miretti *et al*. 2005; Bonhomme *et al*. 2007). If so, *DBB* alleles which at the time of colonization happened to be associated with the *DAB* alleles selectively retained in the HCM and CAR populations could have hitchhiked to high frequency. Demographic processes may have further elevated LD in these populations (Pritchard & Przeworski 2001). On the other hand, the lack of significant LD among *DAB* and *DBB* alleles in the Sudetes may reflect either the break-up of linkage by recombination (SUD populations are closer to the refugium than populations in the HCM and CAR and thus probably older) or less power in LD tests due to high allelic variation in *Meal-DAB*. Linkage with *DAB* would oppose the loss of variation at *DBB* due to genetic drift and make the selective sweep hypothesis unlikely because of high allelic diversity in *DAB*. Therefore, we favour the hypothesis that *DBB* evolves under

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

purifying selection and its synonymous variants are maintained in high frequencies due to linkage with *DAB*.

The three allopatric groups of the Alpine newt in southern Poland (defined as populations in separate mountain ranges), although derived from a single refugium and thus of recent origin (Pabijan & Babik 2006), are differentiated in every part of the nuclear genome studied so far: graft rejection studies crudely assessed the overall MHC differentiation between regions (Rafiński 1970, 1974), the same pattern was subsequently revealed for allozyme (Pabijan *et al*. 2005) and microsatellite markers (Pabijan & Babik 2006). AMOVA showed significant genetic differentiation among groups for both pseudogene loci, but not for *Meal-DAB* or *DBB*. The lack of differentiation in expressed MHC II loci among regions may indicate that selection maintains a more uniform distribution of the MHC alleles.

On the other hand, drift is clearly an important factor shaping the variation at the *Meal-DAB* locus, as shown by a strong correlation between *Meal-DAB* and microsatellite allelic richness at the population level. The similar geographical gradient in both types of markers indicates that variation was reduced during the postglacial colonization of populations now located at the northeastern periphery of the distribution of this species. This may potentially affect the future evolutionary prospects of these populations (Hughes 1991; Hedrick 2001). In the extreme case, as few as three (all nonprivate) alleles (*Meal-DAB*01*, **02* and **04*) were retained in the northernmost populations (HCM). The retention of three highly divergent alleles is consistent with the hypothesis posing that individuals expressing divergent alleles coding for dissimilar MHC II molecules possess a selective advantage because they are able to trigger an immune response against a broader spectrum of antigens (Wakeland *et al*. 1990; Richman *et al*. 2001). The lack of a correlation between pseudogene and microsatellite variation may be explained by the generally limited pseudogene variation and low differentiation in allelic richness among populations. A correlation between variation of MHC and neutral markers has often been observed in various vertebrate species (e.g. Hedrick *et al*. 2001; Hinten *et al*. 2003; Miller & Lambert 2004; Babik *et al*. 2005; Aguilar & Garza 2006; Campos *et al*. 2006). Hence, random drift seems to significantly affect levels of MHC variation on the ecological time scale in many species.

The F_{ST} outlier analysis showed that the MHC loci do not deviate significantly from the genome-averaged level of population differentiation. All possible relationships between levels of genetic structure in neutral markers and MHC genes have been observed in various taxa. Higher differentiation in MHC than in neutral markers may indicate local adaptation in MHC (Landry & Bernatchez 2001; Aguilar & Garza 2006; Ekblom *et al*. 2007), whereas lower differentiation of MHC than neutral loci may occur when spatially uniform selection overrides neutral drift and homogenizes allele frequencies (Sommer 2003; Piertney & Oliver 2006). However, our F_{ST} outlier analysis showed that the MHC loci do not deviate significantly from the genome-averaged level of population differentiation. A lack of differences in the degree of MHC and neutral structure suggests that drift and gene flow are the predominant factors determining the differentiation of both classes of markers (Boyce *et al*. 1997; Hedrick *et al*. 2001).

