MHC class II-assortative mate choice in European badgers (Meles meles)

YUNG WA SIN,*†‡ GEETHA ANNAVI,*†§ CHRIS NEWMAN,* CHRISTINA BUESCHING,* TERRY BURKE,† DAVID W. MACDONALD* and HANNAH L. DUGDALE†¶

*Wildlife Conservation Research Unit (WildCRU), Department of Zoology, Recanati-Kaplan Centre, University of Oxford, Tubney House, Abingdon Road, Abingdon, Oxfordshire OX13 5QL, UK, †NERC Biomolecular Analysis Facility, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, South Yorkshire S10 2TN, UK, ‡Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA, §Biology Department, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Darul Ehsan, Malaysia, ¶Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, 9700CC Groningen, the Netherlands

Abstract

The major histocompatibility complex (MHC) plays a crucial role in the immune system, and in some species, it is a target by which individuals choose mates to optimize the fitness of their offspring, potentially mediated by olfactory cues. Under the genetic compatibility hypothesis, individuals are predicted to choose mates with compatible MHC alleles, to increase the fitness of their offspring. Studies of MHC-based mate choice in wild mammals are under-represented currently, and few investigate more than one class of MHC genes. We investigated mate choice based on the compatibility of MHC class I and II genes in a wild population of European badgers (Meles meles). We also investigated mate choice based on microsatellite-derived pairwise relatedness, to attempt to distinguish MHC-specific effects from genomewide effects. We found MHC-assortative mating, based on MHC class II, but not class I genes. Parent pairs had smaller MHC class II DRB amino acid distances and smaller functional distances than expected from random pairings. When we separated the analyses into within-group and neighbouring-group parent pairs, only neighbouring-group pairs showed MHC-assortative mating, due to similarity at MHC class II loci. Our randomizations showed no evidence of genomewide-based inbreeding, based on 35 microsatellite loci; MHC class II similarity was therefore the apparent target of mate choice. We propose that MHC-assortative mate choice may be a local adaptation to endemic pathogens, and this assortative mate choice may have contributed to the low MHC genetic diversity in this population.

Keywords: Genetic compatibility hypothesis, major histocompatibility complex, MHC-assortative mating, pre- and postcopulatory mate choice, sexual selection

Received 14 January 2014; revision received 22 April 2015; accepted 23 April 2015

Introduction

A central question in the study of sexual selection is what drives the choice of a particular mate over other mates. Indirect genetic benefits have been proposed to underpin mate choice, especially when direct benefits have not been detected (Zelano & Edwards 2002). In theory, females [which are often the choosier sex (Tregenza & Wedell 2000)] are predicted to select males based on ‘good genes’ or genetic compatibility (Trivers 1972; Neff & Pitcher 2005; Kempenaers 2007). Hamilton and Zuk’s ‘good genes’ theory proposed that females select males carrying genes for resistance against the currently prevalent pathogens, to ensure that offspring will carry these good genes (Hamilton & Zuk 1982).

Correspondence: David W. Macdonald, Fax: +44(0)1865611101; E-mail: david.macdonald@zoo.ox.ac.uk

© 2015 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
The good genes hypothesis posits that all breeding individuals in the population will show consistent patterns of mate choice with respect to pathogen prevalence, irrespective of their own genotypes (Neff & Pitcher 2005). Genetic compatibility, by contrast, requires individuals to assess the genotype of potential mates by self-reference to their own genotype (Puurtinen et al. 2005).

In vertebrates, some of the assumed polymorphic genes for resistance in the Hamilton-Zuk model are MHC genes (Milinski 2014). The MHC is a diverse gene family that plays a crucial role in the adaptive immune system, as it encodes cell surface glycoproteins that bind and present antigens to T cells and trigger an immune cascade (Swain 1983). These MHC genes are associated with mate choice (Penn & Potts 1999); for example, MHC-based mate choice was first shown in mice, which preferred the odour of potential mates that had the most dissimilar MHC genes to their own (Yamazaki et al. 1976). Mating with dissimilar mates will, on average, increase the heterozygosity of offspring, and because pathogen recognition is mediated by the sequence of individual MHC alleles, MHC heterozygotes might have an advantage of being susceptible to fewer pathogens than homozygotes (McClelland et al. 2003). The MHC might therefore provide the basis on which individuals discriminate mating partners, to increase the fitness of their offspring (Penn & Potts 1999; Penn et al. 2002; Milinski 2006). While the arms race between pathogens and hosts maintains extreme MHC diversity (Jeffrey & Bangham 2000; Pierrney & Oliver 2006), MHC-related reproductive mechanisms, such as mating preferences, selective fertilization and abortion, also generate MHC diversity (Ziegler et al. 2005; Levlie et al. 2013). MHC genotypes can be detected through olfactory cues (Penn 2002; Leinders-Zufall et al. 2004; Milinski et al. 2005), or through other traits (e.g. ornamentation; vonSchantz et al. 1997; Dunn et al. 2012); thus, the MHC is a highly plausible target for mate choice through indirect genetic benefits.

