Molecular Ecology (2009) doi: 10.1111/j.1365-294X.2008.04057.x

Long-term survival of a urodele amphibian despite depleted major histocompatibility complex variation

W. BABIK,*† M. PABIJAN,‡ J.W. ARNTZEN,§ D. COGĂLNICEANU,¶ W. DURKA* and J. RADWAN**
*Department of Community Ecology, Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Str 4, 06120 Halle/Saale, Germany, †Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland, ‡Department of Comparative Anatomy, Institute of Zoology, Jagiellonian University, Ingardena 6, 30-060 Kraków, Poland, §National Museum of Natural History - Naturalis, PO Box 9517, 2300 RA Leiden, The Netherlands, ¶Faculty of Natural Sciences, University Ovidius Constanta, Bvd Mamaia 124, 900527 Constanta, Romania, **Institute of Nature Conservation, Polish Academy of Sciences, Al. Mickiewicza 33, 31-120 Kraków, Poland*

Abstract

Depletion of polymorphism at major histocompatibility complex (MHC) genes has been hypothesized to limit the ability of populations to respond to emerging pathogens, thus putting their survival at risk. As pathogens contribute substantially to the global amphibian decline, assessing patterns of MHC variation is important in devising conservation strategies. Here, we directly compare levels of MHC class II and neutral variation between multiple populations of the great crested newt (*Triturus cristatus*) from refugial (REF: Romania) and postglacial expansion (PGE: Germany, Poland and UK) areas. REF populations harboured high levels of adaptive variation (24 expressed alleles), exhibiting clear signatures of historical positive selection, which points to the overall importance of MHC class II variation in this species. On the other hand, PGE populations were extremely depauperate (two alleles) but nevertheless have survived for *c*. 10 000 years, since the postglacial expansion. Variation in putative MHC class II pseudogenes, microsatellites and allozymes also showed a significant southern richness–northern purity pattern. The populations in the postglacial expansion area thus provide the clearest example to date of the long-term survival of populations in which MHC variation, historically under positive selection, has been depleted.

Keywords: amphibian decline, glacial refugia, MHC class II, MHC pseudogenes, positive selection, *Triturus cristatus*

Received 21 September 2008; revision revised 12 November 2008; accepted 26 November 2008

Introduction

The role of genetic factors in extinction has been subject to much debate, stirred by Lande's (1988) suggestion that environmental factors should drive species to extinction before genetic factors could have any impact. A recent meta-analysis (Spielman *et al*. 2004) showed that genetic variation in endangered species is substantially lower than in unthreatened species. A well-established consequence of decreased genetic variation is inbreeding depression, a decline in fitness caused by the increase in homozygosity, revealing deleterious mutations at many loci (Hedrick 2001). Computer simulations assuming a level of inbreeding depression found in natural mammalian and avian populations showed that it can indeed decrease the time to extinction by 37% on average (O'Grady *et al*. 2006). Another important risk factor may be associated with the depletion of variation at adaptive loci, such as genes of the major histocompatibility complex (MHC) encoding proteins that play an essential role in the adaptive immunity of gnathostome vertebrates (Klein 1986). Most of the variation in the extremely polymorphic MHC is concentrated at antigen binding sites (ABS) — amino acid residues specifically involved in the recognition and binding of foreign peptides (Garrigan & Hedrick 2003). Empirical studies have documented associations between MHC genotypes and susceptibility to diseases (e.g. Briles *et al*. 1977; Thursz *et al*. 1997; Bonneaud *et al*. 2006). Two mechanisms have generally been invoked in the explanation of high MHC polymorphism:

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

frequency-dependent selection (Snell 1968) and heterozygote advantage (Doherty & Zinkernagel 1975).

It has been suggested that species or populations with low MHC polymorphism may be particularly vulnerable to infection, and consequently may face a higher risk of extinction (Hughes 1991; Hedrick 2001). However, empirical data are equivocal in this respect. There are reports of species possessing relatively high MHC variability despite a lack of or low genetic variation in other parts of the genome, an indication that selection has retained variation in the MHC (Aguilar *et al*. 2004; Jarvi *et al*. 2004). On the other hand, in several mammal species, MHC variation is limited or absent, and despite this, populations do not show signs of increased susceptibility to diseases (Ellegren *et al*. 1993; Mikko *et al*. 1999; Babik *et al*. 2005). There are also examples of a correlation between MHC variation and variation in neutral markers (Hedrick *et al*. 2001; Seddon & Ellegren 2004; Campos *et al*. 2006), indicating the role of stochastic events in shaping MHC diversity. However, as most of these examples concern species which have undergone recent population bottlenecks, the lack of current selection on MHC does not imply that MHC variation will not affect the long-term survival of the species (Radwan *et al*. 2007).

A broader timescale is provided by species which have experienced reductions in effective population sizes, and a consequent reduction in genetic variation, in the temperate and boreal regions during the Pleistocene glacial–interglacial cycles (Hewitt 2004; Babik *et al*. 2005). European and American subspecies of moose show very little MHC diversity, which may be due to a bottleneck that predated the divergence of these subspecies some 100 000 years ago (Mikko & Andersson 1995). However, the MHC diversity in moose preceding the postulated bottleneck is unknown, and this is also the case for other species in which similar scenarios have been proposed (Mikko *et al*. 1999). Here, we investigate patterns of MHC variation in an amphibian species in which the vast majority of current range is of postglacial origin, but refugial populations, harbouring substantial genetic variation, are still present.

Amphibians have been declining worldwide at an unprecedented rate (Stuart *et al*. 2004; 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 have been invoked in this decline, but their relative importance may differ for various taxonomic groups and geographic areas (Beebee & Griffiths 2005). There is a substantial body of evidence showing that infectious diseases, possibly aggravated by climate change, are a major factor contributing to declines in many areas (Pounds *et al*. 2006). 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, as reflected in part by MHC allelic diversity. Information on MHC variation in urodele amphibians is particularly scarce, being restricted to two closely related ambystomatid salamanders (Laurens *et al*. 2001; Bos & DeWoody 2005; Bos *et al*. 2008) and one species of newt (Babik *et al*. 2008).

Here, we describe MHC class II variation in the great crested newt, *Triturus cristatus* (Salamandridae). This species is widely and rather continuously distributed in Eurasia. The glacial refugia of the great crested newt were probably located at the extreme southern fringe of its present-day distribution, i.e. in Hungary, Romania or southwestern Ukraine (Wallis & Arntzen 1989). The species shows homogeneity in allozymes and mtDNA over most of its range with some mtDNA variation in southern Romania and adjacent Serbia (Arntzen 2003). Thus, sampling populations in the putative refugial and postglacial expansion areas and simultaneous analysis of variation at neutral loci should provide insight into the relative contributions of selection and historical demographic events in shaping MHC variation.

We found clear evidence that historically, MHC class II variation in the great crested newt has been shaped by positive selection and that ample variation is still present in refugial populations at the southern fringe of its current distribution. In contrast, almost no MHC variation is present in populations formed after the postglacial expansion of this species, constituting the majority of its contemporary range. The populations from the postglacial expansion area thus provide the clearest, to date, example of the long-term survival of populations whose MHC variation, historically under positive selection, has been depleted.

Materials and methods

Sampling

We collected tail tips of adult or larval newts from seven populations (Table 1; Fig. 1). Three of these comprise widely spaced populations from postglacial expansion areas (PGE) in southern Poland (POL, localities: Inwald, 49°51′N, 19°23′E; Limanowa, 49°43′N, 20°25′E); Germany (GER: Krefeld, 51°20′N, 6°35′E) and central England (UK, localities: Gaddesby, 52°42′N, 0°59′W; Peterborough 52°35′N, 0°15′W; Canterbury, 51°17′N, 1°05′E). Four samples were taken from the postulated refugial area (REF) in Romania (VID: Videle, 44°16′N, 25°31′E; TIR: Tirgoviste, 44°56′N, 25°27′E; MER: Meri, 45°10′N, 23°24′E; PAC: Păclísa, 45°34′N, 22°53′E).

Primer development

The procedure of obtaining the species-specific primers is detailed in Babik *et al*. (2008) and only briefly outlined here. We targeted MHC II second exon because it encodes amino acids involved in antigen presentation and is the most

Table 1 The great crested newt sample sizes (*N*) and descriptive statistics for the MHC class II loci *Trcr-DAB* and *Trcr*-ψ, 5 microsatellite loci and 40 allozyme loci. $N_{alleles}$ is mean number of alleles with standard deviations (for microsatellite and allozyme loci); $N_{private}$ is number of private microsatellite alleles (confined to a single population); N_{loci} is the number and percentage of polymorphic allozyme loci; *R*, allelic richness adjusted to the minimum sample size of five individuals (for MHC and microsatellite data sets) or 10 individuals per population (allozyme data set); H_E, mean expected heterozygosities with standard deviations. PGE, postglacial expansion area; REF, postulated refugial area

		UK: Gaddesby	UK: Peterborough Canterbury	GER: Krefeld	POL: Inwald	POL: Limanowa	**Overall PGE**	MER: Meri	PAC: Păclísa	TIG: Tirgoviste	VID: Videle	**Overall REF**
MHC-II	N	20	—	24	19	—	**63**	5	5	19	8	**37**
	$N_{alleles}$ *Trcr-DAB*	2	—	2	2	—	**2**	5	4	15	7	**24**
	R Trcr-DAB	1.9	—	2	1.4	—	**1.8**	5	4	7.1	5.2	**5.3**
	$N_{alleles}$ *Trcr*-ψ	3	—	3	4	—	**4**	7	6	18	12	**18**
	R Trcr-ψ	3	—	3	3.5	—	**3.2**	7	6	12.7	9.9	**8.9**
microsatellites	N	20	—	24	19	—	**63**	5	5	19	8	**37**
	$N_{alleles}$	7.4 ± 1.02	—	7.2 ± 2.48	7.2 ± 1.94	—	**7.26 ± 1.91**	5.6 ± 0.49	5.4 ± 1.02	9.6 ± 1.85	9.2 ± 1.72	**7.45 ± 2.40**
	$N_{private}$	2	—	3	2	—	**7**	5	1	17	7	**30**
	H_E	0.73 ± 0.15	—	0.76 ± 0.07	0.78 ± 0.08	—	**0.76 ± 0.12**	0.85 ± 0.08	0.83 ± 0.06	0.87 ± 0.05	0.89 ± 0.05	**0.86 ± 0.06**
	R	5	—	4.5	4.9	—	**4.8**	5.6	5.4	7	6.9	**6.2**
allozymes	N	—	10	—	—	10	**20**	—	—	10	11	**21**
	N_{loci}	—	6 (15.0%)	—	—	3 (7.5%)	**7 (17.5%)**	—	—	10 (25.0%)	10 (25.0%)	**12 (30.0%)**
	$N_{alleles}$	—	1.15 ± 0.37	—	—	1.07 ± 0.27	**1.2 ± 0.46**	—	—	1.22 ± 0.51	1.35 ± 0.66	**1.40 ± 0.67**
	H_e	—	0.02 ± 0.07	—	—	0.02 ± 0.07	**0.02 ± 0.07**	—	—	0.07 ± 0.16	0.07 ± 0.16	**0.08 ± 0.17**
	R	—	1.15	—	—	1.07	**1.11**	—	—	1.27	1.34	**1.31**

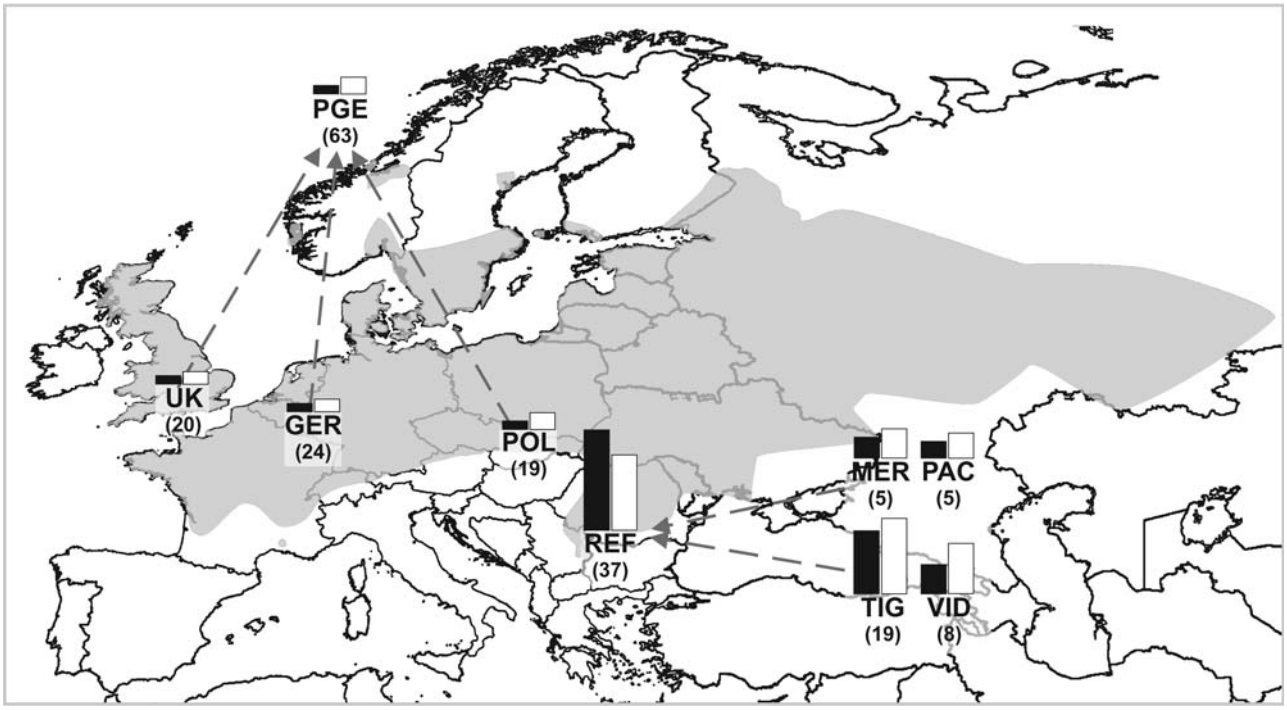

Fig. 1 Populations of the great crested newt studied for MHC II variation. Grey, geographic distribution of *Triturus cristatus*; black bars, number of alleles at the putative expressed locus *Trcr-DAB*; white bars, number of alleles at putative MHC pseudogene loci. PGE, postglacial expansion area; REF, postulated refugial area. Sample sizes in parentheses. Coordinates and further details including populations sampled for allozymes in text and Table 1.

polymorphic part of MHC II genes. Partial MHC II second exon sequences were obtained using primers designed on the basis of *Ambystoma* sequences (Bos & DeWoody 2005). These partial sequences were used to design primers in conserved parts inside the exon which enabled us to obtain sequences of both ends of the exon through vectorette polymerase chain reaction (PCR). In the vectorette PCR approach (Ko *et al*. 2003), 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. By 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 between the specific primer and the RE recognition site, i.e. outside of the region of known sequence. Multiple REs are usually used to ensure that a fragment of sufficient length is obtained. In the case of large, complex genomes, a nested approach, 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.

Two relatively divergent sequences were detected with the vectorette PCR technique. We designed the primer pair TrMHCII10F (5′-GAGTGTCASTTCSTGAACGGCTCTCAG-3′) and TrMHCII5R (5′-CTCACGCYTCCGSTGCTCCATG-3′) to amplify both sequences. Initially, we also tried forward primers TrMHCII1F (5′-TCTCTCCGCAGTGGACTTCGTG-3′) and TrMHCII5F (5′-CTCCGCAGTGGACTTCGTGA-3′) spanning the intron–exon boundary, and thus obtained almost the entire sequence of the exon. However, this primer combination did not amplify all sequences. Therefore, we employed an internal left primer TrMHCII10F located in a conserved part of the exon in an effort to obtain a full representation of the MHC II variation.

MHC genotyping

Total DNA was extracted from tail tips using Genomic Mini kit (A&A Biotechnology). We amplified *Triturus cristatus* MHC in 20 μL reactions containing 1 U of *Taq* polymerase (Fermentas), 1.5 mM $MgCl_2$, 1 μM of labelled primers and 10–100 ng of genomic DNA. The cycling scheme was as follows: 3 min at 94 °C followed by 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 isolating allele sequences. First, the fluorescent single-strand conformational polymorphism (SSCP) analysis was performed as described earlier (Babik *et al*. 2008). Only a few SSCP patterns were detected in PGE populations. Individuals representing unique SSCP patterns (two per pattern where available) were amplified as above but with unlabelled primers. Amplicons were run on SSCP GMA gels (Elchrom Scientific), individual SSCP bands were excised, re-amplified

and sequenced as described in Babik *et al*. (2008). Samples from the REF populations exhibited very complex SSCP patterns and isolating alleles through amplification from SSCP bands proved ineffective because multiple sequences amplified from the majority of SSCP bands. Thus, we applied two additional approaches to characterize the allelic variation in these populations. In the first approach, PCR products were cloned using the pGEM-T Easy vector system (Promega) to determine the sequence of individual alleles. Recombinant clones were detected by blue/white screening, plasmid DNA minipreps were prepared with the Plasmid Mini Kit (A&A Biotechnology) or inserts were directly amplified with M13F and M13R primers in colony PCR. Multiple clones (22.3 per individual on average) containing inserts were sequenced. Sequences were checked and aligned in SeqScape 2.1 (ABI). Additionally, we used primers TrMHCII5F-TrMHCII5R which specifically amplified the majority of the expressed alleles (see Results). This allowed for a control of sampling error in the cloning approach and also discriminated between alleles *Trcr-DAB*01* and *Trcr-DAB*02* which differ only in the 5′ part, not covered by TrMHCII10F-TrMHCII5R amplicons. PCR was performed as described above but with an annealing temperature of 63 °C; PCR products were sequenced directly.

We regarded only sequences derived from 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 some cases, bands showing heterozygous sequences (two alleles amplified from a single band) were used to validate alleles obtained by direct sequencing of PCR products or cloning. Other sequence variants, which may have represented polymerase errors and PCR- or cloning-generated recombinants were excluded from the analyses.

Expression analysis

In order to validate the expression of the assayed loci in the great crested newt, and to check for a possible developmental shift in the expression pattern, we obtained RNA from the spleen of one adult and from the gills of one larva from PGE POL and seven larvae from REF population TIG. Immediately after excision, the tissues were preserved in the RNAlater stabilization reagent (QIAGEN).

RNA was extracted with the RNeasy kit (QIAGEN) from ≤ 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 U/μL) and 1 μL RT and 6 μL of RNase free water. The reaction was incubated at 37 °C for 60 min. The same primers (TRMHCII10F and TrMHCII5R) were used for amplification as described above. PCR products from cDNA were directly sequenced or were run on SCCP GMA gels (Elchrom Scientific) alongside those from genomic DNA. SSCP patterns were visually compared and all clearly visible bands were excised, re-amplified and sequenced.

Microsatellite and allozyme analysis

We scored variation at five microsatellite loci (Krupa *et al*. 2002): *Tcri*13 (PET), *Tcri*35 (FAM), *Tcri*29 (FAM), *Tcri*36 (NED), *Tcri*43 (PET), the names in parentheses referring to the fluorescent dye used to label the F primer for each locus. Note that for locus *Tcri*43, we used other primers than reported in the original paper: F: AAGTAACTGAAA GATAACATGTAG and R: GTTTCTATTCATTTTTGTTAC GCAC (R. Jehle, personal communication).

All loci were amplified in a single multiplex 10 μL reaction containing 5.0 μL of the 2 × Multiplex PCR Master Mix (QIAGEN), 0.2 μM of each primer and 10–100 ng of genomic DNA. The following cycling scheme was used: 95 °C for 15 min, followed by 35 cycles of 30 s at 94 °C, 90 s at 55 °C and 60 s at 72 °C and a final extension step of 30 min at 60 °C. One microlitre of PCR product was mixed with 14 μL of Hi-Di formamide (ABI) and 0.3 μL of GeneScan 500 LIZ size standard (ABI), denatured for 2 min at 95 °C, immediately cooled on ice for 2 min and electrophoresed on an ABI 3130xl. GeneMapper was used for genotyping.

In addition to microsatellites, we analysed 27 protein systems representing 40 presumptive gene loci by means of starch- and polyacrylamide-gel electrophoresis. For laboratory protocols and nomenclature, see Arntzen (2001).

Statistical, population genetic and phylogenetic 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, were computed in MEGA3 (Kumar *et al*. 2004). The phylogenetic relationships among the great crested newt MHC class II sequences were reconstructed using an *A. tigrinum* allele as an outgroup under the Bayesian approach with MrBayes 3.1 (Ronquist & Huelsenbeck 2003) and the maximum-likelihood (ML) approach with PHYML 2.4. (Guindon & Gascuel 2003). The likelihood settings corresponded to the GTR + Γ model with the parameter values estimated from the data. In the Bayesian analysis, priors were set to default values, two independent runs of four Metropolis coupled Markov chain Monte Carlo each (three of them 'heated', temperature = 0.20) were performed for 4×10^6 generations and sampled every 1000 generations. The first 500 trees were discarded as burn-in, resulting in 7000 sampled trees used to calculate the posterior probability (PP) of each bipartition. Robustness of the ML tree was assessed with 1000 bootstrap replicates.

We checked for signatures of recombination in our data set using three methods. Two of these, GeneConv (Padidam *et al*. 1999) and MaxChi2 (Maynard Smith 1992) performed very well in an assessment of 14 recombination detection methods (Posada 2002). They are implemented in the RDP 3 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/).

The impact of historical selection on the MHC sequences was assessed through the *Z*-test of selection in MEGA and by fitting various models of codon evolution available in PAML (Yang 1997). These were: (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). We analysed the locations of codons under positive selection in relation to the residues whose side chains form the antigen-binding groove (ABS), as inferred from the human MHC II molecule structure by Brown *et al*. (1993), as well as pocket-forming residues which determine the specificity of antigen binding (Stern *et al*. 1994). Positively selected codons were identified through the Bayes empirical Bayes procedure (Zhang *et al*. 2005).

The mean number of MHC alleles corrected for sample size, i.e. the allelic richness, was estimated for each population and overall in FSTAT 2.9.3.2 (Goudet 2001) or, when assigning sequences to individual loci was not possible, through randomization tests in PopTools (Hood 2006). We tested for differences in the allelic richness between PGE and REF groups using the Mann–Whitney test. Deviations from Hardy–Weinberg equilibrium in each population for both the microsatellite and allozyme loci were assessed by applying the exact tests in GenePop 3.3 (Raymond & Rousset 1995). *F*-statistics (Weir & Cockerham 1984) and pairwise, multilocus permutation tests of population differentiation were calculated in FSTAT. The sequential Bonferroni procedure was applied where appropriate in order to adjust the significance level to $\alpha = 0.05$. The allelic richness was estimated for each population and overall in FSTAT. An AMOVA test in Arlequin was carried out for each marker class in order to investigate how variation is partitioned within and among colonized vs. refugial populations.

Results

Sequence diversity and phylogenetic analysis

We detected 43 unique sequences (GenBank Accession nos FJ447986–FJ448028) showing homology to the MHC class II second exon in seven populations of the great crested newt. Phylogenetic analysis identified two groups, A and B, with a posterior probability (PP) of 0.99 and 0.92, and bootstrap support (BS) 97 and 40%, respectively (Fig. 2). Only one sequence (*Trcr-DBB*01*) falls outside these groups. Eight A alleles have a deletion of three nucleotides (a single codon, position 78, ABS, Table 2).

MHC expression pattern and the number of MHC II loci

The analysis of cDNA from adult spleen and larval gills (which contain substantial amounts of blood) indicated that not all sequences present in genomic DNA (gDNA) are expressed in these tissues and that the expression pattern

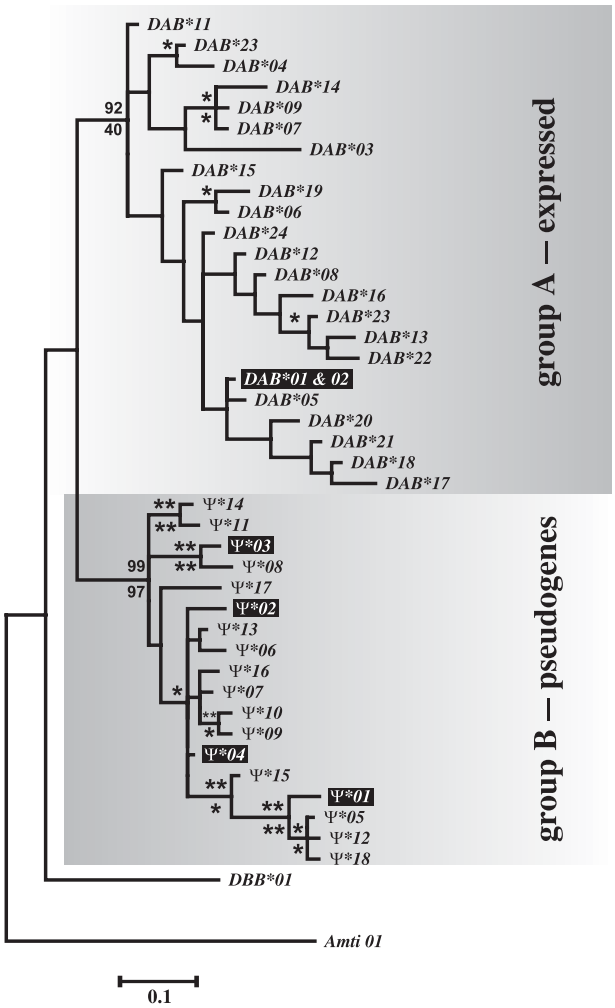

Fig. 2 Phylogenetic relationships of the great crested newt MHC class II alleles. The 50% majority rule tree from the Bayesian analysis is shown. Top branches, posterior probabilities (PP) from the Bayesian analysis; *, PP > 0.95; **, PP = 1; bottom branches, bootstrap support (BS) from the maximum-likelihood analysis: * BS > 70%; ** BS > 95%. Alleles in the black boxes were found both in the PGE and REF populations, the remaining alleles were found only in REF. The tree was rooted using the *Ambystoma tigrinum Amti 01* sequence.

Table 2 Amino acid alignment of the great crested newt MHC II *Trcr-DAB* alleles; amino acids numbered according to their position in the human MHC II beta chain. Shaded columns represent putative antigen binding sites (ABS) (Brown *et al.* 1993); sign '-' denotes codon deletion; 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)

```
                30        40        50        60        70        80
..|....|....|....|....|....|....|....|....|....|....|....|....|...
   x x x           x   x       x  x        xx xx          x
   * * *           *           *           *             *
     * *                       *           *             *
```

*Trcr-DAB*01&02*	RVRLLARWSYNRQQFVHFDSDVGYFVADTEFGEADARYWNSDEAVLEQVRAEVDRFCRYNYGIYEG
*Trcr-DAB*03*	...Y.Q.Y...Q...L.....YE.....P.VKK.....Y...L.G.....H.....
*Trcr-DAB*04*	...F.R.L.....Y.....Y...L..A...Y.....
*Trcr-DAB*05*	.....R...E.....
*Trcr-DAB*06*	...F...L.....R...E.....Y...L..A...-.....
*Trcr-DAB*07*	...Y.Q.F...Q.....P.....YMK.L.G.....L..
*Trcr-DAB*08*	.....PT.....V.....A...-.....
*Trcr-DAB*09*	...Y.Q.F...Q.....P.....YMK.L..A.....L..
*Trcr-DAB*10*	.....Y.....PT.....V.....L..
*Trcr-DAB*11*	...F.D.Y.....Y.....L..
*Trcr-DAB*12*	...F...L.....PT.....A...-.....
*Trcr-DAB*13*	.....P.Y.....Y.....PT.....F.....L..
*Trcr-DAB*14*	...Y.Q.F...Q.....PT.....N..YMK.M.....
*Trcr-DAB*15*	...F...L.....Y.....Y.....TA...-.....L..
*Trcr-DAB*16*	...F.....PT.....V.....DA..A.....L..
*Trcr-DAB*17*	...F.Q.Y.....R...E.....V.....DA.TA...-.....
*Trcr-DAB*18*	...F...L.....R...E.....V.....DA.TA...Y.....
*Trcr-DAB*19*	...F...L.....R.....PT.....Y...L..A...-.....
*Trcr-DAB*20*	.....V.....DA.TA...-.....
*Trcr-DAB*21*	.....R...E.....V.....DA..A...Y.....
*Trcr-DAB*22*	...F.R.L.....PT.....V.....L..
*Trcr-DAB*23*	...F.R.L.....Y.....Y.....L..
*Trcr-DAB*24*	...F...L.....L..A...-.....

does not change throughout the life cycle, i.e. the same alleles are expressed before and after metamorphosis. Only group A sequences were found in cDNA (alleles *Trcr-DAB*01-04, 06, 07, 09, 10, 23*) and in each case, all sequences from group A found in a given individual's gDNA were also found in cDNA. The presence of internal stop codons in three group B alleles (*Trcr-ψ*03, 08, 17*) suggests that they represent nonfunctional pseudogenes rather than non-classical MHC loci.

The number of unique sequences per individual ranged from three to eight. Most of the newts from the PGE populations harboured four sequences, whereas REF newts displayed from three to eight unique sequences. This indicates that up to at least four loci were amplified in the great crested newt. This hypothesis is supported by the pattern seen in the northern PEG populations in which most individuals display the same three sequences from the B group, *Trcr-ψ*01-03*, indicative of three almost monomorphic loci, and one or two sequences from the A group. However, no individual exhibited more than two sequences from the A group. Therefore, we assume that these sequences represent a single expressed locus which we name *Trcr-DAB*. In accordance with the very high allelic diversity in the REF population, most individuals (81%) from these populations were heterozygous; allele frequencies are given in the Table S1, Supporting Information. Genotype frequencies of this locus did not depart significantly from Hardy–Weinberg expectations in any of the populations. Assignment of sequences from the B group to individual loci was not possible on the basis of the phylogenetic clustering of alleles. The presence of sequences *Trcr-ψ*01-03* in all assayed newts from the PGE populations suggests that these represent distinct loci; however, this does not facilitate the allocation of other B group sequences to loci. Allele *Trcr-DBB*01*, detected in two individuals from Romania, may represent a distinct locus, but its expression status remains unknown.

Testing for selection and recombination

The average nucleotide, amino acid, nonsynonymous and synonymous divergences are shown in Table 3 and Table

Table 3 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 neutrality. dS and dN values are given as percentages per site; * denotes significant *P* value

Sites	dN	dS	*Z*	*P*
All sequences				
All	13.5 (2.4)	15.7 (3.9)	–0.532	0.596
ABS	23.9 (7.1)	21.7 (9.5)	0.170	0.866
Non-ABS	9.5 (2.1)	13.9 (3.9)	–1.107	0.271
Group A (putative locus *Trcr-DAB*)				
All	9.9 (2.4)	5.6 (2.1)	1.587	0.115
ABS	18.9 (6.4)	4.2 (3.1)	2.892	0.003*
Non-ABS	6.3 (2.2)	6.3 (2.9)	0.004	0.997
Group B (putative pseudogenes)				
All	6.5 (1.2)	12 (3.4)	–1.650	0.102
ABS	5.6 (1.8)	17.8 (11.7)	–1.014	0.313
Non-ABS	6.9 (1.5)	10.3 (3.6)	–0.945	0.346

S2, Supporting Information. Their inspection reveals contrasting patterns of variation between groups A and B. If all sequences are considered, the mean values represent an average between the opposite tendencies seen in each group and as such are misleading. Therefore, we consider A and B groups separately in most analyses below.

The overall nucleotide and amino acid divergence was three times higher at the ABS than non-ABS sites of A group alleles. The dN/dS ratio was significantly elevated at antigen binding sites (ABS, $Z = 2.89$ $P = 0.003$, *Z*-test of neutrality), but not at the non-ABS sites ($Z = 0.004$, $P = 0.997$; Table 3). Accordingly, the M8 model of codon evolution, which assumes that a fraction of the sites is under positive selection, explained the data better than one dN/dS category or nearly neutral models (Table 4). Of seven codons identified as positively selected under the M8 model by the Bayes empirical Bayes procedure (posterior probability, PP > 95%), four were at putative ABS ($P = 0.086$, Fisher's exact test). When codons with a PP of being under positive selection at > 50% were included, as many as seven of 12 were located at ABS ($P = 0.014$, Fisher's exact test). Among triplets coding for pocket-forming residues, positively selected ones were also over-represented (for PP > 95% $P = 0.012$, for PP > 50% $P = 0.003$).

The pattern of variation in group B was notably different. Neither nucleotide nor amino acid divergence differed between ABS and non-ABS sites. Synonymous divergence at the ABS was more than four times higher than nonsynonymous divergence, indicative of purifying rather than positive selection in the B group (Table 3). A model with a single dN/dS ratio (estimated value 0.190) fit the data best, suggesting that there was no variation in dN/dS among sites and that purifying selection may indeed have been operating (Table 4).

GeneConv and MaxChi2 did not detect recombination in any group. GARD method suggested a single recombination point located at nucleotide position 93 in group A, whereas it did not detect any recombination in group B.

Table 4 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_1, proportion of positively selected sites ($\omega > 1$); ω_1, estimated value of ω for sites under positive selection; p_{beta} and q_{beta}, parameters of beta distribution; δAIC, the difference between the value of the Akaike information criterion (AIC) of a given model and the best model

Model	lnL	δAIC	Parameters
All sequences			
M0 — one ω	–1546.1	111.6	$\omega = 0.361$
M7 — nearly neutral with beta	–1492.0	5.4	$p_{beta} = 0.44$, $q_{beta} = 0.85$
M8 — positive selection with beta ($\omega_0 \leq 1$, $\omega_1 > 1$)	–1487.3	best	$p_{beta} = 0.58$, $q_{beta} = 1.33$, $p_0 = 0.929$, $p_1 = 0.071$, $\omega_1 = 1.959$
Group A (putative locus *Trcr-DAB*)			
M0 — one ω	–934.4	182.2	$\omega = 1.256$
M7 — nearly neutral with beta	–853.0	21.4	$p_{beta} = 0.02$, $q_{beta} = 0.03$
M8 — positive selection with beta ($\omega_0 \leq 1$, $\omega_1 > 1$)	–840.3	best	$p_{beta} = 0.02$, $q_{beta} = 0.06$, $p_0 = 0.874$, $p_1 = 0.126$, $\omega_1 = 4.316$
Group B (putative pseudogenes)			
M0 — one ω	–662.8	best	$\omega = 0.190$
M7 — nearly neutral with beta	–662.5	1.4	$p_{beta} = 2.83$, $q_{beta} = 11.45$
M8 — positive selection with beta ($\omega_0 \leq 1$, $\omega_1 > 1$)	–662.5	5.4	$p_{beta} = 2.83$, $q_{beta} = 11.45$, $p_0 = 1.000$, $p_1 = 0.000$

Table 5 Comparison of F_{ST} indices across all loci and three different population groupings (all, PGE and REF populations) based on five microsatellite loci and 40 allozyme loci in the great crested newt. Mean values with ± 1 SE and 99% CI in parentheses

	F_{ST} all	F_{ST} PGE	F_{ST} REF	F_{IS} all	F_{IS} PGE	F_{IS} REF
Microsatellites	0.105 ± 0.007	0.086 ± 0.041	0.066 ± 0.013	0.026 ± 0.041	0.011 ± 0.028	0.088 ± 0.082
	(0.089; 0.118)	(0.002; 0.181)	(0.039; 0.094)	(–0.058; 0.124)	(–0.060; 0.069)	(–0.039; 0.315)
Allozymes	0.199 ± 0.072	0.049 ± 0.011	0.092 ± 0.114	0.094 ± 0.114;	0.298 ± 0.310	0.067 ± 0.125
	(0.039; 0.375)	(0.00; 0.057)	(–0.033; 0.351)	(–0.161; 0.355)	(–0.133; 0.55)	(–0.199; 0.329)

Microsatellite and allozyme variation

The number of alleles in the five microsatellite loci ranged from 19 (*Tcri*43), through 20 (*Tcri*35, *Tcri*13), 21 (*Tcri*29) to 36 (*Tcri*36). Basic diversity measures are given in Table 1 and allele frequencies in Table S3, Supporting Information. After the sequential Bonferroni correction, only one population, TIR, showed a significant excess of homozygotes at one locus, *Tcri43*. The Micro-Checker software detected a null allele with an estimated frequency of 0.22 (Brookfield 1996, I estimator) in this population. Significant among-population differentiation was detected if all populations were included in the analysis (Table 5). Significant population structure was also evident among PGE and REF population groups (Table 5). All pairwise F_{ST} comparisons among populations were significant ($P < 0.05$), with the exception of MER and PAC after Bonferroni correction (Table S4, Supporting Information). However, these two populations had small sample sizes (five individuals each) which may have decreased the power of the pairwise randomization tests for detecting population differentiation. Linkage disequilibrium was not detected between any of the loci.

A total of 27 out of 40 allozyme loci were monomorphic (allele frequencies are given in Table S5, Supporting Information). This was reflected by the low values for mean numbers of alleles across loci per population, mean expected heterozygosity and mean allelic richness (Table 1). Departures from Hardy–Weinberg equilibrium or linkage equilibrium were not detected in any of the loci. All tests for pairwise population differentiation were significant with the exception of UK and POL (Table S4). Population structure based on allozyme variation was detected only across all populations and not within colonized or refugial population groups (Table 5).

Comparison of MHC and neutral variation between populations from REF and PGE areas

A substantial part of the variation in *Trcr-DAB* was partitioned among the PGE and REF groups (40.9%), in sharp contrast to the microsatellite and allozyme variation, the majority of which was accounted for at the within-population level (Table S6, Supporting Information). We found remarkable differences in the levels of MHC variation between the REF and PGE areas (Table 1). In both A and B groups, the alleles found in the PGE populations represent subsets of REF alleles. Only two alleles from the A group, *Trcr-DAB*01* and *Trcr-DAB*02*, were found in the northern PGE populations. This second allele was only identifiable after amplification with primer TrMHCII5F located slightly upstream of the assayed exon fragment: its sequence was very similar to *Trcr-DAB*01*, differing by three nucleotide and two amino acid substitutions, the differences occurring in the fragment not covered by the TrMHC10F-TrMHC5R amplicons. In contrast, all 24 *DAB* alleles were found in Romanian REF populations, despite the smaller sample size in REF.

The number of alleles in group B also differed strikingly between REF and PGE. Four alleles representing presumably three mostly monomorphic loci were detected in PGE and all 18 alleles were found in REF. Allelic richness adjusted to the smallest sample size of five individuals per population was significantly higher in REF populations for both the A and B groups (in both cases Mann–Whitney $Z = –2.12$, $P = 0.03$). We observed strong correlations between population-level allelic richnesses in *DAB* and putative pseudogenes (Spearman $r_S = 0.88$, $P = 0.008$) *DAB* and microsatellites ($r_S = 0.96$, $P = 0.0005$) as well as between pseudogenes and microsatellites ($r_S = 0.94$, $P = 0.002$).

Clearly, the REF populations contained most of the microsatellite diversity, as shown by measures of allelic richness and the number of private alleles (Table 1). Allelic richness was higher in REF than in PGE populations ($R = 6.2$ vs. 4.8, Mann–Whitney $Z = –2.12$, $P = 0.03$). Private alleles, although found in all samples, were far more numerous in the REF populations (30) than in the PGE populations (7). Allozyme allelic richness was significantly higher in REF ($R_S = 1.31$ vs. 1.11 $P = 0.002$, randomization test permuting individuals between regions).

Discussion

We found evidence for a single expressed MHC class II locus in the great crested newt. Moreover, multiple loci that do not undergo expression and exhibit internal stop codons, and thus probably constitute nonfunctional pseudogenes, are present in this species as well. Only the A group (expressed) sequences bear the unambiguous signal of positive selection at the ABS sites, the pattern commonly reported

for MHC class II. Signatures of positive selection were not detected in the putative pseudogene sequences of the B group. 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 should persist in nucleotide sequences for long periods of evolutionary time (Garrigan & Hedrick 2003), even after the locus becomes non-functional. Thus, the MHC class II pseudogenes present in the great crested newt have either evolved neutrally or are under purifying selection, perhaps because they were never expressed or became pseudogenes very long ago. The excess of synonymous mutations in pseudogene sequences remains unclear. This may be a historical legacy, i.e. the loci could have played some functional role long in the past, before stop codons rendered them non-functional. Potentially, the recombination we have detected may have affected the tests of positive selection (Anisimova *et al*. 2003; Shriner *et al*. 2003). However, evidence for recombination was not compelling in our data, detected by only one of three applied methods; therefore, we regard our inferences on patterns of positive selection as robust.

We detected a clear-cut difference in the level of MHC sequence variation between the PGE and REF populations of *Triturus cristatus*. Extensive variation in both A and B allele groups in REF contrasts with very low variation in the A group and almost no variation in the B group in the PGE. This might be explained by directional selection fixing A and/or B alleles at respective loci, on account of local adaptation. However, the data from putatively neutral allozyme and microsatellite loci suggest a nonadaptive explanation for the observed pattern. The reduced variation observed in the PGE for all classes of genetic markers can be attributed to bottlenecks during the postglacial expansion. Some microsatellite variation may have been restored postglacially in the PGE due to high mutation rates as has been suggested for *Mesotriton alpestris* (Pabijan & Babik 2006). This should not be the case for neutrally evolving sequence markers which generally mutate at a slower pace. Low pseudogene sequence variation in the PGE area may therefore provide additional support for the bottleneck hypothesis. Allozyme and mtDNA data from populations from various parts of the range of *T. cristatus* (Wallis & Arntzen 1989) are also congruent with the bottleneck explanation. For example, restriction fragment length polymorphism analysis shows homogeneity of mtDNA across the extensive range of *T. cristatus* with more variation found only in southern Romania and the adjacent part of Serbia, at the southern margin of the species' range, in the hypothesized refugial area (Wallis & Arntzen 1989). Interestingly, the same two *DAB* alleles (**01* and **02*) are present in all three widely spaced PGE populations. While ongoing gene flow is conceivable between Poland and Germany, the UK population has probably been isolated from the European mainland since the opening of the English Channel *c*. 8500 years ago. The relatively long persistence of both alleles may reflect a large population size since the expansion, and/or balancing selection maintaining both alleles.

Southern Romania and adjacent Serbia host most of the genetic variation in the great crested newt, both in parts of the genome believed to evolve neutrally and in those of likely adaptive significance, making this area a priority for the conservation of this species. The pattern of 'southern richness–northern purity' has been observed in neutrally evolving loci of many species inhabiting regions affected by the Pleistocene climatic oscillations (Hewitt 2004). In the great crested newt, the loss of genetic variation holds also for MHC, implying that demographic events were decisive in shaping the MHC diversity of northern populations. An increasing number of studies suggest that despite clear signs of historical selection, as demonstrated by the excess of nonsynonymous substitutions at the ABS and long-term retention of allelic lineages (Garrigan & Hedrick 2003), on a shorter timescale or in some parts of species' ranges, demographic events contribute extensively to patterns of MHC variation. Manifested as a concordance between MHC and presumably neutral markers (mtDNA, microsatellites), and a lack of deviation from neutral expectations in terms of genotype and allele frequencies, this phenomenon is widespread having been detected, e.g. in the Australian bush rat (Seddon & Baverstock 1999), Sonoran topminnow (Hedrick *et al*. 2001), Scandinavian wolf (Seddon & Ellegren 2004), Eurasian beaver (Babik *et al*. 2005) and coastal steelhead (Aguilar & Garza 2006). These results imply that selection, in spite of leaving a long-lasting signal of positive selection, may not be strong enough to maintain MHC polymorphism through severe population bottlenecks. Studies on other urodele amphibians indicate that selection plays an important role in maintaining MHC diversity, but evidence for the action of drift is also present. In populations of *M. alpestris* inhabiting areas of postglacial expansion, extensive variation was found in one of the expressed MHC class II loci (*DAB*), but the geographic patterns of variation indicate the combined action of selection and drift in this species (Babik *et al*. 2008). In the axolotl, a species with an extremely restricted geographic range and apparently low long-term population size, previous reports of exceedingly reduced MHC class II variation (Laurens *et al*. 2001) have not been confirmed by new data, indicating high allelic variation and evidence of historical positive selection (Richman *et al*. 2007).

It has been hypothesized that limited MHC variation may compromise the ability of populations to respond to pathogen assault, increasing the probability of extinction (Hughes 1991; Hedrick 2001). However, data collected so far are ambiguous in this respect. Indeed, there are examples of species with low MHC variation that are particularly susceptible to diseases (Hedrick 2001), which carry the

potential to extinguish entire species as exemplified by a transmissible clonal tumor in the Tasmanian devil (Siddle *et al*. 2007). There are also demonstrations of selection overcoming drift in retaining MHC variation (Aguilar *et al*. 2004; but see Hedrick 2004). On the other hand, some populations and species seem immune to the presumed adverse affects of limited MHC variation, as evidenced by their large population sizes and recent expansions (Ellegren *et al*. 1993; Mikko *et al*. 1999; Babik *et al*. 2005). However, these examples represent relatively recent bottlenecks, often attributed to human activity, or species whose specific biology has probably not selected for MHC diversity (Trowsdale *et al*. 1989; Sommer *et al*. 2002). The situation is different for the great crested newt in which low MHC variation has seemingly not hampered species survival in the expansion areas (PGE) for thousands of generations. Extremely low MHC polymorphism and geographic uniformity across the extensive postglacial part of the species' range is apparently the direct legacy of its postglacial colonization. Recent population declines observed in some parts of the species' range are rather local events involving habitat degradation, fragmentation and the introduction of predatory fishes (Arntzen 2003), unlikely to affect large-scale patterns of genetic diversity of this still relatively abundant and widespread species. It thus appears that great crested newts have survived in the expansion area for at least several thousand years despite very low genetic variation at the MHC locus. A somewhat similar scenario seem to have occurred in the moose in which the presence of closely related alleles differing mainly by nonsynonymous substitutions has been interpreted as a result of positive Darwinian selection restoring variation after a severe bottleneck dated to *c*. 100 000 years ago (Mikko & Andersson 1995; Mikko *et al*. 1999). Unlike the crested newts, the amount of variation present in moose populations prior to the postulated bottleneck is unknown, which is also the case in the musk ox and the mountain goat for which similar suggestions have been put forward (Mikko *et al*. 1999; Mainguy *et al*. 2007). Thus, our data provide the most convincing evidence for the long-term survival of populations subjected to positive selection on MHC in the past, but whose ample MHC diversity, evident in refugial populations, was extremely reduced during a past bottleneck.

Hedrick (2001) recommended that inbreeding avoidance should have higher priority in conservation programmes than retention of MHC diversity, especially when these two aims are in conflict. Our data support this recommendation by showing that the lack of MHC variation may not constitute a threat to population survival even in a relatively long time perspective. Nevertheless, it cannot be ruled out that populations devoid of MHC variation may be disadvantaged in the future when confronted with newly emerging diseases (Radwan *et al*. 2007; Siddle *et al*. 2007). This may be of special importance in amphibians, given the major role that is attributed to emerging diseases in their worldwide decline (Pounds *et al*. 2006; Lips *et al*. 2008). As shown by the case of the great crested newt, the retention of MHC diversity and avoidance of inbreeding can probably be achieved simultaneously by protecting populations in refugial areas.

Acknowledgements

Sebastian Steinfartz and Meike Thomas kindly provided samples from Germany. 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.

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.

Arntzen JW (2001) Genetic variation in the Italian crested newt, *Triturus carnifex*, and the origin of a non-native population north of the Alps. *Biodiversity and Conservation*, **10**, 971–987.

Arntzen JW (2003) *Triturus cristatus* superspecies. In: *Handbuch der Reptilien und Amphibien Europas* (ed. Böhme W), pp. 421–514. Aula-Verlag, Wiebelsheim. Germany.

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, Pabijan M, Radwan J (2008) Contrasting patterns of variation in MHC loci in the Alpine newt. *Molecular Ecology*, **17**, 2339–2355.

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

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, Gopurenko D, Williams RN, Dewoody JA (2008) Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. *Evolution*, **62**, 1458–1468.

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

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453–455.

Brown JH, Jardetzky TS, Gorga JC *et al*. (1993) 3-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.

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

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.

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

Goudet J (2001) FSTAT, version 2.9.3, a program to estimate and test gene diversities and fixation indices. Available from http://www.unil.ch/izea/softwares/fstat.html.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

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

Hedrick P (2004) Foxy MHC selection story. *Heredity*, **93**, 237–238.

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

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.

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.

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

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

Krupa AP, Jehle R, Dawson DA *et al*. (2002) Microsatellite loci in the crested newt (*Triturus cristatus*) and their utility in other newt taxa. *Conservation Genetics*, **3**, 87–89.

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

Lande R (1988) Genetics and demography in biological conservation. *Science*, **241**, 1455–1460.

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.

Lips KR, Diffendorfer J, Mendelson JR, Sears MW (2008) Riding the wave: reconciling the roles of disease and climate change in amphibian declines. *Public Library of Science, Biology*, **6**, 441–454.

Mainguy J, Worley K, Cote SD, Coltman DW (2007) Low MHC DRB class II diversity in the mountain goat: past bottlenecks and possible role of pathogens and parasites. *Conservation Genetics*, **8**, 885–891.

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, Andersson L (1995) Low major histocompatibility complex class-II diversity in European and North-American moose. *Proceedings of the National Academy of Sciences, USA*, **92**, 4259–4263.

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.

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

O'Grady JJ, Brook BW, Reed DH (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, **133**, 42–51.

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.

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

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.

Radwan J, Kawalko A, Wojcik JM, Babik W (2007) MHC-DRB3 variation in a free-living population of the European bison, *Bison bonasus*. *Molecular Ecology*, **16**, 531–540.

Raymond M, Rousset F (1995) GenePop (version 1.2)—population-genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

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.

Seddon JM, Baverstock PR (1999) Variation on islands: major histocompatibility complex (MHC) polymorphism in populations of the Australian bush rat. *Molecular Ecology*, **8**, 2071–2079.

Seddon JM, Ellegren H (2004) A temporal analysis shows major histocompatibility complex loci in the Scandinavian wolf population are consistent with neutral evolution. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2283–2291.

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, Schwab D, Ganzhorn JU (2002) MHC diversity of endemic Malagasy rodents in relation to geographic range and social system. *Behavioral Ecology and Sociobiology*, **51**, 214–221.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences, USA*, **101**, 15261–15264.

Stern LJ, Brown JH, Jardetzky TS *et al*. (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature*, **368**, 215–221.

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.

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.

Trowsdale J, Groves G, Arnason A (1989) Limited MHC polymorphism in whales. *Immunogenetics*, **29**, 19–24.

Wallis GP, Arntzen JW (1989) Mitochondrial DNA variation in the crested newt superspecies — limited cytoplasmic gene flow among species. *Evolution*, **43**, 88–104.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

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.

WB uses molecular tools to study historical processes at and below the species level. MP is interested in the phylogeny, population genetics and conservation of amphibians. JWA is Curator of Vertebrates as the National Museum of Natural History, Leiden, the Netherlands, and has his main interests in the systematics and evolutionary biology of European amphibians. DC is a Professor at the University Ovidius, interested in amphibian life history and evolution. WD leads the population genetics group at the Department of Community Ecology of the Helmholtz Centre for Environmental Research — UFZ. JR 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.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Allele frequencies at the putative expressed locus *Trcr-DAB*. PGE, postglacial expansion area; REF, postulated refugial area. See text for abbreviations of population names. **01–*24*: alleles *Trcr-DAB*01–*24*

Table S2 The average nucleotide and amino acid distances among the great crested newt MHC II alleles. Nucleotide distance is corrected for multiple substitutions using the Kimura 2-parameter model (K2P), amino acid distance is corrected using expectations from Poisson distribution. ABS, putative antigen binding sites as determined by Brown *et al*. (1993). Distances are given as percentages per site

Table S3 Allele frequencies for five microsatellite loci (*Tcri13*, *Tcri29*, *Tcri35*, *Tcri36*, *Tcri43*). PGE, postglacial expansion area; REF, postulated refugial area. See text for abbreviations of population names. Allele sizes are listed in columns to the left of loci names

Table S4 Population differentiation as estimated by F_{ST} values. Significance assessed by 21 000 and 15 000 permutations for microsatellite and allozyme loci, respectively. Above diagonal, pairwise population differentiation based on five microsatellite loci. Below diagonal, pairwise population differentiation based on 40 allozyme loci. Asterisks denote significance level: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. See text for abbreviations of population names

Table S5 Allele frequencies for 40 allozyme loci. Nomenclature as in Arntzen (2001). PGE, postglacial expansion area; REF, postulated refugial area. See text for abbreviations of population names. Allele sizes are listed in columns to the left of loci names

Table S6 Comparison of AMOVA results for *Trcr-DAB*, microsatellites and allozymes. Percentage of variation explained by among group (PGE and REF), among population within groups and within population variance components with associated significance values based on 1000 permutations: NS, nonsignificant; $^{*}P < 0.05$; $^{***}P < 0.001$

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.