Conclusions

In this first study of MHC II variation in a Urodele amphibian outside the *Ambystoma tigrinum/A. mexicanum* clade, we detected multiple, expressed and nonexpressed MHC II loci in the three allopatric Alpine newt population groups in southern Poland. The expressed loci differed strikingly in levels of variation which may reflect differentiation in selective regimes. We found high allelic variation and strong signatures of historical positive selection at the highly polymorphic *Meal-DAB* locus, whereas very low variation and lack of geographical differentiation of allele frequencies at the apparently linked *Meal-DBB* locus were observed. Pseudogene loci appear to have evolved neutrally. There was a clear geographical pattern in *Meal-DAB* allelic richness which correlated strongly with microsatellite allelic richness, indicating that in addition to positive selection, genetic drift has also shaped *Meal-DAB* variation during or after colonization in postglacial areas.

Acknowledgements

We would like to thank Maria Ogielska and Marcin Liana for help in collecting samples. The work was partially funded by the Alexander von Humboldt Foundation grant 3-Fokoop-POL/1022634.

References

Aguilar A, Garza JC (2006) A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). *Molecular Ecology*, **15**, 923–937.

Aguilar A, Roemer G, Debenham S *et al*. (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences, USA*, **101**, 3490–3494.

Aguilar A, Smith TB, Wayne RK (2005) A comparison of variation between a MHC pseudogene and microsatellite loci of the little greenbul (*Andropadus virens*). *BMC Evolutionary Biology*, **5**, 47.

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

Babik W, Durka W, Radwan J (2005) Sequence diversity of the MHC DRB gene in the Eurasian beaver (*Castor fiber*). *Molecular Ecology*, **14**, 4249–4257.

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

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.

Beebee TJC, Griffiths RA (2005) The amphibian decline crisis: a watershed for conservation biology? *Biological Conservation*, **125**, 271–285.

Berger L, Speare R, Daszak P *et al*. (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences, USA*, **95**, 9031–9036.

Bernatchez L, 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.

Bonhomme M, Blancher A, Jalil MF, Crouau-Roy B (2007) Factors shaping genetic variation in the MHC of natural non-human primate populations. *Tissue Antigens*, **70**, 398–411.

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006) Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, **60**, 383–389.

Bos DH, DeWoody JA (2005) Molecular characterization of major histocompatibility complex class II alleles in wild tiger salamanders (*Ambystoma tigrinum*). *Immunogenetics*, **57**, 775–781.

Bos DH, Waldman B (2006) Evolution by recombination and transspecies polymorphism in the MHC class I gene of *Xenopus laevis*. *Molecular Biology and Evolution*, **23**, 137–143.

Boyce WM, Hedrick PW, MuggliCockett NE *et al*. (1997) Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. *Genetics*, **145**, 421–433.

Briles WE, Stone HA, Cole RK (1977) Marek's disease — effects of B-histocompatibility alloalleles in resistant and susceptible chicken lines. *Science*, **195**, 193–195.

Brown JH, Jardetzky TS, Gorga JC *et al*. (1993) Three-dimensional structure of the human class-II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–39.

Campos JL, Posada D, Moran P (2006) Genetic variation at MHC, mitochondrial and microsatellite loci in isolated populations of Brown trout (*Salmo trutta*). *Conservation Genetics*, **7**, 515–530.

Carrington M, Bontrop RE (2002) Effects of MHC class I on HIV/SIV disease in primates. *Aids*, **16**, S105–S114.

Docherty DE, Meteyer CU, Wang J *et al*. (2003) Diagnostic and molecular evaluation of three iridovirus-associated salamander mortality events. *Journal of Wildlife Diseases*, **39**, 556–566.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at H-2 gene complex. *Nature*, **256**, 50–52.

Ekblom R, Saether SA, Jacobsson P *et al*. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology*, **16**, 1439–1451.

Ellegren H, Hartman G, Johansson M, Andersson L (1993) Major histocompatibility complex monomorphism and low-levels of DNA-fingerprinting variability in a reintroduced and rapidly expanding population of beavers. *Proceedings of the National Academy of Sciences, USA*, **90**, 8150–8153.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

García-Ramos G, Kirkpatrick M (1997) Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, **51**, 21–28.

Garner TWJ, Walker S, Bosch J *et al*. (2005) Chytrid fungus in Europe. *Emerging Infectious Diseases*, **11**, 1639–1641.

Garrigan D, Hedrick PW (2003) Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. *Evolution*, **57**, 1707–1722.

Głowaciński Z, Rafiński J (2003) *Atlas of the Amphibians and Reptiles of Poland Status — Distribution — Conservation*. Inspekcja Ochrony Srodowiska and Instytut Ochrony Przyrody PAN, Warszawa-Kraków.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hauswaldt JS, Stuckas H, Pfautsch S, Tiedemann R (2007) Molecular characterization of MHC class II in a nonmodel anuran species, the fire-bellied toad *Bombina bombina*. *Immunogenetics*, **59**, 479–491.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Hedrick PW, Kim TJ, Parker KM (2001) Parasite resistance and genetic variation in the endangered *Gila* topminnow. *Animal Conservation*, **4**, 103–109.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hinten G, Harriss F, Rossetto M, Braverstock PR (2003) Genetic variation and island biogeography: microsatellite and mitochondrial DNA variation in island populations of the Australian bush rat, *Rattus fuscipes greyii*. *Conservation Genetics*, **4**, 759–778.

Hood GM (2006) POPTOOLS *version 2.7.5*. Available from http://www.cse.csiro.au/poptools.

Hughes AL (1991) MHC polymorphism and the design of captive breeding programs. *Conservation Biology*, **5**, 249–251.

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

Kaufman J, Salomonsen J, Flajnik M (1994) Evolutionary conservation of MHC class I and class II molecules — different yet the same. *Seminars in Immunology*, **6**, 411–424.

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

Klein J (1986) *The Natural History of the Major Histocompatibility Complex*. Wiley & Sons, New York.

Klein J, Bontrop RE, Dawkins RL *et al*. (1990) Nomenclature for the major histocompatibility complexes of different species — a proposal. *Immunogenetics*, **31**, 217–219.

Ko WY, David RM, Akashi H (2003) Molecular phylogeny of the *Drosophila melanogaster* species subgroup. *Journal of Molecular Evolution*, **57**, 562–573.

Kumanovics A, Takada T, Lindahl KF (2003) Genomic organization of the mammalian MHC. *Annual Review of Immunology*, **21**, 629–657.

Landry C, Bernatchez L (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **10**, 2525–2539.

Langefors A, Lohm J, Grahn M, Andersen O, von Schantz T (2001) Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 479–485.

Laurens V, Chapusot C, Ordonez MD *et al*. (2001) Axolotl MHC class II beta chain: predominance of one allele and alternative splicing of the beta 1 domain. *European Journal of Immunology*, **31**, 506–515.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation? *Conservation Biology*, **9**, 753–760.

Lips KR, Brem F, Brenes R *et al*. (2006) Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proceedings of the National Academy of Sciences, USA*, **103**, 3165–3170.

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

Maynard Smith J (1992) Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, **34**, 126–129.

Mikko S, Roed K, Schmutz S, Andersson L (1999) Monomorphism and polymorphism at MHC DRB loci in domestic and wild ruminants. *Immunological Reviews*, **167**, 169–178.

Miller HC, Lambert DM (2004) Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Molecular Ecology*, **13**, 3709.

Miretti MM, Walsh EC, Ke XY *et al*. (2005) A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *American Journal of Human Genetics*, **76**, 634–646.

Morgan JAT, Vredenburg VT, Rachowicz LJ *et al*. (2007) Population genetics of the frog-killing fungus *Batrachochytrium dendrobatidis*. *Proceedings of the National Academy of Sciences, USA*, **104**, 13845–13850.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, **3**, 418–426.

Nei M, Hughes AL (1992) Balanced polymorphism and evolution by the birth-and-death process in the MHC loci. In: *Proceedings of the 11th Histocompatibility Workshop and Conference* (eds Tsuji K Aizawa M Sasazuki T), pp. 27–38. Oxford University Press, Oxford, UK.

Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, **39**, 121–152.

O'Brien SJ, Evermann FF (1988) Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology & Evolution*, **3**, 254–259.

Ohta Y, Goetz W, Hossain MZ, Nonaka M, Flajnik MF (2006) Ancestral organization of the MHC revealed in the amphibian *Xenopus*. *Journal of Immunology*, **176**, 3674–3685.

van Oosterhout C, Joyce DA, Cummings SM (2006) Evolution of MHC class IIB in the genome of wild and ornamental guppies, *Poecilia reticulata*. *Heredity*, **97**, 111–118.

Pabijan M, Babik W (2006) Genetic structure in northeastern populations of the Alpine newt (*Triturus alpestris*): evidence for post-Pleistocene differentiation. *Molecular Ecology*, **15**, 2397–2407.