Individuals could select MHC-similar or MHC-dissimilar mates, but availability of mating partners might also play a role, particularly in group-living species. A preference for MHC-similar males has been reported in a handful of species, potentially due to adaptation to local pathogen pressures (e.g. Yamazaki et al. 1978; Roberts et al. 2005; Bonneauad et al. 2006; Bos et al. 2009; Bollmer et al. 2012). MHC-disassortative rather than MHC-assortative mating is, however, more often detected (Kamiya et al. 2014). Choosing mates with dissimilar MHC can maximize MHC diversity (Ejsmond et al. 2014) and might consequently enhance disease resistance among offspring by the ability to recognize more pathogenic antigens (Doherty & Zinkernagel 1975; Penn et al. 2002). MHC-disassortative mating occurs in a variety of animals (e.g. Landry et al. 2001; Gillingham et al. 2009; Miller et al. 2009; Juola & Dearborn 2012). A recent meta-analysis of 116 effect sizes from 48 studies showed MHC-dissimilar mate choice only in species with multiple MHC loci examined (Kamiya et al. 2014). Mating with a dissimilar partner, with highly divergent MHC genes, might be risky, however, because disruption of co-adapted genes may reduce fitness (Kaufman 1999; Hendry et al. 2000; Neff 2004). Mate choice can, however, produce offspring that have optimal rather than maximal diversity at the MHC loci (Wegner et al. 2003; Milinski 2006). A meta-analysis suggested possible support for the optimizing hypothesis, but pointed out that this should be interpreted with care as it was based on 4 effect sizes from 4 species (Kamiya et al. 2014). The direction of MHC-based mate choice may be context dependent. For example, MHC-based mate choice might play a greater role under circumstances where an extra-pair or extra-group mate has been selected, compared with within-pair/group mate choice. If social mate choice is limited, individuals may settle for a social mate with less optimal MHC compatibility, and in these circumstances, they may be more likely to engage in extra-pair/group matings for MHC-compatible mates (Richardson et al. 2005; Schwensow et al. 2008).

The European badger (Meles meles) is well suited for investigating MHC-based mate choice in the wild, given their mating system, life history characteristics and reproductive biology. In high-density populations, M. meles is group living (Newman et al. 2011), with frequent contacts between neighbouring groups (Macdonald et al. 2008), and has a polygynandrous mating system (i.e. they do not have one exclusive social mate; Dugdale et al. 2007, 2011) with low fecundity (i.e. 1-2 cubs once a year; Macdonald & Newman 2002; Carpenter et al. 2005; Dugdale et al. 2007). Extra-group paternity accounts for >40% of offspring (Carpenter et al. 2005; Dugdale et al. 2007) with no subsequent paternal care (Dugdale et al. 2010). Although European badgers can conceive throughout the year, delayed implantation uncouples mating and parturition (Thom et al. 2004) and they give birth fairly synchronously around February (Yamaguchi et al. 2006). As female badgers are capable of superfetation and embryo re-absorption (Yamaguchi et al. 2006), this extends the opportunity for females to select the most suitable mates, through pre- and postcopulatory mate choice (Andersson & Simmons 2006). Furthermore, we have found no evidence for clear mounting hierarchies in these badgers, that is male mounting frequency is not related to dominance rank; however, male mounting frequency does not correlate with paternity success (Dugdale et al. 2011). Promiscuous mounting may therefore promote sperm...
competition and selection for genetic compatibility through postcopulatory mate choice.

In the European badger, among the four MHC class II genes, the DRB gene is the most variable, with four putatively functional sequences identified (Sin et al. 2012c). Seven putatively functional MHC class I sequences have also been identified in the same study population (Sin et al. 2012b). In contrast to the highly diverse MHC genes in many other species (e.g. Robinson et al. 2003), the low variability of the DRB and class I genes allows the number of loci and haplotypes to be inferred, with at least two DRB loci and two class I loci (Sin et al. 2012b,c). This low MHC diversity therefore provides a good system with which to investigate MHC-based mate choice, in the broad sense of both pre- and postcopulatory mechanisms, because there are no complications that arise from analysing high allelic diversity (Richardson et al. 2005).

To test the genetic compatibility hypothesis, we examined whether badgers selected mates according to their MHC similarity (class II DRB and class I genes), taking into account the number of shared MHC alleles, the functional differences between alleles, and social-group membership. Given that badgers are group-living in high-density populations, and more related to within-group than to neighbouring-group members (Dugdale et al. 2008), MHC similarity is likely to vary with respect to group membership. Additionally, the decision to mate outside of the group may be context dependent, for example in relation to the compatibility of potential within-group mates. We therefore also analysed within-group and neighbouring-group parent pairs separately, in addition to the analyses that combined both within- and neighbouring-group parent pairs. To attempt to distinguish MHC-specific effects from genomewide effects, we then compared MHC similarity with that of potentially neutral microsatellite markers. Mating associated with MHC similarity but not microsatellite similarity would indicate that MHC similarity is the more likely target of mate choice than microsatellite-wide similarity.

Materials and methods

Study site and sample collection

We studied a high-density (44 badgers/km²; Macdonald & Newman 2002; Macdonald et al. 2009) population of badgers in Wytham Woods, a 4-km² mixed coniferous–deciduous woodland in Oxfordshire, UK (51°46′26″N, 1°19′19″W). Trapping events have been typically undertaken three to four times a year since 1987 (for detailed methods, see Macdonald & Newman 2002). Upon first capture, each badger was tattooed with a unique number on their left inguinal region for permanent individual identification. The sex, age class (cub or adult, based on body size) and location (social group) of each badger were recorded. DNA samples were collected: guard hairs were plucked and stored in 80% ethanol, and approximately 3 mL of blood was taken by jugular venipuncture, collected in a vacutainer containing EDTA and stored at –20°C, until DNA was extracted.

Social group ranges were established using bait-marking techniques approximately every 2 years (Macdonald et al. 2008). The mean number of badger social groups in this population, 1987–2005, was 19 ± 2 SE (range = 14–26; Dugdale et al. 2008). These badgers typically exhibit high group fidelity; only 19% of individuals dispersed permanently to other social groups, although temporary intergroup movements do occur (16%, Macdonald et al. 2008). We defined the social group of residence of each individual per year, following the criteria of Macdonald et al. (2008): (a) badgers were assigned to the social group in which they were first trapped as a cub (i.e. natal group) unless they satisfied our dispersal rule (c); (b) badgers first caught as adults were assigned to the social group that they were trapped in most frequently unless dispersal events were recorded; and (c) badgers were recorded as dispersing if the 2 most recent captures, as well as at least 1 of 2 captures before that, were made in a different group to their resident social group.

DNA extraction and MHC gene amplification

Genomic DNA was isolated using the GFX Genomic DNA Purification Kit (Amersham Biosciences, Little Chalfont, UK), or from a minimum of 20 hairs with visible follicles using a Chelex protocol (Walsh et al. 1991). We used published primers to amplify the regions that encode the antigen-binding domain in class II DRB [Meme-DRBex2, for the exon 2 (the forward primer spanned the boundary of intron 1 and exon 2); Sin et al. 2012c] and class I (Meme-MHICex3, for the exon 3; Sin et al. 2012b) genes. Using these primers on 10–30 ng of gDNA, PCR amplification was performed in a 10-μL reaction mix that also contained: 0.5 μM of each primer, 200 μM of each dNTP, 1× PCR buffer (containing MgCl2; Qiagen, Hilden, Germany) and 1 unit 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 s, 59°C for 30 s and 72°C for 60 s, ending with an extension step at 72°C for 10 min.

MHC genotyping

We used reference strand-mediated conformation analysis (RSCA), which can detect genetic variants that differ

© 2015 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd.
at just a single nucleotide (Argüello et al. 1998a,b; Kennedy et al. 2005; Lenz et al. 2009). Fluorescent-labelled reference strands (FLRs) were generated using alleles from closely related species (American badger *Taxidea taxus*, polecat *Mustela putorius*, stoat *Mustela erminea* and mink *Neovison vison*). FLRs were generated by PCR using cloned alleles as templates and a 5'-FAM-labelled primer. Cloning and sequencing details are described in Sin et al. (2012b,c). The same PCR protocol, detailed in the MHC gene amplification section, was used to generate the FLRs, except that the primer proportion was altered to 0.5 µM FAM-labelled primer for Meme-DRBex2R and Meme-MHClex3F, and 0.1 µM unlabelled primer for Meme-DRBex2F and Meme-MHClex3R. Nine FLRs were evaluated for each of the class II DRB and the class I genes, to determine a subset that could best resolve all sequences in the control samples (13 and 12 cloned sequences for MHC class II and class I, respectively). A final set of three FLRs derived from the American badger, polecat and stoat were used for the DRB genes (GenBank Accession no.: KM371113–KM371115), and three FLRs derived from the American badger and mink for the class I genes (KM371116–KM371118).

All the resulting FLRs were diluted 1:5 in ddH2O before use in the hybridization reactions. To form heteroduplexes, 2 µL diluted FLR and 2 µL PCR product (detailed in the MHC gene amplification section) were mixed and incubated in a thermal cycler at 95°C for 10 min, cooled down to 55°C at 1°C/s, hybridized at 55°C for 20 min, cooled to 15°C for 1 min and put on ice for 30 min. The plate was then stored at 4°C. Subsequently, 3 µL of hybridization product was mixed with 7.82 µL water and 0.18 µL Genescan Rox-500 size standard (Applied Biosystems (ABI), Foster City, USA) in a 96-well plate. The samples were then run on an ABI 3100 Genetic Analyzer, using 4% Genescan non-denaturing polymer (ABI), and visualized using matrix Dye set D. The running protocol used an injection voltage of 8 kV, injection time 15 s, run voltage of 15 kV and run temperature of 18°C. The heteroduplex peaks were identified using GENEMAPPER 3.7 (ABI), and their motilities were estimated relative to the ROX size standard. Control alleles from cloned plasmids were included in each run to control for variation between runs. Peaks with the same motility across different FLRs were designated as identical putative alleles. All seven class I alleles and four class II DRB alleles identified were sequenced (Sin et al. 2012b,c).

**Microsatellite typing and parentage analyses**

Individuals were genotyped using 35 microsatellite loci (Annavi et al. 2011). The loci were in Hardy–Weinberg and linkage equilibrium (see Dugdale et al. 2007; Annavi et al. 2014a). Details of candidate parent rules, genotyping procedures and parentage analysis are described elsewhere (Annavi et al. 2014a). Briefly, parentage and sibships were assigned with at least 80% confidence (accounting for genotyping error and unsampled individuals; see Annavi et al. 2014a) using MasterBayes 2.47 (Hadfield et al. 2006) and Colony 2.0 (Wang & Santure 2009), respectively. Candidate fathers included all sexually mature males (i.e. older than 1 year) present in the population in the year before the cub was born (due to delayed implantation; Yamaguchi et al. 2006). Candidate mothers included all females aged two or more resident in the cub’s natal group in the year the cub was born. We defined a parent pair as a male badger and female badger that were assigned parentage of the same cub with at least 80% confidence.

As badgers may be present in the population but not caught in the year of mating (Dugdale et al. 2007), we included individuals as candidate parents for: (a) 2 years beyond their last capture, unless death was confirmed; and (b) the period between their consecutive captures. We analysed mate choice in the 6 years in which the most parent pairs were assigned. See Table S1 (Supporting information) for numbers of assigned parent pairs and candidate fathers in the different analyses. Assigned parent pairs that were resident in the same social group in the year in which the cubs were conceived were categorized as within-group pairs, whereas parents that were from neighbouring social groups were categorized as neighbouring-group pairs.

**Data analysis**

**MHC similarity parameters.** To examine whether assigned parent pairs had more (or less) similar MHC genes than expected under random mating, we examined MHC compatibility using both the extent of allele sharing and the magnitude of the functional differences between genotypes. The allele sharing value indicates the number of class I or II alleles shared between assigned parent pairs. To incorporate the functional similarity between class I and II alleles in the analysis, we also calculated the amino acid distance (proportion of amino acid sites that differed) from pairwise combinations of alleles in assigned parent pairs (Landry et al. 2001). We used the average amino acid distance (Landry et al. 2001; Forsberg et al. 2007; Miller et al. 2009) because individuals might carry different numbers of alleles, and hence, the number of pairwise comparisons could be different between parent pairs. We calculated the amino acid distance for both the entire exon sequences [exon 2 for DRB (285 bp) and exon
Randomization tests. Randomization tests (Landry et al. 2001) enabled us to compare the mean allele sharing value, amino acid distance and functional distance for assigned parent pairs with the frequency distribution of mean values generated from 1000 simulations of the same number of randomly selected parent pairs. Specifically, we disassociated all parent pairs assigned in the parentage analyses and selected parent pairs at random (detailed in the next paragraph) with replacement. We then calculated mean values 1000 times to derive a distribution under random mating. We calculated P-values as the proportion of the total number of iterations greater or smaller than the observed mean (Fisher 1935). We applied a 2-tailed test with $\alpha = 0.05$; any values that fell outside of the 97.5-2.5% confidence interval (CI) were significant. We ran the randomizations separately for MHC class I and class II data sets. All statistical analyses were performed in R 2.15.0 (R Core Development Team 2011).

Forty-three per cent of the paternities were assigned to extra-group males (6 year data set: 98/230 assignments), of which 91% were neighbouring-group males (89/98). We therefore restricted our analyses of extra-group mating to assigned parent pairs from neighbouring-groups only. Likewise, during randomizations, random extra-group parent pairs were only selected from neighbouring-groups. We first compared the MHC similarity parameters of assigned parent pairs [by analysing within- and neighbouring-group parent pairs together, which produced 96% (221/230) of the cubs] with randomly paired assigned mothers and candidate fathers (irrespective of whether the candidate fathers were assigned paternity). This randomization included all within- and neighbouring-group candidate fathers and all candidate mothers that were assigned maternity. This randomization was weighted based on the probability of candidate mothers pairing with a within-group or neighbouring-group mate, according to the proportion of within- versus neighbouring-group paternity in a particular data set for each year (Table S1 in Supporting information). This was required to simulate random mating given that there is invariably a much greater availability of neighbouring males than within-group males; the rate of neighbouring-group mating would be inflated erroneously if the simulations were conducted without weighting.

In addition, badgers are more related to within-group than to neighbouring-group members (Dugdale et al. 2008) and within-group relatedness of assigned mothers and candidate fathers has a negative quadratic effect on the rate of extra-group paternity (Annavi et al. 2014b). We therefore examined the MHC similarity of assigned parent pairs and of randomly paired fathers and mothers in another two sets of randomizations according to social-group membership. First, we compared the MHC similarity parameters of assigned within-group parent pairs with within-group parents that were paired randomly. Our randomizations included all within-group candidate fathers (i.e. including males that were not assigned paternity, in addition to the assigned fathers). Second, we compared the MHC similarity of assigned neighbouring-group parents with random neighbouring-group pairs, where a random mate was selected for each assigned mother from all of their neighbouring candidate fathers.

These randomizations enabled us to compare the MHC similarity of each parent pair to that of the majority of males that a particular female could have mated with, according to social-group membership. Specifically, when females choose a (i) within- or neighbouring-group; (ii) within-group or (iii) neighbouring-group mate, we investigated whether they selected this mate based on MHC compatibility, given the (i) within- and neighbouring-group; (ii) within-group and (iii) neighbouring-group mates they could have chosen, respectively.

To test whether any observed MHC-based mating was a consequence of selection on wider genetic diversity, we also calculated the genetic similarity of assigned parent pairs based on 35 microsatellite loci (Annavi et al. 2011) using a pairwise relatedness measure $R$ (Queller & Goodnight 1989). The mean pairwise relatedness of assigned parent pairs was compared with 1000 simulations of randomly selected parent pairs.

© 2015 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd.
We ran the randomizations of relatedness separately for MHC class I and class II data sets, using the same numbers of parent pairs as for the randomizations of MHC similarity parameters.

Combining probabilities. We assessed the overall significance of detecting larger or smaller mean values of the five genetic similarity measures (allele sharing, amino acid distance of the entire exon sequences, amino acid distance of the ABS, functional distance and relatedness), across years, using Fisher’s method of combining probabilities for independent tests (Sokal & Rohlf 1994), by combining either left- or right-sided \( P \)-values \(( \alpha = 0.025 \)\). The parent pairs, analysed over 6 years, can be considered different reproductive events, because females can prevent males from mounting them \((59/257 = 23\% \text{ of mounting events were failed mountings}; \text{Dugdale et al. 2011})\), so they are not coerced into reproducing with the same male(s) each year (the behavioural mating system is polygynandrous, Dugdale et al. 2011). We analysed 116 females that were assigned maternity; 16 of these females were assigned the same mate in >1 year. We therefore excluded these repeated occurrences of parent pairs by including these parent pairs in only 1 year, which we selected at random, to meet the assumption of combining \( P \)-values from independent tests. Results are considered significant after false discovery rate control (Benjamini et al. 2001) for multiple tests \((n = 10, \alpha = 0.025, \text{adjusted } P \text{-value} = 0.0025–0.025)\) using the same data.

Correlations between variables. We tested for associations between different parameters [i.e. allele sharing, relatedness, pairing type (within- or neighbouring-group) and year] and amino acid distance of assigned within-group and neighbouring-group parent pairs, to strengthen our hypothesis that MHC amino acid distance was the actual target of selection. We ran general linear mixed models, with four fixed main effects: two continuous (allele sharing and relatedness) and two categorical (pairing type and year). We modelled first-order interactions between pairing type, and allele sharing, relatedness and year. We did not include amino acid distance of the ABS, as it was highly correlated with amino acid distance of all sites \([\text{correlation coefficient} = 1.0 \text{ (class II) and } 0.97 \text{ (class I)}]\). We controlled for females that were assigned maternity in more than 1 year, and social groups with more than one assigned mother, by including individual identity and social group residency of assigned mothers as random effects. Model selection was based on Akaike’s information criterion corrected for sample size \((\text{AICc, Akaike 1973})\). Models that better fit the data produce lower AICc values. Multimodel inference (Burnham & Anderson 2002) was performed for models with \(\Delta\text{AICc} < 7\) (Burnham et al. 2011). We report model averaged parameter estimates and their 95% confidence intervals (Anderson 2008).

Results

MHC class II DRB genotypes were determined for 366 individuals \((201 \text{ assigned parent pairs and their cubs})\) and MHC class I genotypes for 356 individuals \((186 \text{ assigned parent pairs and their cubs})\). Candidate males that were not assigned parentage were also genotyped \((250 \text{ and } 263 \text{ individuals for DRB and class I genes, respectively})\) for the randomization analyses. We identified six class II DRB and 13 class I genotypes, from combinations of four class II DRB and seven class I putatively functional sequences \([\text{for sequence information, see Sin et al. 2012a–c}]\). We inferred three DRB and five class I haplotypes \((\text{Fig. 1})\) from parentage assignments, based on the putatively functional sequences and assumption of Mendelian inheritance. These MHC genotyping results fit the parentage data, assuming Mendelian inheritance, suggesting we did not miss any MHC alleles stochastically. However, we could not rule out the possibility that we might not have amplified all MHC loci.

![Diagram](image)

**Fig. 1** The three MHC class II DRB haplotypes and five MHC class I haplotypes in the study population of European badgers \((\text{Meles meles})\), based on putatively functional sequences. It is uncertain whether the MHC class I haplotype 5 comprises \(\text{Meme-MHCI}^*05\) or not. This is because \(\text{Meme-MHCI}^*05\) was present in all examined individuals and the frequency of haplotype 5 was low in the population \((0.9\% \text{ in } 1117 \text{ individuals}; \text{Sin et al. 2014})\), with no homozygote identified.
(a) Combined data set (within- and neighbouring-group parents)

Assigned within-group and neighbouring-group parent pairs had smaller amino acid distances (both at all amino acid sites and the ABS only) and functional distance than random parent pairs, at the MHC class II DRB gene but not at the class I gene, over all years by combining probabilities from 6 years (Table 1). Year-to-year variation in these patterns occurred; for example, in four of 6 years, the assigned within-group and neighbouring-group parent pairs had smaller DRB amino acid distances, compared to random parent pairs, but it did not differ in the other 2 years (Fig. 2; Fig. S1 in Supporting information). MHC allele sharing did not differ between assigned and random parent pairs over all 6 years (Table 1; Fig. S1 in Supporting information). Assigned within-group and neighbouring-group parent pairs were less related than random within-group and neighbouring-group parent pairs overall (Table 1; Fig. S1).

(b) Within-group parents

Over all 6 years, assigned within-group parent pairs shared similar numbers of class I and class II alleles, similar amino acid distances and similar functional distances as randomized within-group pairs (Table 1; Figs S2, S3 in Supporting information). Although there was year-to-year variation (Fig. S2 in Supporting information) overall the results showed that assigned within-group pairs were less related than random within-group pairs (Table 1).

(c) Neighbouring-group parents

Assigned neighbouring-group parent pairs had smaller amino acid distances and functional distance at MHC class II genes than random neighbouring-group pairs, but they had similar amino acid distances at MHC class I genes, similar numbers of class I and II alleles, and similar relatedness to random neighbouring-group parent pairs (Table 1; Figs S2, S3 in Supporting information). Year-to-year variation was detected (Fig. S2 in Supporting information).

Correlations between variables

Greater MHC class I amino acid distances correlated with fewer shared alleles and lower relatedness of assigned within-group and neighbouring-group parent pairs (Fig. 3a). The MHC class I amino acid distances of assigned within-group and neighbouring-group parent pairs was similar in all years, except in 2010 when it was greater than in 1993 (Fig. 3a).

No significant association was found between the amino acid distances of the MHC class II DRB gene (all sites) and allele sharing (Fig. 3b). MHC class II DRB amino acid distances increased with relatedness for assigned within-group parent pairs, but the distances did not vary with relatedness for assigned neighbouring-group parent pairs (Fig. S4 in Supporting information). Amino acid distance in 2008 and 2010 was lower than in 1993 for within-group parent pairs compared to neighbouring-group parent pairs (Fig. 3b).

Discussion

We demonstrate MHC-assortative mate choice in badgers, based on the amino acid distance and functional distance of class II genes of within-group and neighbouring-group parent pairs. In addition, neighbouring-group parent pairs (but not within-group parent pairs) had MHC class II DRB alleles that were smaller in


<table>
<thead>
<tr>
<th>Parent pairs</th>
<th>MHC gene</th>
<th>Allele sharing, $\chi^2_{12}$</th>
<th>Smaller a.a. distance (all), $\chi^2_{12}$</th>
<th>Smaller a.a. distance (ABS), $\chi^2_{12}$</th>
<th>Smaller functional distance, $\chi^2_{12}$</th>
<th>Lower R, $\chi^2_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class I</td>
<td>18.62 (higher)</td>
<td>25.42</td>
<td>24.55</td>
<td>16.43</td>
<td>29.78*</td>
<td></td>
</tr>
<tr>
<td>Class II</td>
<td>12.87 (higher)</td>
<td>40.90***</td>
<td>40.43***</td>
<td>40.64***</td>
<td>28.36*</td>
<td></td>
</tr>
<tr>
<td>Within-group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class I</td>
<td>20.54 (lower)</td>
<td>21.12</td>
<td>20.97</td>
<td>26.86</td>
<td>44.81***</td>
<td></td>
</tr>
<tr>
<td>Class II</td>
<td>13.04 (lower)</td>
<td>21.90</td>
<td>21.35</td>
<td>21.38</td>
<td>42.12***</td>
<td></td>
</tr>
<tr>
<td>Neighbouring-group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class I</td>
<td>16.27 (higher)</td>
<td>20.70</td>
<td>20.79</td>
<td>11.49</td>
<td>10.72</td>
<td></td>
</tr>
<tr>
<td>Class II</td>
<td>12.23 (higher)</td>
<td>32.32**</td>
<td>39.00***</td>
<td>36.48***</td>
<td>9.86</td>
<td></td>
</tr>
</tbody>
</table>

P-values are combined for a higher (or lower) number of shared alleles, smaller amino acid (a.a.) distance (all sites and ABS only), smaller functional distance and lower Queller and Goodnight’s index of pairwise relatedness (R) of assigned parent pairs compared to randomly selected parent pairs. Significant results, after false discovery rate control (number of tests = 10, α = 0.025, adjusted P-value = 0.0025–0.025), are shown in bold (*P < 0.005, **P < 0.002, ***P < 0.001). See supporting information materials for yearly results.
amino acid distance and functional distance than expected from random neighbouring-group parents. In contrast to mate choice for similar MHC genes, parent pairs (both combined data set and within-group) exhibited disassortative mating at microsatellite loci. The different direction of this relationship with the microsatellite versus MHC class II similarity indicated that the smaller MHC amino acid distance of assigned parent pairs versus random pairs was not due to genomewide inbreeding. In combination, these MHC and microsatellite results suggest that MHC functional similarity is the actual target of female choice. Our results indicate that mate choice for MHC similarity might benefit offspring by having less divergent MHC alleles and maintaining co-adapted gene complexes (Kaufman 1999; Hendry et al. 2000; Neff 2004). Because cubs could only be trapped after they have emerged from below the ground when around 12 weeks old (Macdonald & Newman 2002), selective MHC-based mortality before emergence could not be ruled out, although this would not explain the within-/neighbouring-group differences we observed.

We found evidence that badgers selected neighbouring-group mates with reference to MHC class II gene similarity, but this was not observed in within-group pairs. Indeed, the structuring of the badger population into social groups might affect MHC-based mate choice (Huchard et al. 2010), because within-group relatedness was high in this population (Dugdale et al. 2008). Inbreeding avoidance was indicated by analyses based on randomization of within-group and neighbouring-group candidate mate partners, and the significance was much greater when only within-group parents were considered. Potentially, inbreeding avoidance occurred only when candidate mates were more highly related; which was the case in within-group compared to neighbouring-group pairs (Table 1; Fig. S2 in Supporting information; Dugdale et al. 2008). In addition, high relatedness between within-group assigned mothers and candidate fathers was correlated with a higher rate of extra-group paternity (Annavi et al. 2014b). Therefore, the apparent lack of MHC-assortative mate choice in within-group pairs could result from the high level of within-group relatedness and the potential inbreeding avoidance, which could override MHC-assortative mating. Alternatively, the lack of within-group MHC-based mate choice could be an artefact of data selection, if badgers that had fewer MHC-compatible within-group mates only mated outside of their group. However, when all parent pairs were considered in one analysis, MHC-assortative mate choice was detected.

The MHC-assortative mate choice that we demonstrated has been reported by only a few other studies (Yamazaki et al. 1978; Jordan & Bruford 1998; Roberts et al. 2005; Sommer 2005; Bonneaua et al. 2006; Bos et al. 2004 2005 2008 2009 2010 N indicates the number of assigned parent pairs. Significant p-values are shown in bold.

Fig. 2 Mean MHC class II DRB amino acid distances (ABS only) of assigned within-group and neighbouring-group parent pairs (solid line), compared to random within-group and neighbouring-group parent pairs. The frequency distributions (grey bars) are mean values generated from 1000 simulations of random pairings between all assigned mothers and all candidate fathers from both within the group and neighbouring groups. The two-tailed 95% confidence intervals (dashed lines) indicate cut-offs for significant departures from random mating. Randomizations were performed separately for each year: 1993, 2004, 2005, 2008, 2009 and 2010. N indicates the number of assigned parent pairs. Significant p-values are shown in bold.
inference after model selection for models with type and year, respectively. For MHC class I data set, the interaction between pairing type and year was not included in multimodel for all models (95% confidence intervals do not overlap zero have their 95% confidence intervals (CI) and relative importance after model averaging (models with $\Delta$AICc < 7). Estimates where the 95% confidence intervals do not overlap zero have their $y$-axis label in bold. Relative importance (RI) is the sum of Akaike weights for all models ($\Delta$AICc < 7) including the parameter. Neighbouring-group pairs and 1993 were the reference categories for pairing type and year, respectively. For MHC class I data set, the interaction between pairing type and year was not included in multimodel inference after model selection for models with $\Delta$AICc < 7.

Fig. 3 The effect of allele sharing, Queller and Goodnight’s index of pairwise relatedness, pairing type (within-group [WG] or neighbouring-group pair), year and parameter interactions on the amino acid distances (all amino acid sites) of (a) MHC class I and (b) MHC class II DRB genes of assigned within-group and neighbouring-group parent pairs. Parameter estimates are presented with their 95% confidence intervals (CD) and relative importance after model averaging (models with $\Delta$AICc < 7). Estimates where the 95% confidence intervals do not overlap zero have their $y$-axis label in bold. Relative importance (RI) is the sum of Akaike weights for all models ($\Delta$AICc < 7) including the parameter. Neighbouring-group pairs and 1993 were the reference categories for pairing type and year, respectively. For MHC class I data set, the interaction between pairing type and year was not included in multimodel inference after model selection for models with $\Delta$AICc < 7.

2009; Yeates et al. 2009; Bollmer et al. 2012; Gasparini et al. 2015). MHC-assortative mate choice is consistent with mechanisms evolved to limit out-breeding depression, such that MHC-assortative mating (operating through behavioural choice and/or intrinsic physiological mechanisms) would prevent the production of offspring with highly MHC divergent genes, which are hypothesized to be less likely to survive. This survival disadvantage potentially operates through highly divergent MHC genes associated with a smaller T-cell repertoire (Lawlor et al. 1990; Nowak et al. 1992), increased risk of autoimmune diseases and disruption of local adaptations or co-adapted gene complexes (Kaufman 1999; Hendry et al. 2000; Neff 2004). Selection for local adaptation, particularly to endemic pathogenic challenges, could drive mate choice towards genetically similar individuals that carry specific MHC genotypes (Dionne et al. 2007). British badger populations have a reduced genetic diversity compared to other populations in mainland Europe due to a genetic bottleneck (Frantz et al. 2014), and this low genetic diversity may be further exacerbated by MHC class II-assortative mate choice. New MHC genetic variants, arising from mutation, recombination and immigration events, would be selected against through MHC-assortative mate choice if they were very divergent from those alleles already present in the population. Consequently, the frequency of common MHC alleles might be maintained or increased, and new rare alleles should not increase in frequency if sexual selection on this genetic variation was greater than natural selection, such as pathogen-mediated selection.

Studies of MHC-based mate choice can give different patterns depending on the MHC gene examined (Huchard et al. 2013). The divergent patterns of MHC class I and class II genes in contemporary sockeye salmon populations suggest different MHC genes might be responding to different selective pressures (McClelland et al. 2013). MHC-disassortative mating was reported based on MHC class II B but not class I genes in blue petrel (Strandh et al. 2012). The MHC-based mating observed in our study appears to be mediated by the MHC class II gene only, but not by the MHC class I gene. MHC class II molecules principally bind exogenous antigens, such as those derived from extracellular pathogens such as most bacteria, while MHC class I molecules present intracellular antigens (Hughes & Yeager 1998). Bacteria are known to contribute to the actual odour of the host, through metabolization of the scent gland secretion (Singh et al. 1990), offering a potential link between mate choice and MHC genes (Brown 1995). MHC-based odour differences have been proposed to be produced directly by MHC-specific microbial communities or by secondary metabolites generated by those bacteria (Penn & Potts 1998). Badgers exchange information about individual specific
parameters through subcaudal gland scent (Buesching et al. 2002). If the microbiota found in the subcaudal gland secretion (Sin et al. 2012a) is affected by MHC genotypes, then the secondary metabolites generated could encode information on the host’s MHC profile.

There were considerable annual fluctuations in the occurrence of MHC-based mate choice, either relating to real differences in interannual mate-choice criteria, or as an artefact of varying statistical power. We only analysed the 6 years in which the most cubs were produced, which is related to environmental conditions such as food availability (Macdonald et al. 2009); thus, we were unable to identify robustly which annual conditions favour MHC-based mating. Nevertheless, the implication is that MHC-based mating might have an environment-dependent basis if it only happens in specific years related to cub production, given that the 2 years in which the mate choice was not seen were in 2 of the 3 years with the fewest cubs. Consequently, other mechanisms, such as reproductive suppression, which in contrast to other species (e.g. Clutton-Brock et al. 2010) operate more strongly in years with high food availability in the study population (Woodroffe & Macdonald 1995), might also affect MHC-based mate choice. A further possibility is that MHC-based mate choice might be affected by ecological constraints such as interannual variation in disease risk (Jaatinen et al. 2012), and the specific conditions required for MHC-based mate choice might vary over time.

In conclusion, our evidence of MHC class II genes assortative mate choice in badgers contrasts with other research reporting MHC-dissimilar mating patterns. Importantly, the identification of MHC-based patterns depends on the MHC genes examined, mating group compositions and interannual criteria. This highlights the importance of examining both MHC class I and II genes, over a broad time series, within the context of the species’ mating system, to inform the evolution of mate choice mechanisms.

Acknowledgements

We thank P. Nouvellet and S. Ellwood for assistance with badger trapping, and A. Krupa, D. Dawson, G. Horsburgh and M.-E. Mannarelli for assistance with laboratory work at the NERC Biomedical Analysis Facility, Sheffield. We are grateful to three anonymous reviewers for critical comments that greatly improved the manuscript. Y.W.S. was supported by the Croucher Foundation (Hong Kong) and H.L.D.B. by the Netherlands Organization for Scientific Research (NWO) and a NERC fellowship (NE/102178/1). Trapping protocols were subject to ethical review and were performed under Natural England (Badger Act, 1992) Licence (currently 20104655) and UK Home Office (Animals (Scientific Procedures) Act, 1986) Licence (PPL 30/2835).

References


© 2015 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd.
MHC AND MATE CHOICE IN MELES MELES


McClelland EK, Ming TJ, Tabata A et al. (2013) Patterns of selection and allele diversity of class I and class II major histocompatibility loci across the species range of sockeye salmon (*Oncorhynchus nerka*). *Molecular Ecology, 22*, 4783–4800.


Strandh M, Westerdahl H, Pontarp M et al. (2012) Major histocompatibility complex class II compatibility, but not class I, predicts mate choice in a bird with highly developed olfac-
Yamazaki K, Boyse EA, Mike V et al. (1976) Control of mating preference in mice by genes in the major histocompatibility complex. The Journal of Experimental Medicine, 144, 1324–1335.

Y.W.S., H.L.D., C.N. and D.W.M. planned the study. Y.W.S., H.L.D., G.A., C.N. and C.D.B. conducted the field research as part of D.W.M.’s long-term field study. The laboratory work was conducted in T.A.B.’s laboratory, where Y.W.S. genotyped the MHC loci and G.A. and H.L.D. genotyped the microsatellite loci. Y.W.S. analysed the data and created the first draft of the manuscript, to which all the authors then contributed.

Data accessibility
MHC and microsatellite genotyping data, parentage and social group information, and lists of candidate male mating partners are available via Dryad (doi:10.5061/dryad.d8080).

Supporting information
Additional supporting information may be found in the online version of this article.
Table S1 The number of assigned within-group and neighboring-group parent pairs, total number of candidate fathers and mean number of candidate father per female in different years and MHC datasets.
Fig. S1 Mean MHC class I and class II DRB allele sharing, mean amino acid distance (all amino acid sites and ABS only) and mean Queller & Goodnight’s index of pairwise relatedness (Q&G’s r) of assigned within-group and neighboring-group parent pairs (solid line), compared to random within-group and neighboring-group parent pairs.
Fig. S2 Mean MHC class I (a–f) and class II DRB (g–l) allele sharing, mean amino acid distance (all amino acid sites and ABS only) and mean Queller & Goodnight’s index of pairwise relatedness (Q&G’s r) of assigned parent pairs (solid line).
Fig. S3 Mean MHC class I and class II DRB functional distance of combined (both within-group and neighboring group), within-group (WG), and neighboring-group (NG) assigned parent pairs (solid line), compared to random combined, within-group, and neighboring-group parent pairs, respectively.
Fig. S4 Plots of MHC class II DRB amino acid distances (all sites) against Queller & Goodnight’s index of pairwise relatedness, for assigned within-group and neighbouring-group parent pairs.

© 2015 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd.