

MHC class II genes in the European badger (*Meles meles*): characterization, patterns of variation, and transcription analysis

Yung Wa Sin · Hannah L. Dugdale · Chris Newman ·
David W. Macdonald · Terry Burke

Received: 6 June 2011 / Accepted: 28 September 2011 / Published online: 26 October 2011
© Springer-Verlag 2011

Abstract The major histocompatibility complex (MHC) comprises many genes, some of which are polymorphic with numerous alleles. Sequence variation among alleles is most pronounced in exon 2 of the class II genes, which encodes the $\alpha 1$ and $\beta 1$ domains that form the antigen-binding site (ABS) for the presentation of peptides. The MHC thus plays an important role in pathogen defense. European badgers (*Meles meles*) are a good species in which to study the MHC, as they harbor a variety of pathogens. We present the first characterization of MHC class II genes, isolated from genomic DNA (gDNA) and complementary DNA (cDNA), in the European badger. Examination of seven individuals revealed four *DRB*, two

DQB, two *DQA*, and two *DRA* putatively functional gDNA sequences. All of these sequences, except *DRA*, exhibited high variability in exon 2; *DRB* had the highest variability. The ABS codons demonstrated high variability, due potentially to balancing selection, while non-ABS codons had lower variability. Positively selected sites were detected in *DRB* and *DQA*. Phylogenetic analysis demonstrated *trans*-species polymorphism of class II genes. Comparison with cDNA from whole blood revealed that only *DRB* had a transcription pattern reflecting the alleles that were present in the gDNA, while the other three genes had disparities between gDNA and cDNA. Only one sequence was transcribed, even though two gDNA sequences were present, from each of both *DQB* and *DRA*. Our characterization of badger MHC sequences forms a basis for further studies of MHC variability, mate choice, and pathogen resistance in this, and other, species.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-011-0578-9) contains supplementary material, which is available to authorized users.

Y. W. Sin (✉) · C. Newman · D. W. Macdonald
Wildlife Conservation Research Unit, Department of Zoology,
University of Oxford, Recanati-Kaplan Centre,
Tubney House, Abingdon Road, Tubney,
Abingdon, Oxfordshire OX13 5QL, UK
e-mail: yungwa.sin@zoo.ox.ac.uk

Y. W. Sin · H. L. Dugdale · T. Burke
Natural Environment Research Council (NERC) Biomolecular
Analysis Facility, Department of Animal and Plant Sciences,
University of Sheffield,
Western Bank,
Sheffield, South Yorkshire S10 2TN, UK

H. L. Dugdale
Behavioural Ecology and Self-Organization,
University of Groningen,
PO Box 11103, 9700 CC, Groningen, Netherlands

H. L. Dugdale
Theoretical Biology, University of Groningen,
PO Box 11103, 9700 CC, Groningen, Netherlands

Keywords Balancing selection · *DRB* gene ·
Immunogenetics · Mustelid · Pathogen resistance ·
Trans-species polymorphism

Introduction

Major histocompatibility complex (MHC) genes play an important role in the vertebrate immune system and autoimmunity. Cell-surface glycoproteins encoded by the MHC genes are important in both humoral and cell-mediated immune responses, as they bind and present antigens to T-cells and trigger an immune cascade (Swain 1983). The mammalian MHC is organized into functional regions including class I and class II. The genes in each region are subjected to birth-and-death evolution (Piontkivska and Nei 2003; Takahashi et al. 2000), in which new genes are

created by gene duplication, and some genes become nonfunctional through deleterious mutations. MHC class I molecules are primarily responsible for intracellular antigen binding and are expressed on the surface of all nucleated somatic cells (Hughes and Yeager 1998). MHC class II molecules, on the other hand, principally bind peptides from the extracellular environment and are only expressed on antigen-presenting cells such as B-cells and macrophages (Hughes and Yeager 1998). If a helper T-cell recognizes a presented antigen, the antigen-presenting cell is activated to trigger an immune response.

At the molecular level, MHC class II molecules are heterodimers that consists of two *trans*-membrane proteins: an α - and a β -chain, which are encoded by class II genes *A* and *B*, respectively. The α 1- and β 1-domains are regions in the α - and β -chains that contain specific sites forming the antigen-binding site (ABS; Furlong and Yang 2008). In placental mammals, the class II region is divided into subregions, or gene clusters, known in humans as DR, DQ, DP and DM (Kelley et al. 2005; Takahashi et al. 2000). Each cluster contains a functional α -chain gene and one or more functional β -chain gene(s). A similar organization of class II genes has been found in the canine MHC, known as the dog leukocyte antigen (DLA) region (Wagner 2003). Some gene clusters have apparently been lost in some evolutionary lineages. For example, the feline MHC lacks the DQ subregion and does not transcribe DP molecules; however, the feline MHC has an expanded DR subregion (Yuhki et al. 2003). DR and DQ molecules present different peptides and are expressed differentially, according to cell type; for example, DQ is expressed at a lower frequency than DR on peripheral blood monocytes (Alonso et al. 1985; Edwards et al. 1985; Rees et al. 2003).

The high number of MHC alleles and high degree of sequence diversity among alleles are principally due to the numerous and diverse exons that encode the domains forming the ABS (Hughes and Yeager 1998). The nucleotide diversity within MHC genes is generally attributed to balancing selection (Bergstrom and Gyllensten 1995), which encompasses many specific forms of selection that act to maintain large numbers of alleles in populations. Balancing selection greatly enhances the persistence of allelic diversity over long periods of time relative to neutral genetic variation (Richman 2000). Balancing selection, driven principally by ever-evolving pathogens (Spurgin and Richardson 2010), leads to *trans*-species polymorphism (Klein 1987), resulting in orthologous alleles among related species.

There is currently a paucity of studies on the relationships among MHC genotype, parasite burden, and mate choice in wild mammal populations (Bernatchez and Landry 2003). In this study, we characterize the MHC class II genes of the European badger (*Meles meles*), a nocturnal and fossorial carnivore that rests by day in

underground burrow systems, termed setts, and emerges to forage independently at dusk (Roper 2010). *M. meles* has an extensive geographical range, stretching from Europe to Japan and from the Arctic Circle to the Mediterranean, and exhibits considerable socio-spatial variability (Macdonald et al. 2004; Rogers et al. 1997; Roper 2010; Rosalino et al. 2004). *M. meles* has a long mating season (Buesching et al. 2009), delayed implantation (Thom et al. 2004), putative superfoetation (Yamaguchi et al. 2006), and a sensory predisposition toward olfaction (Buesching et al. 2002). In high-density populations, *M. meles* has a polygynandrous mating system (Dugdale et al. 2007, 2011) with high levels of extra-group paternity (Dugdale et al. 2007) and low fecundity (Macdonald et al. 2009). Additionally, it has been the subject of a diverse range of parasitic disease studies (Anwar et al. 2006; Hancox 1980; Lizundia et al. 2011; Macdonald et al. 1999; Newman et al. 2001; Rosalino et al. 2006; San 2007). *M. meles* is thus potentially well suited for investigating how MHC selection, and conferred immunological advantages (e.g., parasite resistance), is modified and regulated by mate choice.

In this study, we (1) characterized the MHC class II genes of *M. meles* from a high-density population and tested for evidence of selection; (2) identified the transcription pattern of alleles by comparing gDNA and cDNA sequences from whole blood samples, which is important as MHC genes identified using gDNA may be nonfunctional (Knapp 2007); and (3) performed phylogenetic analyses to investigate whether *M. meles* MHC sequences belong in monophyletic groups or whether *trans*-species polymorphism occurs.

Materials and methods

Sample collection and nucleic acid isolation

Blood samples were collected ad hoc from seven badgers associated with five different social groups (Macdonald et al. 2008), as part of an ongoing long-term study (see Macdonald et al. 2009) in Wytham Woods, Oxfordshire, UK (global positioning system reference 51:46:26 N, 1:19:19 W; all trapping and handling protocols were subject to ethical review and performed under government licence — the methods are detailed in Macdonald and Newman 2002; Macdonald et al. 2009). The mean of the pairwise Queller and Goodnight (1989) index of relatedness (*R*) of these seven individuals was 0.18 (95% CI=0.10–0.26), slightly lower than half-siblings [0.31 (95% CI=0.30–0.33); Dugdale et al. 2008], but unfortunately, the wild nature of this population restricted sample availability. Nevertheless, no additional *DRB* sequences have been found from genotyping of the entire population (Y.W. Sin, unpublished data). Approximately 3 ml of blood was taken by jugular

venipuncture and collected in a vacutainer containing EDTA. Samples were stored at -20°C until DNA isolation was performed. A single sample from an American badger *Taxidea taxus jacksoni* (sampling locality: Wisconsin, USA) was also examined and the sequences generated were included in the phylogenetic analysis. gDNA was isolated using the GFX Genomic Blood DNA Purification Kit (Amersham Biosciences, Little Chalfont, UK), following the scalable method in the manufacturer's protocol. In order to validate whether the identified alleles were transcribed, a 500- μl blood sample, from each of the seven individuals, was also transferred into RNAprotect Animal Blood Tubes (Qiagen, Hilden, Germany) and stored immediately at -20°C for less than a month before RNA isolation. Total cellular RNA was isolated from each blood sample using an RNeasy Protect Animal Blood Kit (Qiagen).

cDNA synthesis

A standard cDNA synthesis was performed on 2 μg of RNA, quantified using a NanoDrop 8000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA), from each individual. The reverse-transcription reaction components included 4 units of reverse transcriptase (Omniscript, Qiagen), 0.5 mM of each dNTP and 1 μM oligo-dT primer, in 1 \times RT buffer. The reverse transcription reactions were incubated at 37°C for 60 min and stored at -20°C .

Primer design and PCR amplification

To amplify the near full-length transcribed class II genes from cDNA and class II genes in exon 2 from gDNA, published primers of other canivores — the domestic dog *Canis lupus familiaris* (Kennedy et al. 2006; Wagner et al. 1996), DRB exon 2 of the European mink *Mustela lutreola* (Becker et al. 2009), and transcribed class II genes of the California sea lion *Zalophus californianus* (Bowen et al. 2002, 2004) — were tested. Oligonucleotide primers, recognizing highly conserved regions of each of four MHC class II genes (*DRB*, *DQB*, *DQA*, and *DRA*), were also designed using OligoAnalyzer 3.1 (<http://eu.idtdna.com/analyzer/applications/oligoanalyzer>), based on alignments with GenBank's nucleotide sequences from mustelids (*DRB*: EU121854–EU121855, EU263550–EU263558; *DQB*: EU121856–EU121857), canines (*DRB*: AF016910–AF016912, CFU44777–CFU44778, M29611, NC006594, NM001014768; *DQB*: AF016908–AF016909, M90802, NC006594; *DQA*: NC006594, Y07942–Y07944; *DRA*: L37332, NC006594), pinnipeds (*DRB*: AY491456–AY491467; *DQB*: AF503397–AF503401, AF503406–AF503407; *DQA*: AF093799, AF502560–AF502564; *DRA*: AY491450–AY491455),

porcines (*DQA*: NM001130224), and humans (*DQA*: NM002122; *DRA*: NM019111).

PCR amplification, using these primers on 10–30 ng of cDNA/gDNA, was performed in a 20- μl reaction mix that also contained 0.5 μM of each primer (Table 1), 200 μM of each dNTP, 1 \times PCR buffer (containing MgCl_2 ; Qiagen), and 2 units HotStarTaq (Qiagen). The PCR cycle began with incubation at 94°C for 15 min, followed by 30 incubation cycles at 94°C for 30 s, annealing temperature (Table 1) for 30 s, and 72°C for 1–2 min according to amplicon length, ending with an extension step at 72°C for 10 min. The PCR products were electrophoresed on a 1.5% agarose gel and visualized by UV light and ethidium bromide staining. A 100-bp DNA ladder (New England Biolabs, Herts, UK) was used to size the DNA fragments. Bands of expected size were excised from the gel and purified using QIAquick Gel Extraction Kits (Qiagen). PCR products that gave rise to relatively bright bands of the expected size were cloned and sequenced.

Cloning and DNA sequencing

Purified PCR fragments were ligated into a TA cloning vector (pGEM-T Easy Vector Systems, Promega, Madison, WI, USA). Transformation, growth, and blue-white selection were carried out in *Escherichia coli* competent cells (JM109). PCRs were performed on positive clones using M13 forward and reverse primers (Messing 1983). Consensus sequences were created. All amplified fragments were sequenced in both complementary directions using dideoxynucleotide chain termination and loaded on an automated DNA Sequencer (48-well capillary ABI3730, Applied Biosystem, Foster City, CA, USA). Between 6 and 24 clones were sequenced for each individual. Identical sequences were derived from a minimum of two badgers or from independent PCR reactions from the same individual, in compliance with DLA nomenclature rules (Kennedy et al. 1999). Single unique sequences (possible chimeras) were excluded. Nucleotide sequences were analyzed using CodonCode Aligner 3.7.1 (CodonCode, Dedham, MA, USA) and were compared with known MHC class II sequences using the NCBI BLAST program (Altschul et al. 1990). The DNA sequences from the European and American badger were assigned GenBank accession numbers: HQ908092–HQ908108.

Data analyses

Phylogenetic analyses were performed on the consensus alignments of European and American badger *DRB*, *DQB*, *DQA*, and *DRA* exon 2 sequences, against sequences from other Carnivore species available in GenBank (e.g., sea otter *Enhydra lutris*, European mink *M. lutreola*, *Mustela* sp., California sea lion *Z. californianus*, giant panda

Table 1 MHC class II-specific primers for European badgers

	Primer name	Primer sequence	Product size (bp): region amplified	T_a (°C)	Species specificity	Source reference
Amplification of transcribed genes	F: DRBU116	GAGGCTCCTGGATG ACAGCTYTG	776 bp; exon 1–3' UTR	60	Mammalian <i>DRB</i>	Bowen et al. 2004
	R: DRB3'UTR	AGCTGGGAAGGAAG GTTCTTCCT				Bowen et al. 2004
	F: ZcDQBUEX1D	TGTTCCCAATGATGC CTGGGAAG	654 bp; 5'UTR to 5'exon 4	60	Mammalian <i>DQB</i>	Bowen et al. 2002
	R: DQB-L769	CTCTGGGMAGATTC AGACTG				Bowen et al. 2002
	F: Meme-DQAex1F	ACAGAGTTCTGATTC TGGGG	592 bp; exon 1–exon 4	55	Badger <i>DQA</i>	This study
	R: Meme-DQAex4R	CAGCTCTGACATAGG GGTTGGAA				This study
Amplification of exon 2	F: DRAU102	CATAAGTGGAGTCCC WGTGCTAG	647 bp; exon 1–exon 4	55	Badger <i>DRA</i>	Bowen et al. 2004
	R: Meme-DRAex4R	AATGATGCCAACCAG ACCCAC				This study
	F: Meme-DRBex2F	CGTCCCCACAGGAC ATTTC	248 bp; exon 2	59	Badger <i>DRB</i>	This study
	R: Meme-DRBex2R	CTCGCCGCTGCACC				This study
	F: Meme-DQBex2F	TCCCCGCAGAGGATT TCGTG	248 bp; exon 2	55	Badger <i>DQB</i>	This study
	R: Meme-DQBex2R	GCTCACCTCGCCTCTGC				This study
	F: Meme-DQAex2F	TACCAGTCTTATGGTCCC	176 bp; exon 2	48	Badger <i>DQA</i>	This study
	R: Meme-DQAex2R	CATTGGTAGCAGCGGTA TAGTTG				This study
	F: Meme-DRAex2F	AGGACCATGTGATCAT CCAG	209 bp; exon 2	55	Badger <i>DRA</i>	This study
	R: Meme-DRAex2R	ACGTACCATTGGTGTGTTG GGGTG				This study

F Forward, R reverse, T_a annealing temperature, bp base pair

Ailuropoda melanoleuca, and dog *C. lupus familiaris*; accession numbers are provided in the legends of Figs. 1, 2, 3 and 4). Domain borders were assigned based on sequence homology to different HLA genes, respectively (IMGT/HLA database; Robinson et al. 2001). Bayesian phylogenetic inference was performed using MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Data were divided into three partitions based on gene codon positions. We used the web-based application FindModel (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>) to find the best-fit model of nucleotide substitution, which was the general time-reversible model with gamma-distributed rate variation (GTR+ Γ) for both datasets. A Markov chain Monte Carlo (MCMC) search was initiated with random trees and run for 2,000,000 generations, with a sampling frequency of every 100 generations. The standard deviation of split frequencies converged to a value of 0.005 and 0.007 for the α - and β -genes, respectively. We checked for convergence by plotting the likelihood scores against generation and discarded the first 25% of the generations as “burn-in.” Two

separate analyses and four independent chains were then executed to check for convergence of topology.

Rates of non-synonymous (d_N) and synonymous (d_S) substitutions per site were estimated in DnaSP 4.0 (Rozas et al. 2003) and MEGA 4 (Tamura et al. 2007) according to the method of Nei and Gojobori (1986), with Jukes and Cantor (1969) correction. Standard errors were obtained with 1,000 bootstrap replicates. Synonymous and non-synonymous substitutions were calculated separately for the ABS and non-ABS as determined by Reche and Reinherz (2003). CODEML in PAML 4.4b (Yang 2007) was used to check for positive selection sites (PSS) in the $\alpha 1$ and $\beta 1$ -domain of both MHC class II DR and DQ. PSS are indicated by a ratio ω (non-synonymous over synonymous substitutions, d_N/d_S) larger than 1. Three null models of neutral evolution (M0—one ratio, M1a—nearly neutral, and M7—beta) were applied and compared against their nested models allowing for positive selection (M3—discrete, M2a—positive selection, M8—beta and ω , respectively). Model M0 (one ratio) assumes a constant ω ratio for all codons. M1a (nearly neutral) assumes two site

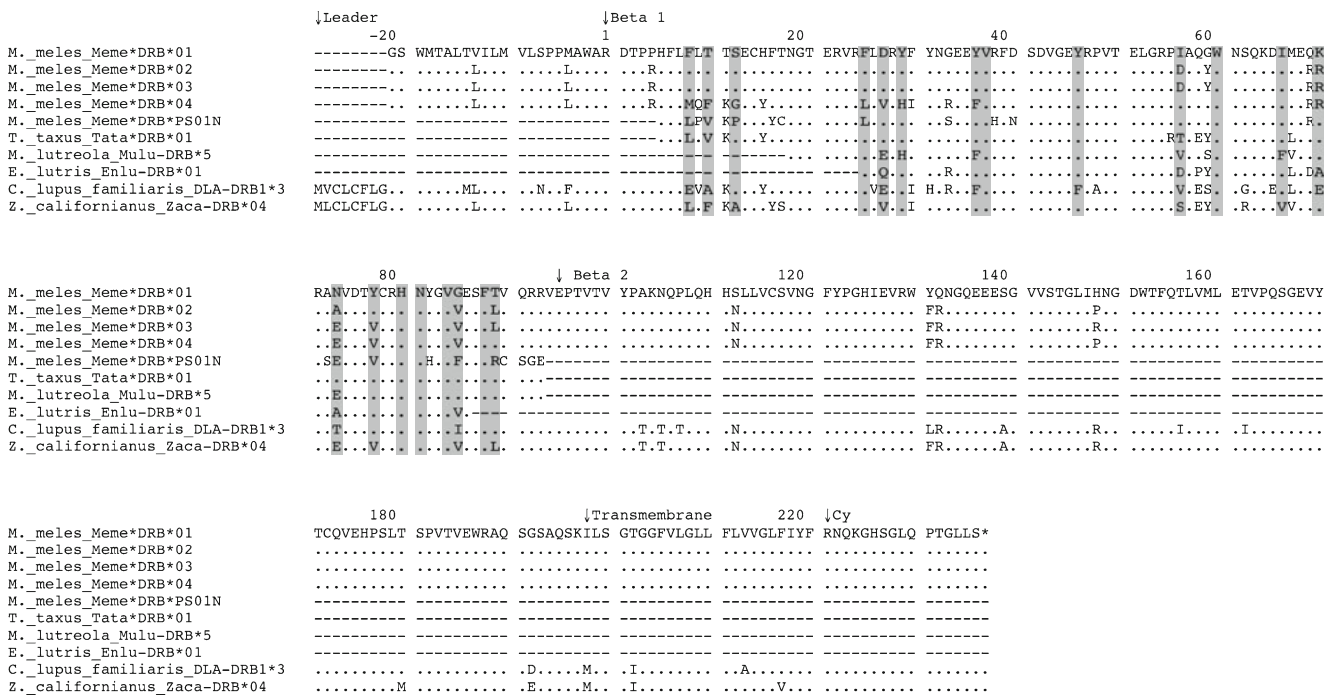


Fig. 1 Amino acid sequence identity for the European badger *Meles meles* class II MHC clones, mustelid, phocine, and canine DRβ sequences. The GenBank accession numbers for DRβ sequences from other mammals are EU263554 (*Mustela lutreola*), EU121855 (*Enhydra lutris*), NM001014768 (*C. lupus familiaris*), and AY491459 (*Zalophus californianus*). The complete amino acid

sequence of *Meme-DRB*01* is shown; numbers above the sequence indicate the codon position in the β domain. Single letters and dots represent amino acids that are distinct from or identical to *Meme-DRB*01*, respectively. Dashes indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with gray boxes

classes [conserved ($0 < \omega_0 < 1$) and neutral ($\omega_1 = 1$)]. M2a (positive selection) adds a third site class ($\omega_2 > 1$) to M1a.

M3 (discrete) estimates the proportion of conserved, neutral, and unrestricted codons from the data. M7 (beta)

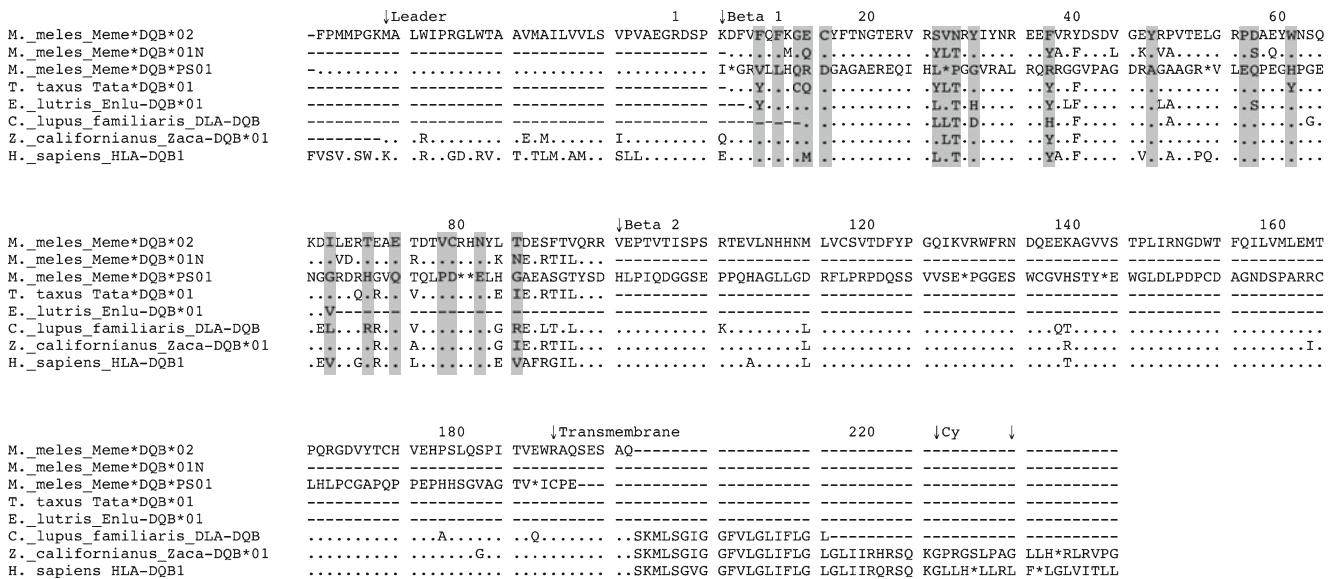


Fig. 2 Amino acid sequence identity for the European badger class II MHC clones, mustelid, phocine, canine, and human DQβ sequences. The GenBank accession numbers for DQβ sequences from other mammals are EU121857 (*Enhydra lutris*), M90802 (*Canis lupus familiaris*), AF503397 (*Zalophus californianus*), and NM002123 (*Homo sapiens*). The complete amino acid sequence of *Meme-*

*DQB*02* is shown; numbers above the sequence indicate the codon position in the β domain. Single letters and dots represent amino acids that are distinct from or identical to *Meme-DQB*02*, respectively. Dashes indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with gray boxes

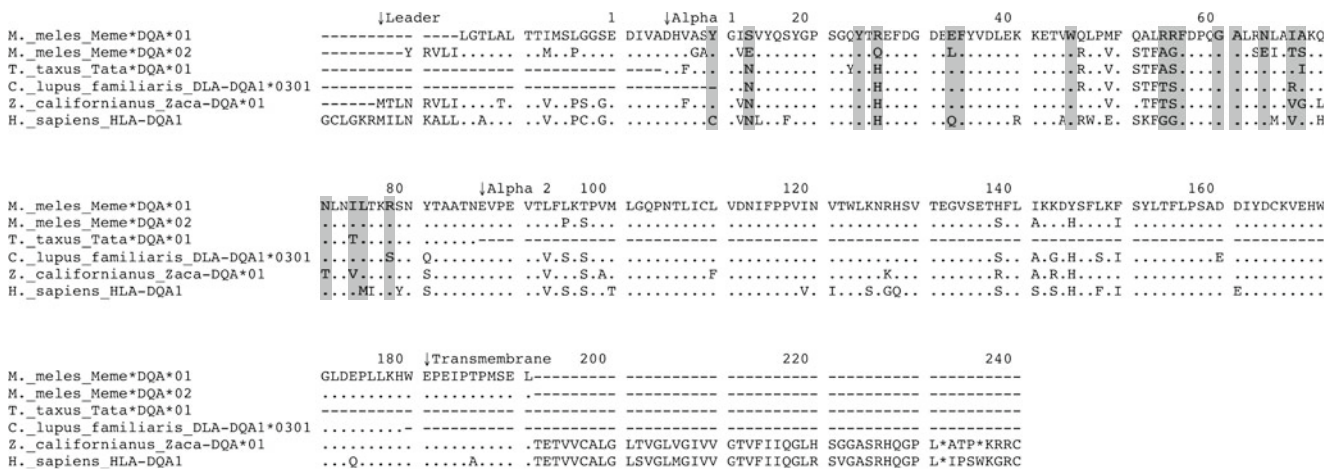


Fig. 3 Amino acid sequence identity for the European badger class II MHC clones, mustelid, phocine, canine, and human DQα sequences. The GenBank accession numbers for *DQA* sequences from other mammals are Y07944 (*Canis lupus familiaris*), AF502560 (*Zalophus californianus*), and NM002122 (*Homo sapiens*). The complete amino acid sequence of *Meme-DQA*01* is shown; numbers above the

sequence indicate the codon position in the α domain. Single letters and dots represent amino acids that are distinct from or identical to *Meme-DQA*01*, respectively. Dashes indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with gray boxes

and M8 (beta and ω) assume an ω ratio with a beta distribution among sites, where M8 adds an extra site class (ω>1). We determined whether the alternative model provided a significantly improved fit, versus the null model, using a likelihood ratio test (LRT), by comparing twice the difference of the log likelihood ratios (2ΔlnL) to a χ² distribution. At each codon, CODEML was used to calculate Bayes Empirical Bayes (BEB) posterior probabilities (Yang et al. 2005), for comparisons of M1a versus M2a and M7 versus M8, to identify codons under positive selection (posterior probabilities>0.95).

Results

MHC class II sequence variability and selection patterns

Five different sequences were isolated for *DRB* genes, three for *DQB*, two for *DQA*, and two for *DRA* from gDNA and cDNA using the primers detailed in Table 1. Two to three putatively functional *DRB* sequences were isolated from each individual, indicating the presence of at least two *DRB* loci. No more than two putatively functional sequences were isolated from each individual for *DQB*, *DQA*, and

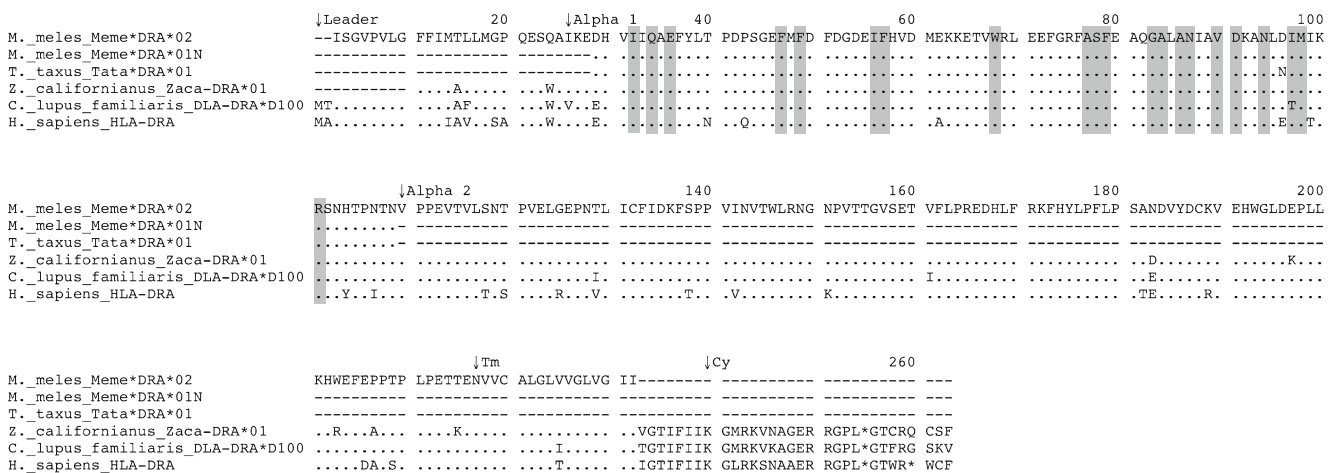


Fig. 4 Amino acid sequence identity for the European badger class II MHC clones, mustelid, phocine, canine, and human DRα sequences. The GenBank accession numbers for *DRA* sequences from other mammals are L37332 (*Canis lupus familiaris*), AY491450 (*Zalophus californianus*), and NM019111 (*Homo sapiens*). The complete amino acid sequence of *Meme-DRα*02* is shown; numbers above the

sequence indicate the codon position in the α domain. Single letters and dots represent amino acid that are distinct from or identical to *Meme-DRα*02*, respectively. Dashes indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with gray boxes

DRA genes, signifying the presence of a single locus. In addition, we inferred the presence of one *DRB* and one *DQB* pseudogene loci (next paragraph). Among the four genes, variation between *DRB* sequences was highest, while that between *DRA* sequences was the lowest (Table 2). The majority of variable sites were in exon 2 (Table 2), where most mutations represented a non-synonymous nucleotide substitution. Hence, DRβ, DQβ, and DQα exhibit more polymorphic amino acid residues among the β1 or α1 domain (Figs. 1, 2 and 3, Table 3); in contrast, the DRα had the least polymorphic amino acid residues (Fig. 4, Table 3).

Pseudogene features (Figs. 1 and 2) were found in two sequences: nucleotide deletions caused a frameshift that occurred at position 355 for *Meme-DRB*PS01N*, and a deletion of 17 nucleotides occurred at the beginning of exon 2 for *Meme-DQB*PS01*, which caused premature stop codons. The deletion in *Meme-DQB*PS01* occurred at the 5' primer annealing site for exon 2 amplification from

Table 2 Sequence polymorphism of MHC class II genes delineated by exon. The number of nucleotide and derived amino acid sequences isolated from seven European badgers *Meles meles* were compared and the numbers of synonymous and non-synonymous nucleotide substitutions are shown

	Exons					
	1	2	3	4	5	6
<i>DRB</i>						
Sequences for comparison	4	4	4	4	4	4
Variable sites	4	43	7	1	0	0
Mutations	4	43	8	1	0	0
Synonymous	2	13	3	1	0	0
Non-synonymous	2	26	5	0	0	0
<i>DQB</i>						
Sequences for comparison	1	2	1	1	–	–
Variable sites	–	39	–	–	–	–
Mutations	–	39	–	–	–	–
Synonymous	–	11	–	–	–	–
Non-synonymous	–	27	–	–	–	–
<i>DQA</i>						
Sequences for comparison	2	2	2	2	–	–
Variable sites	2	38	11	0	–	–
Mutations	2	38	11	0	–	–
Synonymous	0	15	3	0	–	–
Non-synonymous	2	23	8	0	–	–
<i>DRA</i>						
Sequences for comparison	1	2	1	1	–	–
Variable sites	–	3	–	–	–	–
Mutations	–	3	–	–	–	–
Synonymous	–	3	–	–	–	–
Non-synonymous	–	0	–	–	–	–

gDNA, so no sequence was detected from gDNA. Accordingly, these were regarded as pseudogenes and were not included in the nucleotide substitution calculations.

The ratios of non-synonymous to synonymous substitutions for *DRB* and *DQA* were higher in the ABS (Table 4), whereas synonymous substitutions were higher than non-synonymous substitutions in the non-ABS. We inferred PSS for the α1- and β1-domain of MHC DR and DQ using PAML. Models M3 (discrete), M2a (positive selection), and M8 (beta and ω), which allow for positive selection at a subset of sites, gave a significantly better fit than the models without positive selection ($p < 0.01$ in LRTs) in *DRB* and *DQA* but not *DQB* and *DRA* (Table S1). That is, positive selection signals were detected in *DRB* and *DQA*. Parameter estimates (Table S1) suggest that 11.9% (under M2a) and 11.7% (under M8) of *DRB* codon sites are under positive selection with $\omega = 9.2$ (M2a) and $\omega = 9.4$ (M8). Parameter estimates from both M2a and M8 suggest that 27.9% of *DQA* codon sites are under positive selection, with $\omega = 5$ (both M2a and M8). Six *DRB* codon sites (11th, 51st, 60th, 74th, 78th, and 90th) and three *DQA* codon sites (9th, 29th, and 47th) were identified as being under positive selection, with high values under the BEB analysis (Table 5). Four out of six positively selected sites (11th, 74th, 78th, and 90th) in DRβ1 and two out of three positively selected sites (9th and 29th) in DQα1 were within the ABS (gray boxes in Fig. 1 and 3).

Transcription analyses

Of 11 sequences detected in the gDNA, eight were also amplified from the cDNA isolated from whole blood; two were detected inconsistently (Table 6). Pseudogene *Meme-DRB*PS01N* was detected only in gDNA, but *Meme-DQB*PS01* was detected in the cDNA, indicating the presence of a transcribed nonfunctional pseudogene (Fernandez-Soria et al. 1998; Mayer et al. 1993). The transcription of *Meme-DRB*01* and *Meme-DQA*01* was only detected in one out of seven individuals for the *DRB* and *DQA* genes (Table 6), although these alleles were present in their gDNA. In both cases, only one in 20 sequenced clones from cDNA was identified as *Meme-DRB*01* and *Meme-DQA*01*, with the remaining 19 clones determined to be *Meme-DRB*02* and *Meme-DQA*02*, respectively (badger 1, Table 6, Table S2). This implies that some alleles have low levels of transcription in whole blood. *Meme-DRB*04* was transcribed and identified in only a single individual in this study. More than 1,100 individuals from the same population have since been genotyped, and this sequence has been identified in more individuals (Y.W. Sin, unpublished data).

For both *DQB* and *DRA*, only a single sequence was transcribed in all individuals examined, even though two

Table 3 Polymorphisms of four MHC class II molecules according to their domains. The derived amino acid sequences (isolated from seven European badgers *Meles meles*) were compared and the numbers of polymorphic amino acid residues are shown

	Domains					Total
	Leader	$\beta 1/\alpha 1$ domain	$\beta 2/\alpha 2$ domain	<i>Trans</i> -membrane domain	Cytoplasmic domain	
DRβ						
No. of amino acids	21 ^a	95	103	23	16	258
Polymorphic amino acid residues ^b	2	21	4	0	0	27
DQβ						
No. of amino acids	32	90	94	8 ^a	–	224
Polymorphic amino acid residues ^b	–	22	–	–	–	22
DQα						
No. of amino acids	20 ^a	82	94	11 ^a	–	207
Polymorphic amino acid residues ^b	2	18	6	0	–	26
DRα						
No. of amino acids	23 ^a	84	107	16 ^a	–	230
Polymorphic amino acid residues ^b	–	0	–	–	–	0

^a Only part of the domain is included in this study

^b Refer to Table 2 for the number of sequences that were compared

sequences were present in the gDNA of each gene (Table 6). This consistent transcription pattern of the same sequences among all individuals examined was likely also the case for *DQA*; however, *Meme-DQA*01* could be transcribed at such a low level in whole blood that it was not easy to detect by cloning and sequencing. Only *DRB* showed a high degree of variation in transcribed sequences, reflecting the sequences found in the gDNA. From the transcription analysis, it is thus probable that not all MHC class II sequences are transcribed in the blood of *M. meles*.

Phylogenetic analyses

The phylogenetic trees of the class II MHC exon 2 highlight that each gene coding for the β -chain (*DRB* and *DQB*; Fig. 5) and the α -chain (*DQA* and *DRA*; Fig. 6) formed a distinct and highly supported clade. For the *DRB* gene (Fig. 5), mustelid sequences, except for *Meme-DRB*04*, formed a distinct clade [posterior probability (PP) support=95%]. The *M. meles DRB* sequences within this clade did not form a monophyletic group, rather they were interspersed with sequences from *T. taxus* and *E.*

Table 4 Rates (\pm standard error) of non-synonymous (d_N) and synonymous (d_S) substitutions for antigen-binding site (ABS), non-ABS, and combined (ABS +non-ABS) at the European badger *Meles meles* MHC loci

	Position	Number of codons	d_N	d_S	ω
<i>DRB</i>					
	ABS	21	0.315 \pm 0.096	0.181 \pm 0.093	1.74
	Non-ABS	68	0.029 \pm 0.012	0.092 \pm 0.035	0.32
	Combined	90	0.087 \pm 0.021	0.109 \pm 0.031	0.80
<i>DQB</i>					
	ABS	21	0.208 \pm 0.080	0.264 \pm 0.182	0.79
	Non-ABS	68	0.147 \pm 0.042	0.192 \pm 0.083	0.77
	Combined	89	0.162 \pm 0.037	0.208 \pm 0.074	0.78
<i>DQA</i>					
	ABS	19	0.395 \pm 0.171	0.134 \pm 0.108	2.95
	Non-ABS	63	0.090 \pm 0.030	0.287 \pm 0.092	0.31
	Combined	82	0.151 \pm 0.037	0.249 \pm 0.075	0.61
<i>DRA</i>					
	ABS	21	0.000 \pm 0.000	0.104 \pm 0.132	0.00
	Non-ABS	60	0.000 \pm 0.000	0.065 \pm 0.049	0.00
	Combined	81	0.000 \pm 0.000	0.074 \pm 0.046	0.00

ω Ratio of non-synonymous to synonymous nucleotide substitution

Table 5 CODEML detection of amino acid sites under positive selection

	Codon number	Amino acid	Probability $\omega > 1$ (M1a vs. M2a) ^a	Mean $\omega \pm SE$ (M1a vs. M2a)	Probability $\omega > 1$ (M7 vs. M8) ^a	Mean $\omega \pm SE$ (M7 vs. M8)
<i>DRB</i>	11	T	0.904	7.75±2.70	0.955*	8.02±2.26
	51	I	0.937	8.03±2.41	0.968*	8.13±2.11
	60	G	0.961*	8.22±2.15	0.982*	8.24±1.93
	74	N	0.997**	8.49±1.67	0.999**	8.37±1.67
	78	Y	0.999**	8.51±1.64	1.000**	8.38±1.66
	90	T	0.962*	8.22±2.15	0.980*	8.23±1.95
<i>DQA</i>	9	S	0.938	6.99±2.58	0.969*	6.96±2.39
	29	E	0.947	7.07±2.53	0.969*	6.98±2.40
	47	Q	0.971*	7.21±2.36	0.986*	7.08±2.26

ω Ratio of non-synonymous to synonymous nucleotide substitution

^a Positively selected sites were identified in models M2a and M8 by BEB posterior probabilities (Yang et al. 2005)

* $p > 95\%$, ** $p > 99\%$

lutris. For the *DQB* gene (Fig. 5), a distinct clade was formed by *M. meles* sequences: *Meme-DQB*02* and *Meme-DQB*PS01* (PP value=100%); however, this did not cluster with *Meme-DQB*01N*. For the sequences coding for the α -chain (Fig. 6), two *M. meles* *DQA* sequences, and two *M. meles* *DRA* sequences did not form a monophyletic group and were interspersed with other mammalian *DQA* and *DRA* sequences. Although more

than one *DRB* sequence was found in *T. taxus*, only one of them was identified in more than one PCR. In compliance with DLA nomenclature rules (Kennedy et al. 1999), we therefore acknowledged and included only a single sequence for this gene. Only a single sequence was identified for the other three genes.

Discussion

MHC variation and selection

This is the first study to characterize class II MHC genes in European badgers, and importantly, this was performed using both the genome and transcriptome. The *DRB* genes showed higher diversity compared to the *DQB*, *DQA*, and *DRA* genes, which among individuals from this population exhibited a nearly uniform distribution of alleles. Similarly, in other mammals, the *DRB* gene has the highest reported diversity among all class II MHC genes (Bowen et al. 2006b; Doxiadis et al. 2001; Wagner et al. 1999). The other mammalian species tested also exhibited lower diversities for the *DQB* and *DQA* genes (Bowen et al. 2002, 2006b; Doxiadis et al. 2001; Wagner 2003), with the *DLA-DRA* locus appearing monomorphic (Wagner 2003).

In comparison to other mustelid species (Becker et al. 2009; Bowen et al. 2006b), we observed moderate variability in the number of *DRB* sequences for *M. meles*. When compared to closely related species in the Carnivora, the suggested number of loci for the four class II genes that we describe here is identical to reports from a mustelid, *E. lutris* (Bowen et al. 2006b), and a pinniped, *M. angustirostris* (Weber et al. 2004). In addition, one of the *DRB* loci in *M. angustirostris* is monomorphic, while the second is polymorphic with two alleles, which could be similar in *M.*

Table 6 Presence of MHC class II sequences in gDNA and cDNA from seven European badgers *Meles meles*

Gene Sequence	Individual						
	1	2	3	4	5 ^c	6 ^c	7 ^c
<i>DRB</i> <i>Meme-DRB*01</i>	x						
	x	x	x	x	x	x	x
		x				x	x
				x			
<i>DQB</i> <i>Meme-DQB*01N^b</i>							
	x	x	x	x	x	x	x
	x	x			x		
<i>DQA</i> <i>Meme-DQA*01</i>	x						
	x	x	x	x	x	x	x
<i>DRA</i> <i>Meme-DRA*01N^b</i>							
	x	x	x	x	x	x	x

X represents sequence from cDNA. Blank represents no detected sequence. Gray shading indicates sequence from gDNA

^a PS signifies pseudogenes with a frameshift (in *Meme-DQB*PS01*: a deletion occurred at the 5' primer annealing site for exon 2 amplification from gDNA; as a result, no sequence was detected from gDNA)

^b N signifies the presence of genomic sequences, where no cDNA sequences were found

^c Sequences were not determined from gDNA for *DQB*, *DQA*, and *DRA*

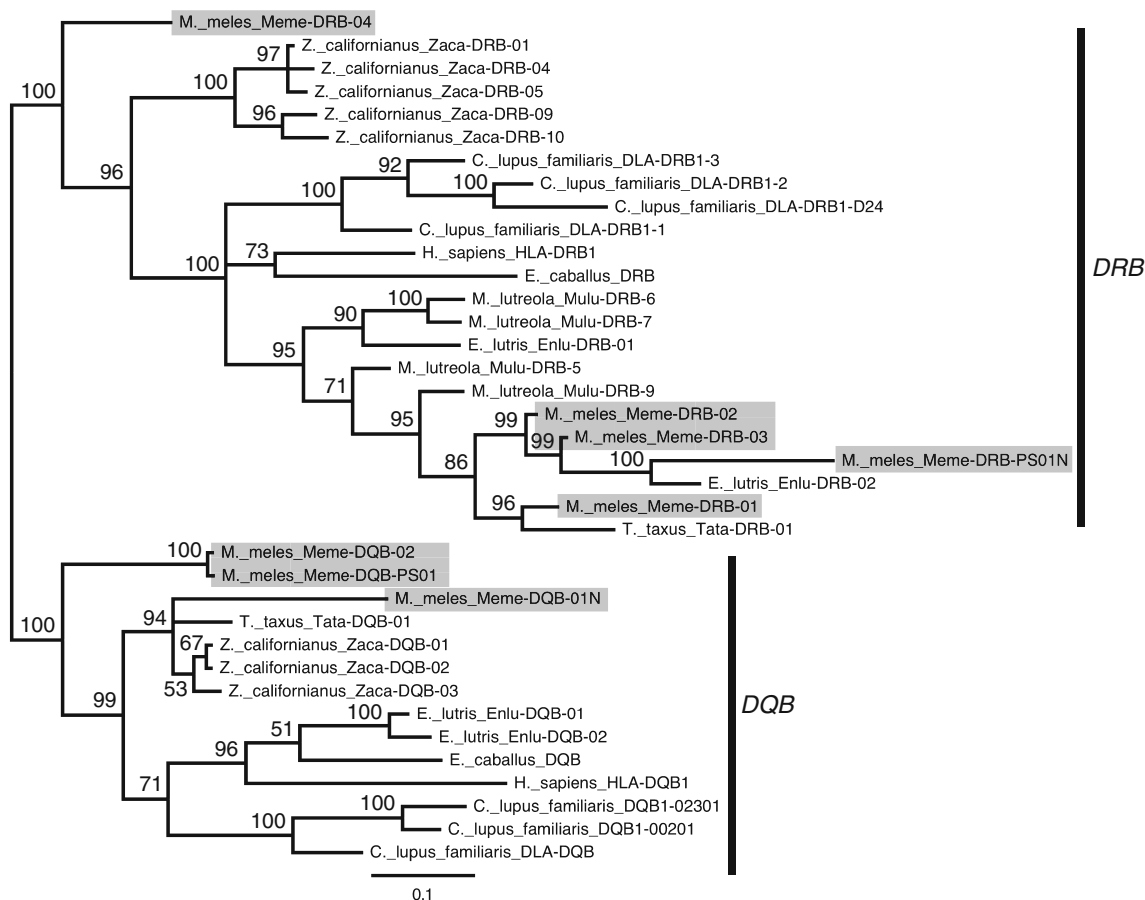


Fig. 5 Phylogenetic tree of MHC class II *DRB* and *DQB* exon 2 sequences from *Meles meles*, *Taxidea taxus*, and other mammal species including mustelid, phocine, canine, equine, and human (GenBank accession numbers are provided in the materials and

methods), based on the 50% majority rule tree from the Bayesian analysis. Bayesian posterior probabilities above 50% are shown above the branches. *M. meles* class II sequences are marked with gray boxes

meles. The number of *DRB* loci in different mammalian species ranges from one to in excess of three (Becker et al. 2009; Bowen et al. 2004; Doxiadis et al. 2001; Takahashi et al. 2000; Yuhki et al. 2003). The MHC diversity in these mammalian species usually arises from a high level of allelic variation at a single locus, and in some species, this locus may be paired with one or two relatively monomorphic loci (Robbins et al. 1997). A few species with multiple *DRB* loci have low polymorphism, but haplotypes are highly variable in their gene content (see Bowen et al. 2004; Doxiadis et al. 2001). In contrast to the *DRB*, the *DQB*, *DQA*, and *DRA* have only one locus in many species (Takahashi et al. 2000). Considering that the four *DRB* alleles that we have characterized are based on seven individuals, which come from a minuscule part of the European badger's range, this suggests that there is a high possibility of detecting more alleles in other populations.

The majority of the polymorphic sites among *M. meles* sequences were found in exon 2, and there were more non-synonymous than synonymous nucleotide substitutions

within this region, contributing to a higher nucleotide and amino acid sequence allelic diversity than within other exons. Exon 2 is responsible for coding the $\alpha 1$ or $\beta 1$ domains, which play a role in presenting antigenic peptides to the immune system (Hughes and Yeager 1998). The high variability in the exon 2 sequences of different alleles allows a population to present a wider repertoire of antigens, increasing their ability to combat parasitic infection (Hughes and Nei 1992).

Within exon 2, some amino acid residues form the ABS (Reche and Reinherz 2003), which is responsible for binding peptides. In the ABS, the rate of non-synonymous substitution (d_N) was higher than the rate of synonymous substitution (d_S) in *DRB* and *DQA*. By contrast, d_S was higher than d_N in the non-ABS. This higher level of d_N in the ABS, which is similar in other species (Cheng et al. 2009; de Bellocq et al. 2009), indicates that this is the target of balancing selection, which would maintain the high levels of genetic diversity that we observed (Richman 2000). Sites outside of the ABS showed a higher d_S than d_N

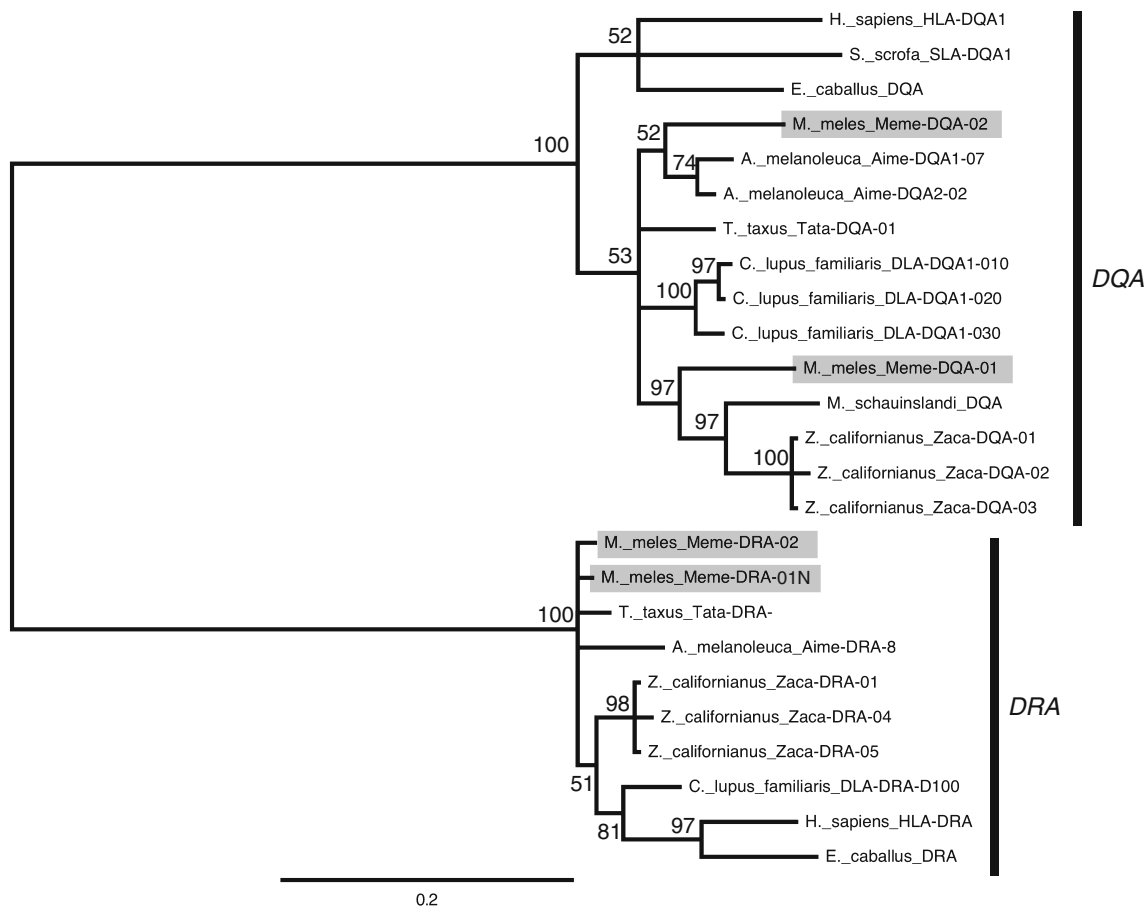


Fig. 6 Phylogenetic tree of MHC class II *DQA* and *DRA* exon 2 sequences from *Meles meles*, *Taxidea taxus*, and other mammal species including mustelid, phocine, canine, equine, and human (GenBank accession numbers are provided in “Materials and

methods”), based on the 50% majority rule tree from the Bayesian analysis. Bayesian posterior probabilities above 50% are shown above the branches. *M. meles* class II sequences are marked with gray boxes

and appeared to have undergone purifying selection, through which non-synonymous substitutions that affect normal gene function will usually be eliminated (Hughes and Yeager 1998).

In our maximum likelihood models, model M0 (one ratio), which assumes a single ω over all codons in the sequence, gave results of $\omega < 1$, revealing that there has been overall purifying selection acting over *M. meles* MHC class II genes. Positive selection most likely affects only a few sites at a few time points (Nielsen 2005; Yang 2002), and so, interpreting the signs of selection from ω averaged over a whole genetic region could be misleading. We detected sites exhibiting positive selection and an excess of non-synonymous substitutions in amino acid residues in *M. meles* MHC II *DRB* and *DQA* genes, which demonstrates that balancing selection plays a role in maintaining these MHC sequence polymorphisms. Most of the sites of positive selection in *DRβ1* and *DQα1* were within the ABS. This is consistent with the theory that balancing selection is mainly act on the ABS and drives the high polymorphism in exon 2 of the MHC class II genes. In five

of the six mammal species studied by Furlong and Yang (2008), *DRβ1* sites 11L, 13F, 57D, 74A, and 86G were under positive selection. We confirm here that the 11th and 74th of these sites are also under positive selection in the European badger, implying that the selective pressures shaping variation at these sites may be driven similarly in different mammals.

Transcription pattern

Many MHC sequences can be detected at the genomic level, but not at the cDNA level (de Groot et al. 2004). This could be because these sequences are pseudogenes or because of fitness advantages conferred by suppressing some inherited genes (DuPasquier et al. 1989): an over-extensive diversity of MHC molecules would minimize the T-cell repertoire of an individual during thymic selection (Nowak et al. 1992; but see Borghans et al. 2003). In terms of how the MHC is related to mate choice and pathogen resistance, nonfunctional MHC sequences (i.e., pseudogenes) should be identified before conducting any analysis;

however, only a few studies (e.g., Babik et al. 2009) have demonstrated the transcription of MHC class II alleles. We found that not all sequences detected in gDNA were detectable in the sequences transcribed from whole blood. Only one *DQB* and *DRA* sequence was found to be transcribed, although two sequences were detected at the genomic level for each of them.

Caution is required when interpreting transcription data, as some of the alleles found in gDNA may be transcribed at very low background levels that are hard to detect (e.g., *Meme-DRB*01* and *Meme-DQA*01* in this study). Alternatively, some alleles may be transcribed at a level high enough to be detected only if the individual is challenged with pathogen invasion, or the transcription of some alleles may vary throughout the life of an individual. In addition, some putative functional sequences may not transcribe at all. Our results suggest that some of the sequences exhibit no transcription, or lower transcription levels than the others, which would be possible if they are situated on different loci, or if there is allele-specific transcription within a locus with transcription controlled by different promoters (Handunnetthi et al. 2010; LeibundGut-Landmann et al. 2004; Villard et al. 1997). Interestingly, the two sequences for which the results suggest a low level of transcription (*Meme-DRB*01* and *Meme-DQA*01*) were both detected in the same individual, potentially because this individual was subjected to a particular immune challenge at the time of sampling. The expression of different MHC genes can vary considerably (Handunnetthi et al. 2010) due to pathogenic challenge (Chen et al. 2006) or differential developmental of the immune system (Handunnetthi et al. 2010). Differential MHC transcription was demonstrated in peripheral blood leukocytes of Steller sea lions (*Eumetopias jubatus*) by using quantitative PCR (Bowen et al. 2006a), which identified extensive variation in the transcription of *DRB* genes even though sequences in the gDNA were identical in all individuals.

Despite the caveats above, we found that *DRB* and *DQA*, where all the detected putative functional sequences were transcribed, showed signals of positive selection. Conversely, models of neutral evolution could not be rejected in *DQB* and *DRA*, where only one of the two putative functional sequences was transcribed. The transcription pattern could be used to explain why no positive selection was detected on *DQB* and *DRA*, as in all probability one of the two sequences was not transcribed and thus not subject to positive selection. The possibility also exists that sequences were missed and it is not possible to prove that certain sequences were not transcribed in other tissues, given that we only tested blood samples. More individuals from different geographical locations would need to be examined to elucidate these transcription patterns, and the diversity of MHC class II genes fully in *M. meles*.

The *DRB* gene exhibited higher variability than the other class II genes among the sampled badgers. Furthermore, all four *DRB* gDNA sequences were transcribed. Although *Meme-DRB*01* was not detected from the cDNA of some individuals, it should still be considered to be a transcribed sequence, as it may potentially have low transcription levels. The *DRB* gene also has the highest variability among all of the class II MHC genes in a variety of other species (Garrigan and Hedrick 2003; Kennedy et al. 2002; Wagner et al. 1999). The *DRB* gene in European badgers is thus a better candidate class II MHC gene, than *DQB*, *DQA*, and *DRA*, for pursuing further studies on MHC and mate choice/pathogen resistance.

Trans-specific evolution

The European badger is classified within the subfamily Melinae, which diverged from the subfamily Taxidiinae somewhere between 16–17 (Marmi et al. 2004) and 20 (Yonezawa et al. 2007) million years ago (Mya). Taxidiinae now only contains the extant American badger, *T. taxus*, which forms the basal group in the family Mustelidae (Koepfli et al. 2008; Yonezawa et al. 2007). The divergence time of Melinae to other Mustelidae subfamilies, i.e., Lutrinae (otters) and Mustelinae, is estimated as 20–12 Mya (Marmi et al. 2004; Yonezawa et al. 2007). The superfamily Musteloidea is a sister taxon to Pinnipedia, and together with Ursidae and Canidae, they form the suborder Caniformia (Flynn et al. 2005).

In contrast to the broader phylogenetic relationship resolved by other studies, the MHC sequences of *M. meles* reported here do not form a monophyletic group, or cluster, with closely related taxa; rather a mixture of MHC alleles among the Pinnipedia, Ursidae, and Canidae was established. This phenomenon has been observed in a number of other MHC studies (Edwards and Hedrick 1998) and is termed *trans-species* polymorphism (Klein 1987). The variability of the MHC genes in the ancestral species may have been retained even when the species diverged, thus an allele may be more closely related to an allele from another species rather than from another allele within the same species. Balancing selection (Penn and Potts 1999), indicated by higher non-synonymous substitution in the ABS, is the proposed mechanism that maintains this ancestral variation, even though *M. meles* diverged from the other species included in this study more than 10 Mya.

Conclusions and implications

This is the first published study to describe MHC sequences in the European badger. The gene-specific primers developed in this study will facilitate future research into the

genetic diversity at the *DRB*, *DQB*, *DQA*, and *DRA* loci in badgers and other closely related mustelid species, and the correlation between these MHC genes and fitness-related traits. Balancing selection acts to maintain genetic diversity, evidenced here by the high diversity of the *DRB* sequences, the detection of positive selection, and *trans*-species polymorphism. Informatively, transcription analysis showed that the gDNA and cDNA sequence patterns were identical in *DRB*, which was moderately variable between individuals. Conversely, there were large disparities between gDNA and cDNA sequences obtained from whole blood in *DQB*, *DQA*, and *DRA*. This highlights the need for future MHC studies to examine transcription patterns before investigating any further questions.

Acknowledgments We thank the Wytham Woods badger team, especially Geetha Annavi, Christina Buesching, Stephen Ellwood, and Pierre Nouvellet, for assistance with badger trapping. We thank Geetha Annavi for providing relatedness data. We also thank Deborah Dawson, Gavin Horsburgh Andy Krupa, and Maria-Elena Mannarelli for assistance with laboratory work. The *T. taxus* samples were generously provided by Alain Frantz and Emily Latch. We thank two anonymous reviewers for their helpful comments on the manuscript. This research was supported by the UK NERC Biomolecular Analysis Facility, the Croucher Foundation (Hong Kong), and the Netherlands Organisation for Scientific Research. All animal handling was carried out under Natural England Licence, currently 20104655, and UK Home Office License PPL 30/2835.

Conflicts of Interest None

References

- Alonso MC, Navarrete C, Solana R, Torres A, Pena J, Festenstein H (1985) Differential expression of HLA-DR and HLA-DQ antigens on normal cells of the myelomonocytic lineage. *Tiss Antig* 26:310–317
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Anwar A, Knaggs J, Service KM, McLaren GW, Riordan P, Newman C, Delahay RJ, Cheesman C, Macdonald DW (2006) Antibodies to *Toxoplasma gondii* in Eurasian badgers. *J Wildl Dis* 42:179–181
- Babik W, Pabijan M, Arntzen JW, Cogalniceanu D, Durka W, Radwan J (2009) Long-term survival of a urodele amphibian despite depleted major histocompatibility complex variation. *Mol Ecol* 18:769–781
- Becker L, Nieberg C, Jahreis K, Peters E (2009) MHC class II variation in the endangered European mink *Mustela lutreola* (L. 1761)—consequences for species conservation. *Immunogenetics* 61:281–288
- Bergstrom T, Gyllensten U (1995) Evolution of Mhc class II polymorphism: the rise and fall of class II gene function in primates. *Immunol Rev* 143:13–31
- Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* 16:363–377
- Borghans JAM, Noest AJ, De Boer RJ (2003) Thymic selection does not limit the individual MHC diversity. *Eur J Immunol* 33:3353–3358
- Bowen L, Aldridge BM, Gulland F, Woo J, VanBonn W, DeLong R, Stott JL, Johnson ML (2002) Molecular characterization of expressed DQA and DQB genes in the California sea lion (*Zalophus californianus*). *Immunogenetics* 54:332–347
- Bowen L, Aldridge BM, Gulland F, VanBonn W, DeLong R, Melin S, Lowenstine LJ, Stott JL, Johnson ML (2004) Class II multiformity generated by variable MHC-DRB region configurations in the California sea lion (*Zalophus californianus*). *Immunogenetics* 56:12–27
- Bowen L, Aldridge B, Beckmen K, Gelatt T, Rea L, Burek K, Pitcher K, Stott JL (2006a) Differential expression of immune response genes in Steller Sea Lions (*Eumetopias jubatus*): an indicator of ecosystem health? *EcoHealth* 3:109–113
- Bowen L, Aldridge BM, Miles AK, Stott JL (2006b) Expressed MHC class II genes in sea otters (*Enhydra lutris*) from geographically disparate populations. *Tiss Antig* 67:402–408
- Buesching CD, Waterhouse JS, Macdonald DW (2002) Gas-chromatographic analyses of the subcaudal gland secretion of the European badger (*Meles meles*). Part I. Chemical differences related to individual parameters. *J Chem Ecol* 28:41–56
- Buesching CD, Heistermann M, Macdonald DW (2009) Seasonal and inter-individual variation in testosterone levels in badgers *Meles meles*: evidence for the existence of two endocrinological phenotypes. *J Comp Physiol A* 195:865–871
- Chen SL, Zhang YX, Xu MY, Ji XS, Yu GC, Dong CF (2006) Molecular polymorphism and expression analysis of MHC class II B gene from red sea bream (*Chrysophrys major*). *Dev Comp Immunol* 30:407–418
- Cheng Y, Siddle HV, Beck S, Eldridge MDB, Belov K (2009) High levels of genetic variation at MHC class II DBB loci in the tamar wallaby (*Macropus eugenii*). *Immunogenetics* 61:111–118
- de Bellocq JG, Suchentrunk F, Baird SJE, Schaschl H (2009) Evolutionary history of an MHC gene in two leporid species: characterisation of Mhc-DQA in the European brown hare and comparison with the European rabbit. *Immunogenetics* 61:131–144
- de Groot N, Doxiadis GG, de Groot NG, Otting N, Heijmans C, Rouweler AJM, Bontrop RE (2004) Genetic makeup of the DR region in rhesus macaques: gene content, transcripts, and pseudogenes. *J Immunol* 172:6152–6157
- Doxiadis GGM, Otting N, de Groot NG, Bontrop RE (2001) Differential evolutionary MHC class II strategies in humans and rhesus macaques: relevance for biomedical studies. *Immunol Rev* 183:76–85
- Dugdale HL, Macdonald DW, Pope LC, Burke T (2007) Polygyny, extra-group paternity and multiple-paternity litters in European badger (*Meles meles*) social groups. *Mol Ecol* 16:5294–5306
- Dugdale HL, Macdonald DW, Pope LC, Johnson PJ, Burke T (2008) Reproductive skew and relatedness in social groups of European badgers, *Meles meles*. *Mol Ecol* 17:1815–1827
- Dugdale HL, Griffiths A, Macdonald DW (2011) Polygynandrous and repeated mounting behaviour in European badgers, *Meles meles*. *Anim Behav*. doi:10.1016/j.anbehav.2011.09.008
- DuPasquier L, Schwager J, Flajnik MF (1989) The immune system of *Xenopus*. *Annu Rev Immunol* 7:251–275
- Edwards SV, Hedrick PW (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol* 13:305–311
- Edwards JA, Jones DB, Evans PR, Smith JL (1985) Differential expression of HLA class II antigens on human fetal and adult lymphocytes and macrophages. *Immunology* 55:489–500
- Fernandez-Soria VM, Morales P, Castro MJ, Suarez B, Recio MJ, Moreno MA, Paz-Artal E, Arnaiz-Villena A (1998) Transcription and weak expression of HLA-DRB6: a gene with anomalies in exon 1 and other regions. *Immunogenetics* 48:16–21

- Flynn JJ, Finarelli JA, Zehr S, Hsu J, Nedbal MA (2005) Molecular phylogeny of the Carnivora (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. *Syst Biol* 54:317–337
- Furlong RF, Yang Z (2008) Diversifying and purifying selection in the peptide binding region of DRB in mammals. *J Mol Evol* 66:384–394
- Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* 57:1707–1722
- Hancox M (1980) Parasites and infectious diseases of the Eurasian badger (*Meles meles* L.): a review. *Mammal Rev* 10:151–162
- Handunnetthi L, Ramagopalan SV, Ebers GC, Knight JC (2010) Regulation of major histocompatibility complex class II gene expression, genetic variation and disease. *Genes Immun* 11:99–112
- Hughes AL, Nei M (1992) Maintenance of MHC polymorphism. *Nature* 355:402–403
- Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet* 32:415–434
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*. Academic Press, New York, pp 21–132
- Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56:683–695
- Kennedy LJ, Altet L, Angles JM, Barnes A, Carter SD, Francino O, Gerlach JA, Happ GM, Ollier WER, Polvi A, Thomson W, Wagner JL (1999) Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA Nomenclature Committee. *Tiss Antig* 54:312–321
- Kennedy LJ, Ryvar R, Gaskell RM, Addie DD, Willoughby K, Carter SD, Thomson W, Ollier WER, Radford AD (2002) Sequence analysis of MHC DRB alleles in domestic cats from the United Kingdom. *Immunogenetics* 54:348–352
- Kennedy LJ, Quarumby S, Happ GM, Barnes A, Ramsey IK, Dixon RM, Catchpole B, Rusbridge C, Graham PA, Hillbertz NS, Roethel C, Dodds WJ, Carmichael NG, Ollier WER (2006) Association of canine hypothyroidism with a common major histocompatibility complex DLA class II allele. *Tiss Antig* 68:82–86
- Klein J (1987) Origin of major histocompatibility complex polymorphism: the *trans*-species hypothesis. *Hum Immunol* 19:155–162
- Knapp LA (2007) Selection on MHC? A matter of form over function. *Heredity* 99:241–242
- Koepfli K-P, Deere KA, Slater GJ, Begg C, Begg K, Grassman L, Lucherini M, Veron G, Wayne RK (2008) Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biol* 6:10. doi:10.1186/1741-7007-1186-1110
- LeibundGut-Landmann S, Waldburger J, Krawczyk M, Otten LA, Suter T, Fontana A, Acha-Orbea H, Reith W (2004) Specificity and expression of CIITA, the master regulator of MHC class II genes. *Eur J Immunol* 34:1513–1525
- Lizundia R, Newman C, Buesching CD, Ngugi D, Blake D, Sin YW, Macdonald DW, Wilson A, McKeever D (2011) Evidence for a role of the host-specific flea (*Paraceras melis*) in the transmission of *Trypanosoma (Megatrypanum) pestanaei* to the European badger. *PLoS One* 6:e16977
- Macdonald DW, Newman C (2002) Population dynamics of badgers (*Meles meles*) in Oxfordshire, UK: numbers, density and cohort life histories, and a possible role of climate change in population growth. *J Zool* 256:121–138
- Macdonald DW, Anwar M, Newman C, Woodroffe R, Johnson PJ (1999) Inter-annual differences in the age-related prevalences of *Babesia* and *Trypanosoma* parasites of European badgers (*Meles meles*). *J Zool* 247:65–70
- Macdonald DW, Newman C, Dean J, Buesching CD, Johnson PJ (2004) The distribution of Eurasian badger, *Meles meles*, setts in a high density area: field observations contradict the sett dispersion hypothesis. *Oikos* 106:295–307
- Macdonald DW, Newman C, Buesching CD, Johnson PJ (2008) Male-biased movement in a high-density population of the Eurasian badger (*Meles meles*). *J Mammal* 89:1077–1086
- Macdonald DW, Newman C, Nouvellet PM, Buesching CD (2009) An analysis of Eurasian badger (*Meles meles*) population dynamics: implications for regulatory mechanisms. *J Mammal* 90:1392–1403
- Marmi J, Lopez-Giraldez JF, Domingo-Roura X (2004) Phylogeny, evolutionary history and taxonomy of the Mustelidae based on sequences of the cytochrome b gene and a complex repetitive flanking region. *Zool Scr* 33:481–499
- Mayer WE, O’huigin C, Klein J (1993) Resolution of the HLA-DRB6 puzzle: a case of grafting a de novo generated exon on an existing gene. *Proc Natl Acad Sci U S A* 90:10720–10724
- Messing J (1983) New M13 vectors for cloning. *Methods Enzymol* 101:20–78
- Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Newman C, Macdonald DW, Anwar MA (2001) Coccidiosis in the European badger, *Meles meles* in Wytham Woods: infection and consequences for growth and survival. *Parasitology* 123:133–142
- Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197–218
- Nowak MA, Tarczy-Hornoch K, Austyn JM (1992) The optimal number of major histocompatibility complex molecules in an individual. *Proc Natl Acad Sci U S A* 89:10896–10899
- Penn J, Potts WK (1999) The evolution of mating preferences and major histocompatibility complex genes. *Am Nat* 153:145–164
- Piontkivska H, Nei M (2003) Birth-and-death evolution in primate MHC class I genes: divergence time estimates. *Mol Biol Evol* 20:601–609
- Queller DC, Goodnight KF (1989) Estimating relatedness using genetic-markers. *Evolution* 43:258–275
- Reche PA, Reinherz EL (2003) Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *J Mol Biol* 331:623–641
- Rees LE, Ayoub O, Haverson K, Birchall MA, Bailey M (2003) Differential major histocompatibility complex class II locus expression on human laryngeal epithelium. *Clin Exp Immunol* 134:497–502
- Richman A (2000) Evolution of balanced genetic polymorphism. *Mol Ecol* 9:1953–1963
- Robbins F, Hurley CK, Tang T, Yao H, Lin YS, Wade J, Goeken N, Hartzman RJ (1997) Diversity associated with the second expressed HLA-DRB locus in the human population. *Immunogenetics* 46:104–110
- Robinson J, Waller M, Parham P, Bodmer JG, Marsh SGE (2001) IMGT/HLA database — a sequence database for the human major histocompatibility complex. *Nucleic Acids Res* 29:210–213
- Rogers LM, Cheeseman CL, Mallinson PJ (1997) The demography of a high-density badger (*Meles meles*) population in the west of England. *J Zool* 242:705–728
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Roper TJ (2010) *Badger*. Collins, London
- Rosalino LM, Macdonald DW, Santos-Reis M (2004) Spatial structure and land-cover use in a low-density Mediterranean population of Eurasian badgers. *Can J Zool* 82:1493–1502

- Rosalino LM, Torres J, Santos-Reis M (2006) A survey of helminth infection in Eurasian badgers (*Meles meles*) in relation to their foraging behaviour in a Mediterranean environment in southwest Portugal. *Eur J Wildl Res* 52:202–206
- Rozas J, Sánchez-DelBarrio J, Messenguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2469–2497
- San EDL (2007) Ectoparasite categories of the Eurasian Badger (*Meles meles*) in Western Switzerland. *Wildl Biol Pract* 3(2):52–59
- Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc R Soc B* 277:979–988
- Swain SL (1983) T cell subsets and the recognition of MHC class. *Immunol Rev* 74:129–142
- Takahashi K, Rooney AP, Nei M (2000) Origins and divergence times of mammalian class II MHC gene clusters. *J Hered* 91:198–204
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Thom MD, Johnson DDP, Macdonald DW (2004) The evolution and maintenance of delayed implantation in the Mustelidae (Mammalia: Carnivora). *Evolution* 58:175–183
- Villard J, Lisowska-Groszpiette B, vanden Elsen P, Fischer A, Reith W, Mach B (1997) Mutation of RFXAP, a regulator of MHC class II genes, in primary MHC class II deficiency. *N Engl J Med* 337:748–753
- Wagner JL (2003) Molecular organization of the canine major histocompatibility complex. *J Hered* 94:23–26
- Wagner JL, Burnett RC, Storb R (1996) Molecular analysis of the DLA-DR region. *Tiss Antig* 48:549–553
- Wagner JL, Burnett RC, Storb R (1999) Organization of the canine major histocompatibility complex: current perspectives. *J Hered* 90:35–38
- Weber DS, Stewart BS, Schienman J, Lehman N (2004) Major histocompatibility complex variation at three class II loci in the northern elephant seal. *Mol Ecol* 13:711–718
- Yamaguchi N, Dugdale HL, Macdonald DW (2006) Female receptivity, embryonic diapause, and superfetation in the European badger (*Meles meles*): implications for the reproductive tactics of males and females. *Q Rev Biol* 81:33–48
- Yang Z (2002) Inference of selection from multiple species alignments. *Curr Opin Genet Dev* 12:688–694
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591
- Yang Z, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22:1107–1118
- Yonezawa T, Nikaido M, Kohno N, Fukumoto Y, Okada N, Hasegawa M (2007) Molecular phylogenetic study on the origin and evolution of Mustelidae. *Gene* 396:1–12
- Yuhki N, Beck T, Stephens RM, Nishigaki Y, Newmann K, O'Brien SJ (2003) Comparative genome organization of human, murine, and feline MHC class II region. *Genome Res* 13:1169–1179