Pabijan M, Babik W, Rafinski J (2005) Conservation units in north-eastern populations of the Alpine newt (*Triturus alpestris*). *Conservation Genetics*, **6**, 307–312.

Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology*, **265**, 218–225.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics*, **37**, 637–669.

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

Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW (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, Buckley TR (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.

Pounds JA, Bustamante MR, Coloma LA *et al*. (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, **439**, 161–167.

Pritchard JK, Przeworski M (2001) Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics*, **69**, 1–14.

Rafiński JN (1970) Population studies of homograft rejection in the newt *Triturus alpestris* (Laur.). *Bulletin de l'Academie Polonaise Des Sciences*, **18**, 425–428.

Rafiński JN (1974) Studies on the genetic structure of the Alpine newt *Triturus alpestris* (Laur.), populations. *Acta Biologica Cracoviensia*, **27**, 51–68.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population-genetics software for exact tests and ecumenicism. *Journal of Heredity*, **27**, 248–249.

Reusch TBH, Langefors A (2005) Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback *Gasterosteus aculeatus*. *Journal of Molecular Evolution*, **61**, 531–U545.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Richman AD, Herrera LG, 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.

Richman AD, Herrera G, Reynoso VH, Mendez G, Zambrano L (2007) Evidence for balancing selection at the DAB locus in the axolotl, *Ambystoma mexicanum*. *International Journal of Immunogenetics*, **34**, 475–478.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Sammut B, Du Pasquier L, Ducoroy P *et al*. (1999) Axolotl MHC architecture and polymorphism. *European Journal of Immunology*, **29**, 2897–2907.

Shriner D, Nickle DC, Jensen MA, Mullins JI (2003) Potential impact of recombination on sitewise approaches for detecting positive natural selection. *Genetical Research*, **81**, 115–121.

Siddle HV, Kreiss A, Eldridge MDB *et al*. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. *Proceedings of the National Academy of Sciences, USA*, **104**, 16221–16226.

Snell GD (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 (2003) Effects of habitat fragmentation and changes of dispersal behaviour after a recent population decline on the genetic variability of noncoding and coding DNA of a monogamous Malagasy rodent. *Molecular Ecology*, **12**, 2845–2851.

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

Sotiropoulos K, Eleftherakos K, Dzukic G *et al*. (2007) Phylogeny and biogeography of the alpine newt *Mesotriton alpestris* (Salamandridae, Caudata), inferred from mtDNA sequences. *Molecular Phylogenetics and Evolution*, **45**, 211–226.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Stuart SN, Chanson JS, Cox NA *et al*. (2004) Status and trends of amphibian declines and extinctions worldwide. *Science*, **306**, 1783–1786.

Sullivan J, Joyce P (2005) Model selection in phylogenetics. *Annual Review of Ecology Evolution and Systematics*, **36**, 445–466.

Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*, **124**, 967–978.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA 4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Thursz MR, Thomas HC, Greenwood BM, Hill AVS (1997) Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nature Genetics*, **17**, 11–12.

Tournefier A, Laurens V, Chapusot C *et al*. (1998) Structure of MHC class I and class II cDNAs and possible immunodeficiency linked to class II expression in the Mexican axolotl. *Immunological Reviews*, **166**, 259–277.

Wakeland EK, Boehme S, She JX *et al*. (1990) Ancestral polymorphisms of MHC class-II genes — divergent allele advantage. *Immunologic Research*, **9**, 115–122.

Willi Y, van Buskirk J, Hoffmann A (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 433–458.

Wilson DJ, McVean G (2006) Estimating diversifying selection and functional constraint in the presence of recombination. *Genetics*, **172**, 1411–1425.

Wittzell H, Bernot A, Auffray C, Zoorob R (1999) Concerted evolution of two MHC class II B loci in pheasants and domestic chickens. *Molecular Biology and Evolution*, **16**, 479–490.

Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences*, **13**, 555–556.

Zhang JZ, Nielsen R, Yang ZH (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, **22**, 2472–2479.

W. Babik uses molecular tools to study historical processes at and below the species level. M. Pabijan is interested in the phylogeny, population genetics and conservation of amphibians. J. Radwan is a Professor at the Institute of Nature Conservation, Polish Academy of Sciences and at the Institute of Environmental Sciences, Jagiellonian University. His interests focus on sexual selection, its evolutionary-genetic background and conservation genetic implications.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd