Evolution of MHC class I genes in the European badger (Meles meles)

Yung Wa Sin1,2, Hannah L. Dugdale2,3,4, Chris Newman1, David W. Macdonald1 & Terry Burke2

1Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, Recanati-Kaplan Centre, Tubney House, Abingdon Road, Tubney, Abingdon, Oxfordshire OX13 5QL, United Kingdom
2NERC Biomolecular Analysis Facility, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, South Yorkshire, S10 2TN, United Kingdom
3Behavioural Ecology and Self-Organization, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands
4Theoretical Biology, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands

Keywords
Balancing selection, birth-and-death evolution, concerted evolution, major histocompatibility complex, orthology, trans-species polymorphism.

Abstract
The major histocompatibility complex (MHC) plays a central role in the adaptive immune system and provides a good model with which to understand the evolutionary processes underlying functional genes. Trans-species polymorphism and orthology are both commonly found in MHC genes; however, mammalian MHC class I genes tend to cluster by species. Concerted evolution has the potential to homogenize different loci, whereas birth-and-death evolution can lead to the loss of orthologs; both processes result in monophyletic groups within species. Studies investigating the evolution of MHC class I genes have been biased toward a few particular taxa and model species. We present the first study of MHC class I genes in a species from the superfamily Musteloidea. The European badger (Meles meles) exhibits moderate variation in MHC class I sequences when compared to other carnivores. We identified seven putatively functional sequences and nine pseudo-genes from genomic (gDNA) and complementary (cDNA) DNA, signifying at least two functional class I loci. We found evidence for separate evolutionary histories of the α1a and α2/α3 domains. In the α1 domain, several sequences from different species were more closely related to each other than to sequences from the same species, resembling orthology or trans-species polymorphism. Balancing selection and probable recombination maintain genetic diversity in the α1 domain, evidenced by the detection of positive selection and a recombination event. By comparison, two recombination breakpoints indicate that the α2/α3 domains have most likely undergone concerted evolution, where recombination has homogenized the α2/α3 domains between genes, leading to species-specific clusters of sequences. Our findings highlight the importance of analyzing MHC domains separately.

Introduction
The major histocompatibility complex (MHC) is of particular importance to the study of evolutionary genetics owing to its pattern of molecular evolution. MHC is a diverse gene family that plays a crucial role in the vertebrate adaptive immune system and in autoimmunity. Cell surface glycoproteins, encoded by the MHC genes, are vital in both humoral and cell-mediated immune responses, as they bind and present antigens to T cells and trigger an immune cascade (Swain 1983). MHC genes are classified into groups including class I and class II. MHC class II molecules 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). MHC class I genes comprise classical (class Ia) and nonclassical (class Ib) loci that differ in polymorphism, structure, function, and expression pattern (Parham and Ohta 1996; Rodgers and Cook 2005). MHC class Ia molecules are responsible primarily for intracellular antigen binding and are expressed on the surface of all
nucleated somatic cells (Bjorkman and Parham 1990). Because of this crucial role in the immune system, MHC genes are under constant selective pressures due to challenges from parasites and pathogens (Jeffrey and Bangham 2000; Piertney and Oliver 2006). This arms race between pathogens and hosts is posited to be the driving force for the extreme diversity in MHC genes (e.g., class Ia genes such as HLA-A, HLA-B, and HLA-C in humans; Piertney and Oliver 2006).

The high degree of diversity among MHC alleles arises from the diverse exons that encode the domains forming the antigen-binding site (ABS; Hughes and Yeager 1998). The nucleotide diversity within the MHC genes has been attributed to balancing selection (Bergström and Gylensten 1995), such as overdominance and frequency-dependent selection, which act to maintain large numbers of alleles in populations. The persistence of ancestral allelic diversity over long periods of time, relative to neutral genetic variation (Richman 2000), is enhanced substantially by balancing selection, which leads to high levels of allelic diversity within species (Hughes and Yeager 1998; Penn et al. 2002). Some mammalian MHC allelic lineages are more than a million years old and are maintained after speciation (Figueroa et al. 1988). Phylogenetic reconstructions therefore reveal trans-species polymorphism (Klein 1987; Klein et al. 1998, 2007), where alleles between species are more closely related (or even identical) than alleles within species. Phylogenetic reconstructions can also reveal orthologous relationships, by which sequences group with more than 50 alleles identified (Wagner 2003). Studies with more than 50 alleles identified (Wagner 2003). Studies of the MHC class I genes in many carnivores, such as those in the superfamily Musteloidea, are however lacking. Here, we characterize the MHC class I genes of the European badger (Meles meles). Meles meles is well suited to investigate how MHC selection and conferred immunological advantages (e.g., pathogen and parasite resistance) are regulated by mate choice in the wild. Meles 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 endoparasitic disease studies (Macdonald et al. 1999; Anwar et al. 2000, 2006; Newman et al. 2001; Rosalino et al. 2006; Nouvellet et al. 2010; Lizundia et al. 2011). In particular, M. meles is a wildlife reservoir of Mycobacterium bovis (Delahay et al. 2001; Mathews et al. 2006; Riordan et al. 2011); an intracellular bacteria that is the cause of bovine tuberculosis (bTB) in cattle and wildlife. MHC class I-dependent immunity is known to play an important role in the eradication of M. bovis in mice (Ladel et al. 1995). Meles meles with different MHC genotypes, or the presence/absence of certain MHC alleles, may therefore...
have differential susceptibilities to *M. bovis*, which could contribute to genetic-based bTB control strategies in badgers.

In this study we: (1) characterized the MHC class I genes of *M. meles* from a high-density population and tested for evidence of selection and recombination; (2) identified the transcription pattern by comparing genomic DNA (gDNA) and complementary DNA (cDNA) sequences from whole blood samples, which is important as MHC genes identified using gDNA may be nonfunctional; and (3) performed phylogenetic analyses to investigate whether *M. meles* sequences belong to monophyletic groups, or whether sequences transcend species boundaries. Characterization of MHC class I genes in *M. meles* will clarify whether these have a more rapid turnover rate than class II loci (Sin et al. 2012) and facilitate elucidation of the underlying evolutionary processes within different regions. Moreover, the development of MHC markers will facilitate studies on the relationship between genetics and disease in this controversial animal, as well as other closely related species.

**Materials and Methods**

**Sample collection and nucleic acid isolation**

Blood samples were collected from 11 badgers that resided in eight different social groups (Sin et al. 2012) in Wytham Woods, Oxfordshire, UK (global positioning system reference 51°46′26″N, 1°19′19″W). All trapping and handling protocols are detailed in Macdonald and Newman (2002). These protocols were subject to ethical review and were performed under Natural England Licence (currently 20104655) and UK Home Office Licence (PPL 30/2835). 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. 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 11 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). Methods of cDNA synthesis are detailed in Sin et al. (2012).

**Primer design and polymerase chain reaction (PCR) amplification**

MHC class I molecules are heterodimers, that consist of an α chain and a β2-microglobulin (β2m) molecule. The α chain is composed of three extracellular domains α1, α2, and α3 (encoded by exon 2, 3, and 4, respectively), a transmembrane domain (exon 5), and a cytoplasmic domain (exon 6; Fig. 1; Bjorkman and Parham 1990). The α1 and α2 domains are α chain regions that comprise specific sites forming the ABS.

To amplify the class I genes from *M. meles*, we tested published primers used successfully in other carnivores (Zhong et al. 1998; Aldridge et al. 2006). Oligonucleotide primers, which recognize highly conserved regions of the MHC class I genes, were also designed using OligoAnalyzer 3.1 (Owczarzy et al. 2008), based on alignments with GenBank’s nucleotide sequences from domestic dog (*Canis lupus familiaris*; AF218297, AF218299, AF218301, AF218303, DQ056267, DQ056268, and M32283), harbour seal (*Phoca vitulina*; U88874), domestic cat (Felis catus; M26318, U07670, U07672 and U07674), horse (*Equus caballus*; NM001123381), and human (*Homo sapiens*; AF287959, U03907, NM002117, and NM002127).

Using these primers on 10–30 ng of cDNA/gDNA, PCR amplification 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× PCR buffer (containing MgCl2; Qiagen), and 2 units of HotStarTaq (Qiagen). The PCR cycle began with incubation at 94°C for 15 min, followed by 35 incubation cycles at 94°C for 30 sec, annealing temperature (Table 1) for 30 sec, and 72°C for 60–90 sec according to amplicon length (60 sec for amplification of exon 2 or exon 3 only), ending with an extension step at 72°C for 10 min. The PCR products were electrophoresed on a 1.5% agarose gel and visualized using ultraviolet 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 cloned using the pGEM-T Easy Vector Systems (Promega, Madison, WI). The cloning and sequencing procedure is detailed in Sin et al. (2012). Between nine and 72 clones were sequenced for each individual.

Figure 1. Schematic representation of the positions of the primers used for amplification of MHC class I sequences from cDNA/gDNA. α1–3 domains are labeled at the exons encoding them.
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 site substitutions were analyzed using CodonCode Aligner 3.7.1 (CodonCode Corporation, Dedham, MA) and were compared with known MHC class I sequences using the NCBI BLAST program (Altschul et al. 1990). The DNA sequences from *M. meles* were assigned the GenBank accession numbers JQ425427—JQ425447.

### Data Analyses

#### Selection and recombination

Selection at the amino acid level was measured as the rates of nonsynonymous ($d_N$) and synonymous ($d_S$) substitutions per codon site, 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 derived from 1000 bootstrap replicates. Synonymous and nonsynonymous substitutions were calculated separately for the ABS and non-ABS, as determined by Bjorkman et al. (1987). CODEML in PAML 4.4b (Yang 2007) was used to check for positively selected sites (PSS) in the α1 and α2 domains, which are indicated where the ratio $ω$ (nonsynonymous/synonymous substitution rate ratio, $d_N/d_S$) exceeds 1, meaning that nucleotide mutations that alter the amino acid sequence of a protein occur more frequently than nucleotide mutations that do not alter amino acids. Thus, $ω > 1$ is indicative of positive selection whereby beneficial amino acid changes are fixed. Codon-based likelihood analysis was used to test for evidence of positive selection, using several models: M1a (nearly neutral), M2a (positive selection), M7 (beta), and M8 (beta and $ω$). The assumptions of these models are detailed in Yang et al. (2000, 2005). Two null models of neutral evolution (M1a and M7) were applied and compared against their nested models, which allow stringent testing for positive selection (M2a and M8, respectively, which assume a different distribution of mutations; Anisimova et al. 2001, 2002). We determined whether the alternative models (M2a and M8) provided a significantly improved fit, versus their null models (M1a and M7, respectively), using a likelihood ratio test (LRT), which compares twice the difference of the log likelihood ratios ($2\Delta\ln L$) to a $\chi^2$ distribution. To identify codons under positive selection (posterior probabilities >0.95), through comparisons of M1a versus M2a and M7 versus M8, CODEML was used to calculate the Bayes Empirical Bayes (BEB) posterior probabilities (Yang et al. 2005) at each codon. Different domains were analyzed separately, as their evolutionary histories could be different due to recombination events in the intervening intron.

Recombination analyses were performed on the nucleotide alignment spanning exon 2, intron 2, and exon 3 using RDP3 alpha 44 (Martin et al. 2010). Four methods, that is, RDP (Martin and Rybicki 2000), GENECONV (Padidam et al. 1999), MaxChi (Smith 1992), and Bootscan (Martin et al. 2005), were applied in the first run to detect recombination events. Default settings were applied with a maximum $P$-value of 0.05, applying Bonferroni correction for multiple comparisons. Any recombination signals, which were detected by at least three methods, were then rechecked with all available methods (Martin et al. 2010). Any putative recombination events detected were verified further by examination of MaxChi plots and matrices and Neighbor-Joining trees from the inferred fragments, to assess recombinant designation and breakpoint placement. Only the recombination events that were confirmed after these procedures were considered significant. The effects of recombination and gene conversion on sequence evolution are similar in small sequence fragments, therefore we did not differentiate between them, and we refer to them as recombination sensu lato hereafter (Richman et al. 2003; Burri et al. 2008).

### Phylogenetic Analyses

Phylogenetic analyses were performed on the consensus alignments of *M. meles* MHC class I exon 2, exon 3, and exon 4 sequences, against sequences from other species available in GenBank (e.g., *C. lupus familiaris* [accession number: M32283, NM001014378, NM001014379, NM001014767, NM001020810, U55029], *P. vitulina* [PVU88874], *F. catus* [U07672, U07674], *E. caballus* [LOC100056062], *H. sapiens* [AF287959, NM002117.4, NM002127.5, U03907], *Ailuropes melanoleuca* [EU162658, EU162659], *Leopardus...
were highly divergent from the
Meme-MHC I in exon 2, exon 3, and exon 4, re-
signifies pseudogenes with a frameshift,
Monachus schauinslandi
was identical (Figs. 3 and 4C) yet
Bos taurus
was highly divergent (9.3
were identified and expressed in
0.84%); 7.0% (SE
were only de-
Acinonyx jubatus
and
and
were only de-
1648
[07678], Acinonyx jubatus [07666], Bos taurus
[AB245424], and Monachus schauinslandi [Aldridge et al.
2006]). Tree reconstruction was performed separately on
three domains, in order to maximize the detection of differ-
ent evolutionary histories due to genetic exchange. Domain
borders were assigned according to Koller and Orr (1985).

Phylogenetic networks allow visualization of reticulate
phylogenetic signals (Huson and Bryant 2006) whereas
phylogenetics trees may poorly describe complex evolution-
ary scenarios; thus, phylogenetic networks provide an effec-
tive way of evaluating evolutionary relationships involving
gene duplication and recombination. We used the Neighbor-
Net algorithm in SplitsTree 4.12.3 (Huson and Bryant 2006)
to analyze the phylogenetic relationships for exon 2, exon 3,
and intron 2. We constructed a Neighbor-Net network based
uncorrected p-distances and conducted 1000 bootstrap
replicates to estimate their support.

Bayesian phylogenetic inference was performed using
MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003) to ana-
lyze the phylogenetic relationship of exon 4. We used the
web-based application FindModel (http://www.hiv.lanl.gov/ content/sequence/findmodel/findmodel.html) to ascertain
the best-fit model of nucleotide substitution, which was iden-
tified as the Hasegawa–Kishino–Yano plus gamma model
(HKY+Γ). A Markov chain Monte Carlo (MCMC) search
was initiated with random trees and run for 3,000,000 gener-
ations, sampling every 100 generations. Data were partitioned
by gene codon positions. Two separate analyses and four in-
dependent chains were executed. Convergence was indicated
when the average standard deviation of split frequencies was
less than 0.01 (Ronquist et al. 2005). We also checked for con-
vergence by plotting the likelihood scores against generations
and discarded the first 25% of the generations as “burn-in.”

Results

Diversity and transcription of MHC class I sequences

Sixteen different sequences were isolated from gDNA and
cDNA, using the primers detailed in Table 1, Figure 1, and
Table S1. A maximum of four putatively functional sequences
was derived from a single individual, indicating the pres-
ence of at least two loci. Transcription analysis showed that
nine of 16 sequences detected in the gDNA were also ampli-
fied from the cDNA, which was isolated from whole blood
(Fig. 2; Table 2), while four sequences were detected only in
the gDNA (Fig. 2).

Pseudogene features (Fig. 2) were found in four se-
quencies: nucleotide deletions caused a frameshift for Meme-
MHC I*PS01, Meme-MHC I*PS02, Meme-MHC I*PS03, and
Meme-MHC I*PS04N, which caused premature stop codons
(Fig. 2; PS signifies pseudogenes with a frameshift, N sig-
nifies the presence of genomic sequences, where no cDNA
sequences were found). Pseudogene Meme-MHC I*PS04N
was detected only in gDNA, whereas Meme-MHC I*PS01,
Meme-MHC I*PS02, and Meme-MHC I*PS03 were only de-
tected in the cDNA, indicating the presence of a transcribed
nonfunctional pseudogene (Mayer et al. 1993; Fernandez-
Soria et al. 1998). The nucleotide deletion in Meme-MHC
I*PS01 and Meme-MHC I*PS03 occurred at the 5’ primer
annealing site for exon 2 amplification from gDNA; thus, no
sequence was detected from gDNA. Nucleotide insertions
or deletions were detected in two of four sequences that
only amplified in the exon 2 region (Meme-MHC I*09N,
Meme-MHC I*PS10, Meme-MHC I*11, and Meme-MHC
I*PS12N; Fig. 2). The intron 2 of Meme-MHC I*PS08N
is 45–56 bp longer than other putatively functional sequences; a
longer intron was also found in a DLA pseudogene (DLA-53;
Wagner 2003; Figs. 3 and 4C). Accordingly, the above se-
quencies were regarded as pseudogenes and were not included
in the nucleotide substitution calculations. Only sequences
without pseudogene features, and with both exon 2 and exon
3 amplified from both gDNA and cDNA, were regarded as
functional sequences and included in further analyses of
nucleotide substitution and selection (i.e., Meme-MHC
I*01–Meme-MHC I*07; Table 2; Fig. 2).

Of the seven putatively functional sequences, all were de-
tected in both gDNA and cDNA (Table 2; Fig. 2). Meme-MHC
I*05 was detected in all individuals. By contrast, Meme-MHC
I*01 and Meme-MHC I*02 were identified and expressed in
only one individual in this study (Table 2). One thousand one
hundred and fifty-three individuals from the same popula-
tion have been genotyped subsequently and these sequences
have been identified in more individuals (330 for Meme-
MHC I*01 and 132 for Meme-MHC I*02; Y. W. Sin, unpubl.
data).

The majority of variable sites in the identified sequences
were in exon 2 and exon 3, where most mutations repre-
sented a nonsynonymous nucleotide substitution (Table 3).
As a consequence, there were more polymorphic amino acid
residues among the α1 and α2 domains (Fig. 2; Table 3),
which are involved in antigen binding. The average sequence
divergence was 6.5% (SE = 0.62%); 7.0% (SE = 0.84%); and
1.0% (SE = 0.25%) in exon 2, exon 3, and exon 4, re-
spectively. Two of the sequences, Meme-MHC I*02 and 04,
were highly similar, differing from each other at only one
nucleotide position in exon 2 (Fig. 4A) and five positions
in exon 3 (Fig. 4B). Meme-MHC I*03 was highly divergent
(9.3–9.6%) from Meme-MHC I*01, 02 and 04; Meme-MHC
I*07 was highly divergent (9.3–9.6%) from Meme-MHC
I*02 and 04, in the exon 2 region. In the exon 3 region,
Meme-MHC I*05 and 06 were highly divergent from the
other five sequences (8.8–12.3%). The intron 2 region of
Meme-MHC I*05 and 06 was identical (Figs. 3 and 4C) yet
both of them were highly divergent from other putatively
functional sequences (28.3–30.1%).

1648
© 2012 The Authors. Ecology and Evolution published by Blackwell Publishing Ltd.
Figure 2. Amino acid sequence identity for class I sequences of *Meles meles* and seven other mammals (*Phoca vitulina*, *Ailuropoda melanoleuca*, *Canis lupus familiaris*, *Felis catus*, *Acinonyx jubatus*, *Equus caballus*, and *Homo sapiens*; GenBank accession numbers are provided in the Materials and Methods). The complete amino acid sequence of *Meme-MHC I*\(^*\)03 is shown. PS signifies pseudogenes with a frameshift. N signifies the presence of genomic sequences, where no cDNA sequences were found. Single letters and dots within the alignment/sequence represent amino acids that are distinct from or identical to *Meme-MHC I*\(^*\)03, respectively. Dashes (−) indicate missing sequences. Numbers above the sequence indicate the codon position. Arrows above the sequence label the beginning of a domain. Asterisks (∗) indicate amino acid residues pointing toward the postulated antigen-binding site, which were defined according to Bjorkman et al. (1987). Carets (∧) indicate residues pointing up on an alpha-helix, postulated to interact with peptides and/or T-cell receptors (TCRs), and dots above the sequence (.) indicate residues on an alpha-helix that is pointing away from the antigen-binding site, postulated to interact with TCRs (Bjorkman et al. 1987). Conserved sites that bind the peptide N- and C-termini are marked with gray boxes. The location of the N-linked glycosylation site is at position 88 (CHO). Disulphide bonds formed between cysteine residues are shown with a line spanning the two cysteine residues. Residues that form the β-sheet or α-helix, and residues that influence the binding of the CD8 glycoprotein, are marked under the alignment. The columns on the left of the sequences indicate whether the sequences were found in gDNA and/or cDNA.
Table 2. Presence of MHC class I sequences in genomic DNA (gDNA) and complementary DNA (cDNA) from 11 Meles meles. The total number of clones with different sequences found is given. Total number of clones: numbers out of brackets are clones of polymerase chain reaction (PCR) products from gDNA (numbers in brackets are clones of PCR products from cDNA).

| Individual        | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | All total no. of clones |
|-------------------|---|---|---|---|---|---|---|---|---|----|----|------------------------|
| Meme-MHC*I*01    | X |   |   |   |   |   |   | X |   |    |    | 3 (2)                  |
| Meme-MHC*I*02    |   | X |   |   |   |   |   |   |   |    |    | 8 (5)                  |
| Meme-MHC*I*03    |   |   | X |   |   |   |   |   |   |    |    | 3 (27)                 |
| Meme-MHC*I*04    |   | X | X | X | X | X |   | X |   |    |    | 7 (72)                 |
| Meme-MHC*I*05    | X | X | X | X | X | X | X | X |   |    |    | 88 (49)                |
| Meme-MHC*I*06    |   | X |   |   |   |   |   |   | X |    |    | 12 (14)                |
| Meme-MHC*I*07    | X |   |   |   |   |   |   |   | X |    |    | 5 (26)                 |

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

The M. meles sequences also coded for other conserved residues, characteristic of class I molecules, which are important for the molecular structure and peptide binding.

Figure 3. Nucleotide sequence identity for the MHC class I intron 2 of Meles meles clones and C. lupus familiaris. The GenBank accession numbers for sequences from C. lupus familiaris are NW876254.1 and US50209.1. The complete nucleotide sequence of Meme-MHC*I*03 is shown; numbers above the sequence indicate the nucleotide position. PS signifies pseudogenes with a frameshift. N signifies the presence of genomic sequences, where no cDNA sequences were found. Single letters and dots (.) represent nucleotides that are distinct from or identical to Meme-MHC*I*03, respectively. Dashes (-) indicate missing sequences. The degenerate 13-bp sequence motif (CCNCNNTNNCCNC) that is crucial in crossover events at human recombination hotspots (Myers et al. 2008) is marked with gray boxes. The nucleotide breakpoint for Meme-MHC*I*05 and Meme-MHC*I*06 is marked with an asterisk (*) above the alignment.
Phylogenetic analyses

The phylogenetic tree of α1 (Fig. 4A) highlights that sequences from *M. meles* did not form a monophyletic clade, rather they were intermingled with phocine, canine, and ursine sequences. No sequences from the other species included in the analysis formed a monophyletic clade, except for those from *H. sapiens* and *M. schauinslandi*. But the support of clades with mixed sequences from difference species was not strong (e.g., bootstrap value = 56 for the clustering of *Aime-128* and *Meme-MHC I*^*01*, *02*, and *04*). Some *M. meles* pseudogenes, however, formed highly or moderately supported clades without grouping with any putatively functional sequences (e.g., *Meme-MHC I*^*0PS02* and *Meme-MHC I*^*0PS03*; *Meme-MHC I*^*0PS08N*, *Meme-MHC I*^*09N*, and *Meme-MHC I*^*0PS10*). For the sequences encoding the α2 domain (Fig. 4B), all *M. meles* sequences, except for *Meme-MHC I*^*0PS08N*, formed a monophyletic clade. Within this clade, pseudogenes *Meme-MHC I*^*0PS04N*, *Meme-MHC I*^*0PS02*, and *Meme-MHC I*^*0PS03* grouped together. Sequences from *A. melanoleuca* and *H. sapiens* also formed monophyletic clades. For α3 sequences, *M. meles* formed a distinct and highly supported clade (posterior probability [PP] support = 1.0; Fig. 7) that grouped together with *P. vitulina* and *A. melanoleuca* (PP = 0.67). This clade further grouped with all carnivores in our comparison including canine and feline sequences (PP = 0.86), separating it from equine, bovine, and human sequences.

Discussion

Diversity and transcription of MHC class I sequences

This is the first study to characterize MHC class I genes in *M. meles*. Moreover, this work was performed using both the genome and transcriptome. The sequences identified included pseudogenes and putatively functional sequences that encode domains necessary to form functional class I molecules. The major structural features (Kaufman et al. 1994) that distinguish class Iα molecules are all present in these putatively functional sequences, including highly conserved amino acid residues that bind the N- and C-termini of the peptide, cysteine (C) residues that form the disulphide bonds, an N-linked glycosylation site, a threonine (T) residue for interaction with TAP complex, and regions to interact with CD8 glycoproteins on the T cells. In addition, we demonstrate that all putatively functional sequences were expressed in the RNA level in whole blood. These features, together with the detected polymorphisms, indicate that these sequences belong to at least two class Iα loci, in contrast to class Ib genes that are typically monomorphic (Rodgers and Cook 2005).
The variability in the number of class I sequences in *M. meles* is intermediate compared to other carnivores. The most closely related species for which class I genes have been characterized is the Hawaiian monk seal (*M. schauinslandi*; Aldridge et al. 2006), which has at least two loci. No variability was found, however, within more than 80 individuals of this endangered seal species. Within the infraorder Arctoidea, the giant panda (*A. melanoleuca*) has two classical genes (*Aime*-152 and *Aime*-128; Pan et al. 2008), plus a non-classical gene closely related to DLA-79. *Aime*-152 appears to be monomorphic, while nine *Aime*-128 sequences were found in five individuals (Pan et al. 2008). The carnivores that, to date, have had their MHC organization characterized most thoroughly are *C. lupus familiaris* and *F. catus*. In the feline leukocyte antigen (FLA), there are at least three class Ia genes, in common with human and murine MHCs (Yuhki et al. 2007). In DLA (Burnett et al. 1997; Kennedy et al. 1999; Wagner et al. 1999; Wagner 2003), there is one class Ia gene (DLA-88), three class Ib genes (DLA-79, -12, and -64), and two pseudogenes (DLA-53 and -12a). Among these, DLA-88 is the most polymorphic gene, with more than 50 alleles identified (Wagner 2003). As discussed above, the number of class I genes and their organization can differ between species greatly (Kelley et al. 2005). Mammals usually possess...
1–3 class I genes (Nei et al. 1997) and have a variable range of polymorphism (Parham and Ohta 1996; Wagner 2003). Considering the extensive geographical range of *M. meles* and their considerable socio-spatial variability (Macdonald et al. 2004; Rosalino et al. 2004; Newman et al. 2011), it is highly likely that more sequences will be detected as and when other populations are examined.

Our transcription analysis demonstrated that not all the detected sequences were expressed in whole blood. In addition to the seven putatively functional sequences, for which the sequences detected from gDNA and cDNA were identical; nine sequences were identified as pseudogenes. The phylogenetic analyses indicate that *Meme-MHC I* *PS02*, *Meme-MHC I* *PS03*, and *Meme-MHC I* *PS04N*, which form a
Evolution of MHC Class I Genes in *Meles meles*

Y. W. Sin et al.

**Figure 5.** Rates (± standard error) of nonsynonymous (*d*_N; solid marker) and synonymous (*d*_S; open marker) substitutions and ratio of *d*_N to *d*_S (*ω*; gray bar) for antigen-binding site (ABS), non-ABS, and combined (ABS + non-ABS) at the three α domains of the *Meles meles* MHC class I loci. The number of codons for each region is given under the x-axis. Asterisk (‘∗’) indicates no ABS in α3 domain.

strongly supported clade, belong to the same pseudogene locus, with *Meme-MHC I*^∗^PS02 and *Meme-MHC I*^∗^PS03 both present in two individuals and *Meme-MHC I*^∗^PS04N detected in another two individuals. Another grouping of *Meme-MHC I*^∗^PS08N, *Meme-MHC I*^∗^09N, and *Meme-MHC I*^∗^PS10 indicated that these belong to two closely related pseudogene loci, as 1−3 sequences of this group were detected from seven individuals separately. Other studies have shown that many MHC sequences can be detected at the genomic level, but not at the cDNA level (de Groot et al. 2004). The presence of expressed nonfunctional MHC pseudogenes has also been reported (Mayer et al. 1993; Fernandez-Soria et al. 1998), even in the class II genes of *M. meles* (Sin et al. 2012). This is concordant with the finding that the MHC class I and class II regions have duplicated many times, generating many pseudogenes in addition to novel functional genes (Beck et al. 1999).

Evolution of the MHC class I genes

Our phylogenetic analyses reveal the evolutionary histories of the class I domains we examined to be very different from each other. Within the α1 domain, there were more nonsynonymous than synonymous nucleotide substitutions in both the ABS and non-ABS, contributing to a higher nucleotide and amino acid sequence diversity. Evidence of positive selection was detected in the α1 domain, in which PSS were all within the ABS (Fig. 6A). *Meles meles* sequences of the α1 domain were not monophyletic within this species but intermingled with sequences from other species (e.g., *P. vitulina*); a phenomenon characteristic of MHC genes. This could be due to trans-species polymorphism (Klein 1987; Hughes and Yeager 1998; Klein et al. 1998), whereby balancing selection maintains this ancestral variation over a long period of time (many generations), even after species divergence (Penn and Potts 1999; Bernatchez and Landry 2003), leading to differences between gene tree and species tree. As the α1 domain is responsible for antigen binding, the diversifying and balancing selection that drives and maintains this ABS polymorphism permits a population to present a wider repertoire of antigens, thus increasing its ability to combat pathogenic and parasitic infections (Hughes and Nei 1992; Hughes and Yeager 1998). The clustering of sequences among species could also be due to orthology, whereby sequences from an orthologous gene will cluster together, which may produce a similar clustering pattern as that observed with trans-species polymorphism. Without the availability of information about loci identity, it was not possible to disentangle between signals for orthology and trans-species polymorphism. Recombination between exons that encode ABSs can also increase allelic diversity in MHC genes (Ohta 1991; Jakobsen et al. 1998; Shum et al. 2001; Bos and Waldman 2006). We detected a recombination event at the 3′ end of exon 2 that may have increased variation in *M. meles* class I.
Evolution of MHC Class I Genes in *Meles meles*

Table 4. Positively selected sites (PSS), parameter estimates, log-likelihood values of different models of codon evolution, and summary of test statistics for the likelihood-ratio test (LRT) of the α-1 and α-2 domain of MHC class I genes in *Meles meles*.

| Domain | Model name | Log-likelihood | Parameter estimates | PSS | LRT | Test statistic | P-value |
|--------|------------|----------------|--------------------|-----|-----|----------------|---------|
| α-1    | M1a        | −588.31        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.78          | <0.002  |
|        | M2a        | −581.92        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |
|        | M7 beta    | −581.95        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 11.36          | 0.001   |
|        | M8 beta and | −588.32        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |
| α-2    | M1a        | −556.92        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |
|        | M2a        | −556.92        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |
|        | M8 beta    | −556.92        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |
|        | M8 beta and | −556.92        | α0 = 0.51, α1 = 0.04 | 0   | 0   | 12.74          | 0.002   |

Log-likelihood values and estimated parameters for each model were calculated using CODEML in PAML 4.4b (Yang 2007). ω0 and ω1 are the parameters of the beta distribution. PSS were identified in models M2a and M8 by Bayes Empirical Bayes (BEB) with posterior probability greater than 0.95 (Yang et al. 2005); sites with PP > 0.95 are shown in bold. *NS* indicates that the test statistic to a χ2 distribution.

As is the case for α-1 domain, the α-2 domain functions as an antigen-binding domain and showed a higher non-synonymous/synonymous rate ratio (ω) within the ABS than in the non-ABS (Fig. 6B). Comparisons of maximum likelihood models, which allow for positive selection (M2a and M8) relative to their corresponding null models, detected no significant positive selection acting on this region, however. A possible explanation for less intense diversifying selection on the α-2 domain would be that it has more conserved sites than the α-1 domain and so purifying selection may act to maintain the structural features of the MHC molecule. Our phylogenetic analyses also showed that all *M. meles* α-2 domain sequences were grouped together except for one pseudogene (*Meme-MHC 1*PS08N clustering with DLA-64 but with low bootstrap support [72.4]; Fig. 4B); a very different pattern to the pattern that we observed in the α-1 domain. The birth-and-death model of evolution predicts that this type of clustering by species (Nei et al. 1997; Piontkivska and Nei 2003) will result in less orthologous relationships, due to higher rates of gene duplication and deletion. Mechanistically, a differential rate of birth-and-death evolution explains the more rapid turnover rate of class I genes than class II genes in mammals (Hughes and Nei 1989; Piontkivska and Nei 2003) and possibly reveals why sequences are more closely related within *M. meles* than to sequences from other species. The α-1 sequences we observed, however, showed that ancestral polymorphism or orthology at multiple lineages was maintained after the analyzed species diverged from one another, which indicates that different selective pressures act on different class I gene regions. Under this birth-and-death model, balancing selection at the α-1 domain should maintain ancestral polymorphism even after species divergence, while strong purifying selection and recent duplication are required to produce a species-specific clustering at the α-2 domain.

Alternatively, and more plausibly, species-specific gene clusters may result from concerted evolution (Hess and Edwards 2002; Joly and Rouillon 2006), where genes are homogenized by recombination. Recombination in the α-2 coding region could produce *M. meles* sequences that are more similar to one another than to the sequences from other species. Under this scenario, a recombination breakpoint is needed to separate selection on the α-1 and α-2 regions, given that they exhibit different evolutionary histories. Our recombination analysis exposed that there are indeed recombination breakpoints located in the middle of the α-1 domain and in the middle of intron 2, which is between the α-1 and α-2 coding exons. These recombination breakpoints indicate the probable crossover location and they most likely
contributed to the independent evolutionary histories of the α1 and α2 domains by recombination in the α2 encoding region. Recombination within intron 2 has also been reported in other species (Holmes and Parham 1985; Shum et al. 2001; Bos and Waldman 2006), which gives greater support to the hypothesis of concerted evolution through shuffling of the α1 and α2 domains. These recombination breakpoints are shared among some M. meles sequences, demonstrating that
breakpoints are not random and that these recombination events are due to in vivo recombination but not in vitro chimera formation. Coincidentally, the degenerate 13-bp sequence motif, which is crucial in the crossover event of human recombination hotspots (Myers et al. 2008), was also found in the M. meles intron 2 sequences in which the recombination breakpoints were identified (Fig. 3). Variation in the zinc-finger protein PRDM9, however, affects species-specific binding to the sequence motif (Myers et al. 2010); hence, more studies are needed to elucidate the recombination mechanism in M. meles.

Phylogenetic relationships of the α3 encoding sequences were more similar to that of the α2 than the α1 domain, demonstrating species-specific clustering of M. meles α3 sequences. This indicates that a recombination event most plausibly separated the α1 domain and the remaining 3’ segments of the gene, leading to inconsistencies between the α1 gene tree and the α2 and α3 gene trees. However, the α3 domain is not antigen binding and it shows a very low number of mutations and polymorphic amino acid residues. Compared with the antigen-binding domains that are under positive selection, which increases diversity, strong purifying selection could eliminate mutations leading to unfavorable changes in the molecular structure or the regions that interact with T cells (Kaufman et al. 1994). The relatively conserved α3 domain is thus used to reconstruct evolutionary relationships more frequently than the α1 and α2 domains (e.g., Glaberman and Caccone 2008). Here, the phylogenetic relationships of the α3 sequences follow the molecular phylogeny of the extant Carnivora (Flynn et al. 2005; Fulton and Strobeck 2006), in which Musteloidea, Pinnipedia, and Ursoidea together form the infraorder Arctoidea. This infraorder then groups with Canidae to form the suborder Caniformia and together
with Feliformia comprise the order Carnivora. An exception is *P. vitulina* that forms a clade with *A. melanoleuca* instead of *M. meles*. Musteloidea diverged from Ursoidae and Pinnipedia around 36 and 35.5 million years ago, respectively (Bininda-Emonds et al. 1999). The proximity of these divergence times might lead to incomplete lineage sorting (Maddison and Knowle 2006), which is common in highly polymorphic genes (Lu 2001).

**Conclusions**

Our findings highlight the importance of examining gene regions separately in order to obtain a more comprehensive understanding of MHC evolution. While the MHC class II genes (e.g., *DRB* and *DQB*; Sin et al. 2012) in *M. meles* did not exhibit extensive trans-species polymorphism, it nevertheless was observed. The class II genes are therefore likely to be undergoing balancing selection, whereas phylogenetic inconsistency between the α1 and α2/α3 domains of the class I genes indicates that these domains have different evolutionary histories. Concerted evolution provides a more plausible hypothesis than birth-and-death evolution in the α2 and α3 domains, given the intron 2 breakpoint and separate evolutionary history of the α1 domain. Probable recombination has homogenized the α2 and α3 encoding exons, after ancestral gene duplication. As the α2 domain is antigen binding, whereas α3 is not, these domains were subject to different selective pressures leading to the observed difference in the degree of sequence divergence over evolutionary time. Separated from the α2 and α3 encoding exons, the α1 domain is likely to be subject to balancing selection (Nei et al. 1997). The high number of pseudogenes we found is typical in class I loci, which often show a high turnover rate, generating an abundant number of pseudogenes with various degrees of divergence from their functional counterparts (Hughes 1995; Cadavid et al. 1996; Beck et al. 1999; Piontkivska and Nei 2003). Nonfunctional sequences could serve as a reservoir for genetic exchange (Ohta 1991; de Groot et al. 2004), which would be the case here if the recombination event happened after the parental sequence *Meme-MHC* I 1*PS08N* lost its function.

This is the first study to characterize MHC class I genes in a species within the superfamily Musteloidea. Given the polygynandrous mating system of *M. meles*, where high levels of extra-group paternity are observed (Dugdale et al. 2007), further studies of mate choice and MHC would be informative. Examination of the association between MHC genotypes and pathogens could also have significant implications for research into bTB epidemiology in badgers (Allen et al. 2010).

**Acknowledgments**

We thank the Wytham Woods badger team, especially C. Buesching, G. Annavi, P. Nouvellet, and S. Ellwood for assistance with badger trapping. We also thank A. Krupa, D. Dawson, G. Horsburgh, and M.-E. Mannarelli for assistance with laboratory work at the Natural Environment Research Council (NERC) Biomolecular Analysis Facility, Sheffield. Y. W. S. was supported by the Croucher Foundation (Hong Kong) and H. L. D. by the Netherlands Organization for Scientific Research (NWO) and a NERC fellowship (NE/I021748/1). Open access publication charges were paid by the “Incentive Fund Open Access—Publications 2012 BOO,” NWO (grant 036.001.773 awarded to H. L. D.).

**References**

Aldridge, B. M., L. Bowen, B. R. G. Smith A. Antonelis, F. Gulland, and J. L. Stott. 2006. Paucity of class I MHC gene heterogeneity between individuals in the endangered Hawaiian monk seal population. Immunogenetics 58: 203–215.

Allen, A. R., G. Minozzi, E. J. Glass, R. A. Skuce, S.W. J. McDowell, J. A. Woolliams, and S. C. Bishop. 2010. Bovine tuberculosis: the genetic basis of host susceptibility. Proc. R. Soc. Lond. B Biol. Sci. 277:2737–2745.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Anisimova, M., J. P. Bielawski, and Z. Yang. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol. Biol. Evol. 18:1585–1592.

Anisimova, M., J. P. Bielawski, and Z. Yang. 2002. Accuracy and power of Bayes prediction of amino acid sites under positive selection. Mol. Biol. Evol. 19:950–958.

Anwar, A., J. Knaggs, K. M. Service, G. W. McLaren, P. Riordan, C. Newman, R. J. Delahay, C. cheesman, and D. W. Macdonald. 2006. Antibodies to *Toxoplasma gondii* in Eurasian badgers. J. Wildl. Dis. 42:179–181.

Anwar, M., C. Newman, D. Macdonald, M. Woolhouse, and D. Kelly. 2000. Coccidiosis in the European badger (*Meles meles*) from England, an epidemiological study. Parasitology 120:255–260.

Beck, S., D. Geraghty, H. Inoko, and L. Rowen. 1999. Complete sequence and gene map of a human major histocompatibility complex. Nature 401:921–923.

Bergström, T., and U. Gyllensten. 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., and C. Landry. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? J. Evol. Biol. 16:363–377.

Bininda-Emonds, O. R. P., J. L. Gittleman, and A. Purvis. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biol. Rev. 74:143–175.
Bjorkman, P. J., and P. Parham. 1990. Structure, function, and diversity of class I major histocompatibility complex molecules. Annu. Rev. Biochem. 59:253–288.

Bjorkman, P. J., M. A. Saper, B. Samraoui, W. S. Bennett, J. L. Strominger, and D. C. Wiley. 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature 329:518–518.

Bos, D. H., and B. Waldman. 2006. Evolution by recombination and transspecies polymorphism in the MHC class I gene of *Xenopus laevis*. Mol. Biol. Evol. 23:137–143.

Buesching, C. D., J. S. Waterhouse, and D. W. Macdonald. 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. 28:241–256.

Buesching, C. D., M. Heistermann, and D. W. Macdonald. 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.

Burnett, R. C., S. A. DeRose, J. L. Wagner, and R. Store. 1997. Molecular analysis of dog leukocyte antigen class I sequences including three complete genes, two truncated genes and one full length processed gene. Tissue Antigens 49:484–495.

Burri, R., H. N. Hirzel, N. Salamin, A. Roulin, and L. Fumagalli. 2008. Evolutionary patterns of MHC class II B in owls and their implications for the understanding of avian MHC evolution. Mol. Biol. Evol. 25:1180–1191.

Cadavid, L. F., A. L. Hughes, and D. I. Watkins. 1996. MHC class I-processed pseudogenes in New World primates provide evidence for rapid turnover of MHC class I genes. J. Immunol. 157:2403–2409.

de Groot, N., G. G. Doxiadis, N. G. de Groot, N. Otting, C. Heijmans, A. J. M. Rouweler, and R. E. Bontrop. 2004. Genetic makeup of the DR region in thresus macaques: gene content, transcripts, and pseudogenes. J. Immunol. 172:6152–6157.

Delahay, R. J., C. L. Cheeseman, and R. S. Clifton-Hadley. 2001. Wildlife disease reservoirs: the epidemiology of *Mycobacterium bovis* infection in the European badger (*Meles meles*) and other British mammals. Tuberculosis 81:43–49.

Dugdale, H. L., D. W. Macdonald, L. C. Pope, and T. Burke. 2007. Polygynandry, extra-group paternity and multiple-paternity litters in European badger (*Meles meles*) social groups. Mol. Ecol. 16:5294–5306.

Dugdale, H. L., A. Griffiths, and D. W. Macdonald. 2011. Polygynandry and repeated mounting behaviour in European badgers, *Meles meles*. Anim. Behav. 82:1287–1297.

Durairaj, M., R. Sharma, J. C. Varghese, and K. P. Kame. 2003. Requirement for Q226, but not multiple charged residues, in the class I MHC CD loop/D strand for TCR-activated CD8 accessory function. Eur. J. Immunol. 33:676–684.

Fernandez-Soria, V. M., P. Morales, M. J. Castro, B. Suarez, M. J. Recio, M. A. Moreno, E. Paz-Artal, and A. Arnaiz-Villena. 1998. Transcription and weak expression of *HLA-DRB6*: a gene with anomalies in exon I and other regions. Immunogenetics 48:16–21.

Figueroa, E., E. Guntner, and J. Klein. 1988. MHC polymorphism pre-dating speciation. Nature 335:265–267.

Fitch, W. M. 2000. Homology: a personal view on some of the problems. Trends Genet. 16:227–231.

Flynn, J., J. A. Finarelli, S. Zehr, J. Hsu, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Fulton, T. L., and C. Strobeck. 2006. Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene datasets. Mol. Phylogenet. Evol. 41:165–181.

Glaborman, S., and A. Caconn. 2008. Species-specific evolution of class I MHC genes in iguanas (Order: Squamata; Subfamily: Iguaninae). Immunogenetics 60:371–382.

Hess, C. M., and S. V. Edwards. 2002. The evolution of the major histocompatibility complex in birds. Bioscience 52:423–431.

Holmes, N., and P. Parham. 1985. Exon shuffling in vivo can generate novel HLA class I molecules. EMBO J. 4:2849–2854.

Houlden, B. A., W. D. Greville, and W. B. Sherwin. 1996. Evolution of MHC class I loci in marsupials: characterization of sequences from koala (*Phascolarctos cinereus*). Mol. Biol. Evol. 13:1119–1127.

Hughes, A. L. 1995. Origin and evolution of HLA class I pseudogenes. Mol. Biol. Evol. 12:247–258.

Hughes, A. L., and M. Nei. 1989. Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. Mol. Biol. Evol. 6:559–579.

Hughes, A. L., and M. Nei. 1992. Maintenance of MHC polymorphism. Nature 355:402–405.

Hughes, A. L., and M. Yeager. 1998. Natural selection at major histocompatibility complex loci of vertebrates. Annu. Rev. Genet. 32:415–434.

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254–267.

Jakobsen, I. B., S. R. Wilson, and S. Easteal. 1998. Patterns of reticulate evolution for the classical class I and II HLA loci. Immunogenetics 48:312–323.

Jeffrey, K. J., and R. M. Bangham. 2000. Do infectious diseases drive MHC diversity? Microbes Infect. 2:1335–1341.

Joly, E., and V. Rouillon. 2006. The orthology of HLA-E and H2-Qa1 is hidden by their concerted evolution with other MHC class I molecules. Biol. Direct 1:2.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 in H. N. Munro, ed. Mammalian protein metabolism. Academic Press, New York.

Kaufman, J., J. Salomonsen, and M. Flajnik. 1994. Evolutionary patterns of MHC class II B in owls and other birds. Trends Genet. 10:163–171.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.

Klein, J. C. Varghese, and M. A. Nedbal. 2005. Molecular phylogeny of the Carnivorina (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54:317–337.
Kelley, L., W. Walter, and J. Trowsdale. 2005. Comparative genomics of major histocompatibility complexes. Immunogenetics 56:683–695.

Kennedy, L. J., L. Altet, J. M. Angles, A. Barnes, S. D. Carter, O. Francincio, J. A. Gerlach, G. M. Happ, W. E. R. Ollier, A. Polvi, et al. 1999. Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA Nomenclature Committee. Tissue Antigens 54:312–321.

Kennedy, L. J., J. M. Angles, A. Barnes, S. D. Carter, O. Francincio, J. A. Gerlach, G. M. Happ, W. E. R. Ollier, W. Thomson, and J. L. Wagner. 2001. Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: second report of the ISAG DLA Nomenclature Committee. Tissus Anticinstes 58:55–70.

Klein, J. 1987. Origin of major histocompatibility complex polymorphism—the trans-species hypothesis. Hum. Immunol. 19:155–162.

Klein, J., A. Sato, N. Nagl, and C. O’Hugín. 1998. Molecular trans-species polymorphism. Annu. Rev. Ecol. Syst. 29:1–21.

Klein J., A. Sato, and N. Nikolaidis 2007. MHC, TSP, and the origin of species: from immunogenetics to evolutionary genetics. Annu. Rev. Genet. 41:281–304.

Koller, B. H., and H. T. Orr. 1985. Cloning and complete sequence of an HLA-A2 gene: analysis of two HLA-A alleles at the nucleotide level. J. Immunol. 134:2727–2733.

Lizundia, R., C. Newman, C. D. Buesching, D. Ngugi, D. Blake, Y. W. Sin, D. W. Macdonald, A. Wilson, and D. McKeever. 2011. Evidence for a role of the host-specific flea (Paraceras melis) in the transmission of Trypanosoma (Megatrpanyum) pestanaei to the European badger. PLoS ONE 6:e16977.

Lu, Y. 2001. Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of Solanaceae. Heredity 86:195–205.

Macdonald, D. W., and C. Newman. 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, D. W., M. Anwar, C. Newman, R. Woodroffe, and P. J. Johnson. 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, D. W., C. Newman, J. Dean, C. D. Buesching, and P. J. Johnson. 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, D. W., C. Newman, P. M. Nouvellet, and C. D. Buesching. 2009. An analysis of Eurasian badger (Meles meles) population dynamics: implications for regulatory mechanisms. J. Mammal. 90:1392–1403.

Maddison, W. P., and L. L. Knowle. 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55:21–30.

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

Martin, D. P., D. Posada, K. A. Crandall, and C. Williamson. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res. Hum. Retrovir. 21:98–102.

Mathews, F., D. W. Macdonald, G. M. Taylor, M. Gelling, R. A. Norman, P. E. Hones, R. Foster, C. M. Gower, S. Varley, A. Harris, et al. 2006. Bovine tuberculosis (Mycobacterium bovis) in British farmland wildlife: the importance to agriculture. Proc. R. Soc. Lond. B Biol. Sci. 273:357–365.

Mayer, W. E., C. O’huigin, and J. Klein. 1993. Resolution of the HLA-DRB6 puzzle: a case of grafted de novo generated exon on an existing gene. Proc. Natl. Acad. Sci. USA 90:10720–10724.

Mesa, C. M., K. J. Thulien, D. A. Moon, S. M. Veniamin, and K. E. Magor. 2004. The dominant MHC class I gene is adjacent to the polymorphic TAP2 gene in the duck, Anas platyrhynchos. Immunogenetics 56:192–203.

Myers, S., C. Freeman, A. Auton, P. Donnelly, and G. McVean. 2008. A common sequence motif associated with recombination hot spots and genome instability in humans. Nat. Genet. 40:1124–1129.

Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman, T. S. MacFie, G. McVean, and P. Donnelly. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327:876–879.

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

Nei, M., and A. P. Rooney. 2005. Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. 39:121–152.

Nei, M., X. Gu, and T. Sitnikova. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc. Natl. Acad. Sci. USA 94:7799–7806.

Newman, C., D. W. Macdonald, and M. A. Anwar. 2001. Coccidiosis in the European badger, Meles meles in Wytham Woods: infection and consequences for growth and survival. Parasitology 123:133–142.

Newman, C., Y.-B. Zhou, C. D. Buesching, Y. Kaneko, and D. W. Macdonald. 2011. Contrasting sociality in two widespread, generalist, mustelid genera, Meles and Martes. Mamm. Study 36:169–188.

Nouvellet, P., C. D. Buesching, H. L. Dugdale, C. Newman, and D. W. Macdonald. 2010. Mouthing off about developmental
Evolution of MHC Class I Genes in Meles meles

Y. W. Sin et al.

stress: individuality of palate marking in the European badger and its relationship with juvenile parasites. J. Zool. 283:52–62.

Ohta, T. 1991. Role of diversifying selection and gene conversion in evolution of major histocompatibility complex loci. Proc. Natl. Acad. Sci. USA 88:6716–6720.

Owczarzy, R., A. V. Tatarov, Y. Wu, J. A. Manthey, K. A. McQuisten, H. G. Almabrazzi, K. F. Pedersen, Y. Lin, J. Garretson, N. O. McEntaggart, et al. 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. Nucleic Acids Res. 36:W163–W169.

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

Pan, H.-J., Q.-H. Wan, and S.-G. Fang. 2008. Molecular characterization of major histocompatibility complex class I genes from the giant panda (Ailuropoda melanoleuca). Immunogenetics 60:185–193.

Parham, P., and T. Ohta. 1996. Population biology of antigen presentation by MHC class I molecules. Science 272:67–74.

Peace-Brewer, A. L., L. G. Tussey, M. Matsui, G. Li, D. G. Quinn, and J. A. Frelinger. 1996. A point mutation in HLA-A*0201 results in failure to bind the TAP complex and to present virus-derived peptides to CTL. Immunity 4:505–514.

Penn, D. J., and W. K. Potts. 1999. The evolution of mating preferences and major histocompatibility complex genes. Am. Nat. 153:145–164.

Penn, D. J., K. Damjanovich, and W. K. Potts. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. Proc. Natl. Acad. Sci. USA 99:11260–11264.

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

Piontkivska, H., and M. Nei. 2003. Birth-and-death evolution in primate MHC class I genes: divergence time estimates. Mol. Biol. Evol. 20:601–609.

Rada, C., Lorenzi, R., S. J. Powis, J. V. D. Bogaerde, P. Parham, and J. C. Howard. 1990. Concerted evolution of class I genes in the major histocompatibility complex of murine rodents. Proc. Natl. Acad. Sci. USA 87:2167–2171.

Richman, A. 2000. Evolution of balanced genetic polymorphism. Mol. Ecol. 9:1953–1963.

Richman, A., L. Herrera, D. Nash, and M. Schierup. 2003. Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in Peromyscus maniculatus. Genet. Res. 82:89–99.

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

Ronquist, F., J. P. Huelsenbeck, and P. van-der-Mark. 2005. MrBayes v. 3.1 Manual. Available at: http://mrbayes.sourceforge.net/manual.php

Rosalino, L. M., D. W. Macdonald, and M. Santos-Reis. 2004. Spatial structure and land-cover use in a low-density Mediterranean population of Eurasian badgers. Can. J. Zool. 82:1493–1502.

Rosalino, L. M., J. Torres, and M. Santos-Reis. 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.

Swain, S. L.. 1983. T cell subsets and the recognition of MHC class II antigens. J. Immunol. 130:2977–2981.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596–1599.

Thom, M. D., D. D. P. Johnson, and D. W. Macdonald. 2004. The evolution and maintenance of delayed implantation in the Mustelidae (Mammalia:Carnivora). Evolution 58: 175–183.

Vogel, T. U., D. T. Evans, J. A. Urvater, D. H. O’Connor, A. L. Hughes, and D. I. Watkins. 1999. Major histocompatibility complex class I genes in primates: co-evolution with pathogens. Immunol. Rev. 167:327–337.

Wagner, J. L.. 2003. Molecular organization of the canine major histocompatibility complex. J. Hered. 94:23–26.

Wagner, J. L., R. C. Burnett, and R. Storb. 1999. Organization of the canine major histocompatibility complex: current perspectives. J. Hered. 90:35–38.
Evolution of MHC Class I Genes in *Meles meles*

Y. W. Sin et al.

Yamaguchi, N., H. L. Dugdale, and D. W. Macdonald. 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.. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yang, Z., R. N. Nielsen Goldman, and A.-M. K. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155:431–449.

Yang, Z., W. S. W. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107–1118.

Yuhki, N., T. Beck, R. Stephens, B. Neelam, and S. J. O’Brien. 2007. Comparative genomic structure of human, dog, and cat MHC: HLA, DLA, and FLA. J. Hered. 98:390–399.

Zhong, J. F., J. T. Harvey, and J. T. Boothby. 1998. Characterization of a harbor seal class I major histocompatibility complex cDNA clone. Immunogenetics 48:422–424.

Supporting Information

Additional Supporting Information may be found online on Wiley Online Library.

Table S1. Different *Meles meles* MHC class I alleles could be obtaining by PCR amplification using the primers marked with ticks.

Please note: Wiley-Blackwell is 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.