Evolution, 2023, **77(1)**, 221–238
https://doi.org/10.1093/evolut/qpac014
Advance access publication 9 December 2022
Original Article

Both selection and drift drive the spatial pattern of adaptive genetic variation in a wild mammal

Derek Kong Lam[1], Alain C. Frantz[2], Terry Burke[3], Eli Geffen[4], Simon Yung Wa Sin[1]

[1]School of Biological Sciences, The University of Hong Kong, Hong Kong, China
[2]Musée National d'Histoire Naturelle, Luxembourg, Luxembourg
[3]Ecology and Evolutionary Biology, School of Biosciences, University of Sheffield, Sheffield, United Kingdom
[4]School of Zoology, Tel Aviv University, Tel Aviv, Israel

Corresponding author: School of Biological Sciences, The University of Hong Kong, Pok Fu Lam Road, Hong Kong, China. Email: sinyw@hku.hk

Abstract

The major histocompatibility complex (MHC) has been intensively studied for the relative effects of different evolutionary forces in recent decades. Pathogen-mediated balancing selection is generally thought to explain the high polymorphism observed in MHC genes, but it is still unclear to what extent MHC diversity is shaped by selection relative to neutral drift. In this study, we genotyped MHC class II DRB genes and 15 neutral microsatellite loci across 26 geographic populations of European badgers (*Meles meles*) covering most of their geographic range. By comparing variation of microsatellite and diversity of MHC at different levels, we demonstrate that both balancing selection and drift have shaped the evolution of MHC genes. When only MHC allelic identity was investigated, the spatial pattern of MHC variation was similar to that of microsatellites. By contrast, when functional aspects of the MHC diversity (e.g., immunological supertypes) were considered, balancing selection appears to decrease genetic structuring across populations. Our comprehensive sampling and analytical approach enable us to conclude that the likely mechanisms of selection are heterozygote advantage and/or rare-allele advantage. This study is a clear demonstration of how both balancing selection and genetic drift simultaneously affect the evolution of MHC genes in a widely distributed wild mammal.

Keywords: balancing selection on supertype, genetic drift, major histocompatibility complex, European badgers, pathogen-mediated selection, genetic diversity

How the diversity of polymorphic genes is maintained in natural populations is one of the central questions in evolutionary biology. Balancing selection maintains genetic diversity at loci important for adaptation, while positive selection and demographic processes such as genetic drift may cause an allele to either be fixed or lost randomly (Nei et al., 1975). The extent to which genetic variation at adaptive loci is shaped by selection has been an area of intense interest. The major histocompatibility complex (MHC) genes are ideal candidates for studying this topic, due to their highly polymorphic nature and important role in the adaptive immune system (Piertney & Oliver, 2006; Radwan et al., 2020; Spurgin & Richardson, 2010). They encode a variety of cell-surface glycoproteins, which bind pathogen peptides and present them to T cells to trigger the adaptive immune response (Matsumura et al., 1992; Swain, 1983). Both selection and neutral processes could play a role in shaping MHC diversity, however, how these two distinct forces interact to shape the MHC diversity, specifically how variation at the genetic and functional levels is preserved across a large spatial scale for a long period of time, remains unclear.

The maintenance of MHC polymorphism in wild populations has been generally attributed to pathogen-mediated balancing selection (PMBS), with relatively less attention on the roles of sexual selection (Bernatchez & Landry, 2003; Edwards & Hedrick, 1998; Jeffery & Bangham, 2000; Martinsohn et al., 1999; Penn & Potts, 1999; Radwan et al., 2020; Spurgin & Richardson, 2010). A strong theoretical framework has been established to explain the mechanisms of PMBS, including heterozygote advantage, fluctuating selection, and rare-allele advantage/negative frequency dependent selection (Apanius et al., 1997; Bernatchez & Landry, 2003; Doherty & Zinkernagel, 1975; Hill, 1991; Hughes & Nei, 1988; Spurgin & Richardson, 2010; Takahata & Nei, 1990). The divergent allele advantage hypothesis was later developed based on heterozygote advantage (Wakeland et al., 1990). However, it remains difficult to tease apart the specific mechanism, since the alternatives are not mutually exclusive (Bernatchez & Landry, 2003; Piertney & Oliver, 2006). One approach to determining the mechanisms driving MHC variation is to compare the spatial structuring at MHC relative to that at neutral loci in a natural setting (Figure 1; Alcaide et al., 2008; Ekblom et al., 2007; Evans et al., 2010; Li et al., 2016; Radwan et al., 2020; Spurgin & Richardson, 2010). Under different pathogen-mediated selection mechanisms, the spatial structuring of MHC genes relative to neutral loci could be either stronger or weaker (Figure 1). If the structuring at MHC is similar to that of neutral loci, it implies neutral processes could be the dominating force shaping MHC diversity. However, the specific nature of the rare alleles affected

Received December 5, 2021; revisions received October 3, 2022; accepted November 4, 2022

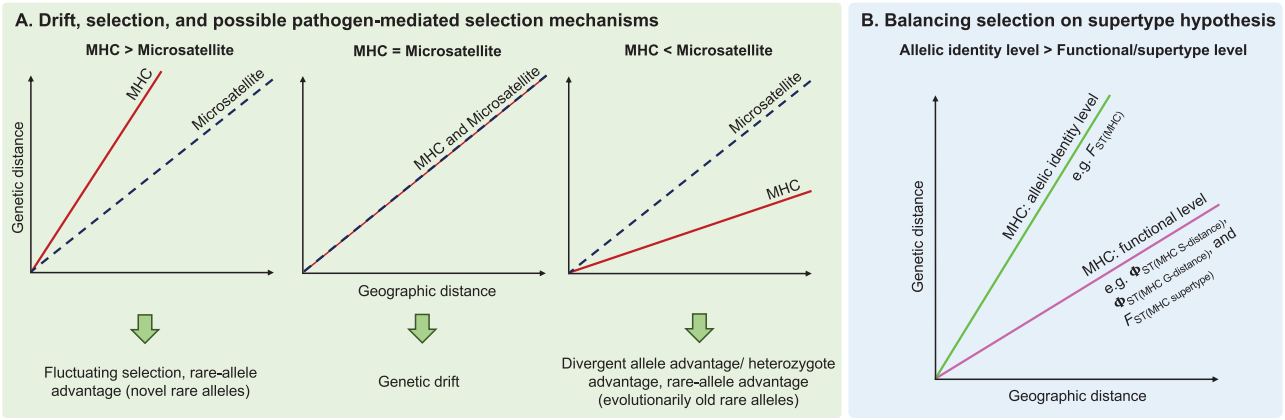

Figure 1. Predictions of the geographic structuring at major histocompatibility complex (MHC) arising from (A) different models of pathogen-mediated selection and (B) balancing selection on supertype hypothesis. (A) In continuous populations where the gene flow is restricted by dispersal, three potential scenarios of population structuring between MHC and neutral markers were predicted (i.e., stronger structuring at MHC, equal structuring at MHC and neutral loci, and weaker structuring at MHC). A steeper slope indicates a higher level of population structuring. Stronger structuring at MHC than microsatellites indicates higher divergence at MHC than neutral expectation, which could be driven by fluctuating selection or rare-allele advantage that favors novel rare alleles due to the distinct pathogen-mediated selective pressure in different local environments. Fluctuating selection is defined by the fluctuations in the direction of selection due to spatial-temporal heterogeneity in composition and abundance of pathogens caused by other factors than host-pathogen coevolution (Spurgin & Richardson, 2010). Rare-allele advantage (aka. negative-frquence dependecne seletion, NFDS) is charaterized by the cyclical host-pathogen co-evolutionary arms race. If drift plays a dominating role in shaping MHC diversity, the structuring at MHC is predicted to be similar to that of microsatellites, assuming a similar mutation rate. On the other hand, a weaker structuring at MHC than microsatellite due to sharing of similar MHC alleles by distant populations could be driven by divergent allele advantage/heterozygote advantage and rare-allele advantage that favors evolutionarily old rare alleles. (B) Balancing selection on supertype hypothesis proposes that MHC alleles should be grouped into immunological supertypes based on similarities of the physicochemical properties of the peptide-binding sites, and alleles in each supertype should have a similar binding repertoire. This hypothesis predicts a weaker population structuring of MHC at the functional/supertype level than that at the level of allelic identity, due to the action of balancing selection on MHC supertypes. The statistics that can be used to investigate between-population MHC diversity are indicated below the lines (see Materials and methods and Supplementary Material S1 for the details of the statistics).

by negative frequency dependent selection further complicates the issue, as whether the rare allele is evolutionarily old (i.e., previously common in the population but now rare) or newly mutated can be confounded with the outcomes of heterozygote advantage/divergent allele advantage and fluctuating selection (Spurgin & Richardson, 2010). Both fluctuating selection and negative frequency dependent selection that favors newly mutated rare alleles can result in greater structuring at MHC than neutral loci, while heterozygote advantage and negative frequency dependent selection favoring evolutionarily old rare alleles have the opposite effect (Spurgin & Richardson, 2010).

Indeed, different patterns of geographic structuring of MHC have been detected in wild populations, but a stronger structuring at MHC than neutral loci has been more commonly reported (e.g., Cammen et al., 2011; Eizaguirre et al., 2012; Ekblom et al., 2007; Landry & Bernatchez, 2002; Li et al., 2016; Loiseau et al., 2009; Miller et al., 2001). Most studies conclude that this pattern of population structuring was due to diversifying selection across populations driven by locally specific pathogens. Such fluctuating selection, defined by fluctuations in the direction of selection due to spatial-temporal heterogeneity in composition and abundance of pathogens caused by external abiotic or biotic factors, can result in a higher spatial structuring at MHC over neutral loci and maintain MHC diversity (Figure 1; Radwan et al., 2020; Spurgin & Richardson 2010).

In addition to selection, genetic drift and demographic events will also shape MHC variation. When population size becomes small and gene flow is restricted, the effect of genetic drift may exceed that of balancing selection, causing a reduction in or loss of MHC diversity (Campos et al., 2006; Miller et al., 2010; Miller & Lambert, 2004; Strand et al., 2012; Zhai et al., 2017). However, high MHC polymorphism may persist in some small populations or after a genetic bottleneck (Aguilar et al., 2004; Minias et al., 2019; Zhang et al., 2018). On the other hand, strong gene flow can result in there being no spatial structuring in MHC over a large geographic scale (Gillingham et al., 2017). The extent to which, and under what conditions, adaptive MHC genes are influenced by neutral drift and gene flow relative to selection is still not well understood.

To date, many different statistical measures were used to capture MHC diversity. For example, studies investigating spatial MHC diversity often report only one population statistic (e.g., F_{ST} or Φ_{ST}) based on either individual alleles or allelic lineages (Babik et al., 2008; Biedrzycka & Radwan, 2008; Bichet et al., 2015; Ekblom et al., 2007; Landry & Bernatchez, 2002; Loiseau et al., 2009). The commonly used F_{ST} based on the individual allele does not include the information at the sequence level. However, the unit for selection to act on (e.g., allelic identity vs. sequence) could vary among study systems. Only reporting one of the statistics might therefore bias the conclusions toward either drift or selection due to the complex nature of MHC evolution. To consider the functional aspect, MHC alleles can be grouped into immunological supertypes based on the similarity of their binding repertoires (Sidney et al., 1996). Earlier models of PMBS have not considered this immunological aspect of MHC genes, and earlier studies very often did not test different levels (e.g., supertype vs. allelic levels) of MHC diversity systematically. The balancing selection on supertype hypothesis has recently been proposed for explaining trans-species polymorphism (Lighten et al., 2017; but see Ejsmond et al., 2018; Herdegen-Radwan et al., 2021), but it may also explain the maintenance of diversity and lack of spatial structuring

of MHC genes across populations. The hypothesis proposes that different evolutionary mechanisms can operate at the supertype versus allelic levels, in which balancing selection is more likely to operate at the functional level defined by the supertypes. A weaker structuring at the functional level of MHC genes is therefore predicted under this hypothesis (Figure 1; Herdegen-Radwan et al., 2021), while neutral processes like genetic drift may have a larger effect at the allelic level. However, empirical support for this hypothesis is still rare (Herdegen-Radwan et al., 2021; Lighten et al., 2017; Vlček et al., 2016).

Additionally, relative signals of diversifying selection, balancing selection, or genetic drift on MHC genes has been proposed to depend on the spatial or temporal scales of the study (Herdegen et al., 2014). Comprehensive knowledge and sampling are therefore crucial to a proper understanding of the underlying mechanisms that have shaped MHC diversity. The European badger (*Meles meles*) is an ideal species in this regard, as it has been extensively studied regarding its population structure (Frantz et al., 2014; Guerrero et al., 2018; Pope et al., 2005), demographic history (Frantz et al., 2014; Pope et al., 2005), mating system (Annavi et al., 2014; Dugdale et al., 2007; Sin et al., 2015), communication (Buesching et al., 2016; Sin et al., 2012c), pathogens and immunity (Corner et al., 2011; Bilham et al., 2013; Lizundia et al., 2011; Sin et al., 2014, 2016), and MHC genes (Abduriyim et al., 2017, 2019; Sin et al., 2012a, b, 2015). *M. meles* is a nocturnal and fossorial carnivore. It has an extensive geographic distribution across Europe, extending from the British Isles to the Middle East (Macdonald et al., 2004; Rogers et al., 1997; Rosalino et al., 2004; Roper, 2010). Badger numbers in multiple European populations were reduced historically due to human-mediated habitat modification and/or disease management in the mid-20th Century (Griffiths & Thomas, 1993; van der Zee et al., 1992; Wilson et al., 1998). Some northern populations (e.g., Sweden, Britain, Ireland, Denmark, Norway, etc.) were shown to have experienced genetic bottlenecks, which resulted in reduced neutral genetic diversity (Frantz et al., 2014; Pope et al., 2005). Furthermore, peripheral populations in Iberia, Ireland, Britain, Denmark, and Scandinavia were shown to be well differentiated from the remaining mainland populations, probably due to post-glacial expansion processes (Frantz et al., 2014). However, it remains unknown whether such demographic processes also left a signal in adaptive genes, e.g., the MHC genes, in *M. meles*. Furthermore, a pronounced isolation-by-distance (IBD) relationship detected at both the local and global scales (Pope et al., 2005), and the philopatric behavior of this species (Woodroffe et al., 1995), suggest that gene flow is limited, which makes *M. meles* an ideal candidate in which to study the relative roles of drift and selection.

The MHC genes of *M. meles* have been characterized for both class I (mainly responsible for recognizing intra-cellular peptides) and class II genes (mainly responsible for recognizing extra-cellular peptides), with evidence of balancing selection being found (Abduriyim et al., 2017, 2019; Sin et al., 2012a, b). MHC-assortative mate choice based on the DRB gene, the most polymorphic MHC class II gene (Sin et al., 2012b) was detected in a British population (Sin et al., 2015). This mate choice behavior was proposed to explain the low MHC polymorphism observed in the British population, in addition to the effects of genetic drift due to bottlenecks. Meanwhile, support for PMBS was also found in *M. meles* in a British population (Sin et al., 2014). Whether those populations outside Britain that did not go through bottlenecks also exhibit a low degree of MHC polymorphism is unknown. Continent-wide sampling offers the potential to provide a deeper understanding of MHC evolution in the species.

In this study, we investigated whether MHC diversity in *M. meles* was shaped by selection, drift, or gene flow. We genotyped the MHC class II DRB genes and 15 microsatellite loci across 26 geographic populations, which spanned most of its biogeographic range. By comparing the population genetic structure of MHC and microsatellite loci, we determined the relative roles of past demographic events, genetic drift, and selection in shaping spatial MHC diversity, and identified the likely PMBS mechanism (See Supplementary Material S1 for the analytical framework). We conducted the analysis at three levels (i.e., allelic identity, amino acid sequence, and supertype) to investigate the genetic structuring of MHC relative to that of microsatellites. This analytical approach allows us to test the effects of evolutionary forces at different levels, and specifically to test the balancing selection on supertype hypothesis. This study clearly demonstrates how both selection and genetic drift can simultaneously influence the evolution of a functionally important gene family.

Materials and methods

Sample collection

We collected tissue or hair samples from a total of 499 individuals from one Middle Eastern country (Israel) and 16 European countries (including Austria, Belgium, Croatia Denmark, France, Germany, Britain, Ireland, Israel, Italy, Luxembourg, Poland, Portugal, Serbia, Sweden, and Switzerland). Since most of the samples in the present study are the same as in previous studies (Frantz et al., 2014; Pope et al., 2005), we followed previously defined criteria to assign samples into several geographic populations. Only samples from Great Britain and Poland were sub-divided into different populations, thus giving rise to a total of 26 geographic populations in our study (Supplementary Material S2). Tissue samples from road-killed or legally harvested individuals were stored in 96% ethanol until DNA extraction. DNA was extracted from tissue samples using an ammonium acetate-based salting-out procedure (Miller et al., 1988), while a Chelex protocol (Walsh et al., 2013) was used for hair samples.

MHC genotyping

The ability of MHC class II molecules to distinguish different suits of pathogen peptides relies on their peptide-binding sites (PBSs; Swain, 1983), which are on the β1 domain encoded by exon 2. We, therefore, genotyped the exon 2 of the MHC class II DRB genes, in which at least two loci were found (Abduriyim et al., 2017; Sin et al., 2012b, 2015). Individuals were genotyped using reference strand-mediated conformation analysis (RSCA) and cloning and sequencing. We first amplified exon 2 of the DRB genes using published primers designed specifically to badgers (Meme-DRBex2F and MEME-DRBex2R; Sin et al., 2012b). The amplified MHC sequences were then genotyped using RSCA following Sin et al. (2015). The identities of the alleles in RSCA were confirmed by cloning and sequencing following the protocols in Sin et al. (2012b, 2015). Each clone were both sequenced and analyzed by RSCA. Each RSCA allele were verified by

five different clones from three individuals on average. All RSCA alleles were unique MHC alleles, except for one RSCA allele that contains two alleles (*Meme-DRB*04* and *Meme-DRB*13*). These two alleles differed by just a single base pair could not be separated by RSCA only, and were sequenced using the newly-designed forward primer 5′-CTGATGCAGTTTAAGGGC-3′, and reverse primer 5′-CTCGCCGCTGCACC-3′ (Meme-DRBex2R) from Sin et al. (2012b).

Microsatellite genotyping

Out of 499 individuals, 421 individuals originating from all countries were genotyped using 18 microsatellite markers (*Mel 101–Mel 104*, *Mel 106–Mel 108*, *Mel102*, *Mel112–Mel115* and *Mel117* (Carpenter et al., 2003); *Mel1*, *Mel10*, *Mel12*, *Mel14I*, and *Mel15* (Domingo-Roura et al., 2003)) following procedures described in Frantz et al. (2014). Since three loci (*Mel12*, *Mel102*, and *Mel110*) were poorly genotyped for some populations, only 15 loci were included in downstream analyses. In Great Britain, microsatellite data were not obtained for five populations (i.e., Ipswich, Carlisle, Leeds, Middlesbrough, and Scarborough).

Data analysis

Phylogenetic analysis

Phylogenetic trees were constructed using MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003). Model testing was first performed using JModelTest (Posada, 2008) to ascertain the best-fit model of nucleotide substitution, which was identified as the general time-reversible model with gamma distributed rates. A Markov chain Mont Carlo (MCMC) sampling procedure was run for 30,000,000 generations. The mean standard deviations of split frequencies were <0.01. The first 25% of the tree samples were discarded as burn-in. MHC DRB exon 2 sequences of *M. meles* and other related species (*Meles anakuma*, *Meles canescens*, *Meles leucurus*, *Gulo gulo*, *Mustela itatsi*, *Mustela sibirica*, *Mustela lutreola*, *Enhydra lutris*, *Taxidea taxus*, *Zalophus californianus*, *Canis latrans*, *Canis lupus*, and *Canis familiaris*; see Figure 4 for the accession number) were included in the phylogenetic analysis to determine if there is a signal of trans-species polymorphism in DRB genes, which is a common phenomenon of MHC genes. The DRB exon 2 sequence from *Homo sapiens* was also included as an outgroup.

Selection analysis

We detected the presence of recombination using GARD (Kosakovsky Pond et al., 2006) implemented in HyPhy (Kosakovsky Pond et al., 2019) prior to selection inference. One significant recombination breakpoint was detected, thus data were partitioned into two parts and tested for positive and purifying selection using MEME (Murrell et al., 2012), FEL (Kosakovsky Pond & Frost, 2005), FUBAR (Murrell et al., 2013), and SLAC (Kosakovsky Pond & Frost, 2005) implemented in HyPhy. In addition, the rates of synonymous (d_S) and non-synonymous (d_N) substitutions were also calculated separately for PBS and non-PBS using HyPhy implemented in MEGA7 (Kumar et al., 2016), according to the method of Nei and Gojobori (Nei & Gojobori, 1986) with Jukes and Cantor (Jukes & Cantor, 1969) correction. The prior assignment of PBS and non-PBS were based on human MHC class II HLA molecules (Reche & Reinherz, 2003). A ratio ω (non-synonymous over synonymous substitutions, d_N/d_S) larger than 1 indicates positive selection.

MHC supertype grouping

We grouped the identified MHC alleles into four immunological supertypes based on similarities of the physicochemical properties of the putative PBSs and positively selected sites inferred from the selection analysis (Herdegen-Radwan et al., 2021; Lighten et al., 2017). The positively selected sites are likely to be involved in peptide binding, because their positions are near the putative PBSs and they were under positive selection. We first used principal component analysis (PCA; *R* package *stats*) and multidimensional scaling (MDS; *R* package *vegan*) to examine the relationships among all putatively functional DRB alleles in *M. meles* (including previously identified alleles from GenBank) based on the amino acid distance. Both Sandberg's and Grantham's distances (Grantham, 1974; Sandberg et al., 1998) were used to calculate the amino acid distance between each pair of alleles (*R* package *MHCtools*). The Sandberg's distance depends on five amino acid physicochemical descriptors (*z1*, hydrophobicity; *z2*, steric bulk; z3, polarity; *z4* and *z5*, electronic effects), while Grantham's distance depends on three properties of an amino acid (composition, polarity, and molecular volume). Next, we used ward hierarchical clustering method to group the alleles into supertypes using the five amino acid physicochemical descriptors. Results from the PCA, MDS, and clustering analyses suggest that alleles identified in this study can be grouped into four supertypes (Supplementary Material S4).

Within population diversity analysis

It is common to amplify more than one MHC locus at a time, making it difficult to assign specific alleles to their corresponding locus due to concerted evolution, recombination between alleles and possible copy number variation (Abduriyim et al., 2017; Eimes et al., 2011; Ekblom et al., 2007; Gillingham et al., 2017; Lighten et al., 2014; Loiseau et al., 2009; Miller et al., 2010; Siddle et al., 2010; Strand et al., 2012; Wong et al., 2022). We therefore used several commonly employed statistical indices in place of the more usual locus-based statistics (e.g., observed heterozygosity). We first used average percentage difference (APD) to reflect general MHC within-population diversity. The APD was calculated by averaging the percentage of nucleotide alleles not shared between two individuals in the populations. The APD does not rely on allelic frequency (Gillingham et al., 2017; Lynch, 1990, 1991; Miller et al., 2010). Next, we used mean number of sequences per individual (MHC/Ind) to estimate MHC heterozygosity in a population; total sequences per population size (MHC/Pop) and theta k (θ_k) to estimate MHC allelic richness for each population; and π (mean number of pairwise differences between all pairs of MHC alleles) to reflect nucleotide diversity of MHC in a population. The use of MHC/Ind as an estimator of MHC heterozygosity assumes that there is no copy number variation. Both θ_k and π were calculated using Arlequin 3.1 (Excoffier & Lischer, 2010). Since few individuals (N = 4–9) were sampled for most of the populations in Great Britain, and small sample sizes are often associated with inaccurate estimates and unreliable results (Fumagalli, 2013; Pruett & Winker, 2008; Subramanian, 2016), all individuals from Great Britain were combined into one single population in the following analysis. We considered that this treatment would not significantly affect our results and interpretations because (a) all populations in Great Britain showed very similar levels of MHC diversity (Figure 1 and Table 1); and (b) our aim is to investigate the spatial MHC diversity over the

Table 1. Summary statistics for MHC and microsatellite diversity among populations.

Country	Region within country	MHC								Microsatellites					
		N	MHC/Ind	Total sequences	MHC/Pop	Private sequences	APD	θ_k	π	*N*	No. of alleles	Private alleles	*He*	*Ho*	Allelic richness
Austria	-	23	3.26	10.00	0.43	0.00	0.71	2.88	22.05	19	4.73	0.07	0.60	0.53	4.47
Belgium	-	28	3.57	10.00	0.36	0.00	0.60	2.57	19.46	27	4.40	0.00	0.61	0.55	4.07
Croatia	-	18	3.06	11.00	0.61	1.00	0.58	3.86	16.81	17	5.73	0.07	0.66	0.65	5.38
Denmark	-	29	2.83	9.00	0.31	0.00	0.55	2.38	19.42	28	3.93	0.13	0.48	0.47	3.59
France	-	21	3.29	8.00	0.38	0.00	0.54	2.13	21.24	20	4.33	0.00	0.51	0.52	3.99
Germany	-	16	3.06	9.00	0.56	0.00	0.71	2.97	19.57	16	4.73	0.00	0.63	0.62	4.59
Great Britain	Ipswich	6	2.17	3.00	0.50	0.00	0.11	0.89	13.09	25	4.33	0.00	0.54	0.54	3.94
Great Britain	Carlisle	5	2.80	3.00	0.60	0.00	0.13	0.85	21.07	-					-
Great Britain	Leeds	6	2.17	3.00	0.50	0.00	0.11	0.89	13.09	-					-
Great Britain	Middlesbrough	6	2.50	3.00	0.50	0.00	0.29	0.82	20.44	-					-
Great Britain	Scarborough	25	2.40	3.00	0.12	0.00	0.25	0.49	20.80	-					-
Great Britain	Scotland (North)	7	2.29	3.00	0.43	0.00	0.16	0.80	15.29	29	3.53	0.00	0.50	0.45	3.19
Great Britain	Scotland (South)	4	2.75	3.00	0.75	0.00	0.17	0.99	21.20	-					-
Great Britain	Wales (Northeast)	9	2.22	4.00	0.44	1.00	0.27	1.21	11.90	-					-
Great Britain	Wales (South)	6	2.00	2.00	0.33	0.00	0.00	0.39	8.53	30	4.07	0.13	0.52	0.53	3.48
Great Britain	Total	74	2.37	4.00	0.05	1.00	0.29	0.60	17.89	84	3.98	0.04	0.52	0.51	3.54
Ireland	-	14	2.36	4.00	0.29	0.00	0.36	0.96	20.21	14	3.33	0.00	0.53	0.45	3.32
Israel	-	23	3.09	8.00	0.35	1.00	0.61	2.11	24.48	23	4.27	1.53	0.59	0.53	4.02
Italy	-	23	2.83	9.00	0.39	0.00	0.57	2.61	21.92	21	5.00	0.27	0.61	0.50	4.47
Luxembourg	-	25	3.04	12.00	0.48	0.00	0.71	3.77	19.85	19	5.20	0.07	0.62	0.56	4.79
Poland	Eastern Poland	27	3.15	12.00	0.44	0.00	0.69	3.58	20.85	24	5.07	0.07	0.65	0.65	4.65
Poland	Western Poland	10	2.90	8.00	0.80	0.00	0.69	3.30	17.34	8	4.67	0.00	0.68	0.69	-
Portugal	-	8	3.25	7.00	0.88	0.00	0.63	2.80	21.25	8	3.27	0.00	0.51	0.53	-
Serbia	-	21	3.19	12.00	0.57	0.00	0.73	4.00	22.71	20	6.20	0.13	0.68	0.69	5.82
Spain	-	31	3.10	9.00	0.29	0.00	0.67	2.24	21.04	30	4.60	0.13	0.53	0.50	3.91
Sweden	-	29	2.93	5.00	0.17	1.00	0.51	0.99	21.48	29	3.27	0.00	0.50	0.46	3.01
Switzerland	-	14	3.07	8.00	0.57	0.00	0.57	2.63	20.14	14	4.80	0.20	0.60	0.59	4.73

Note. For MHC, MHC/Ind represents the estimated MHC heterozygosity in each population, calculated as the mean number of alleles per individual. Total alleles is the total number of MHC alleles found in each population. MHC/pop represents the total number of MHC alleles divided by the sample size (*N*), which estimates allelic richness. APD measures the average percentage difference in the number of nucleotide alleles between two individuals at the MHC in the population. $\theta_k$measures MHC allelic diversity in the population. π measures mean number of pairwise differences between all pairs of MHC alleles in a population. For microsatellites, all measure are represented as the mean value. H_e = mean expected heterozygosity at microsatellites. H_o = mean observed heterozygosity at microsatellites.

whole range of distribution instead of at local scale. Next, we compared the above measures of MHC within-population diversity using Pearson's product–moment correlation test in *R* (package *stats*), with the *p* values corrected by the Benjamini–Hochberg procedure (Haynes, 2013). If a population has a high number of alleles and high heterozygosity, then the inter-individual difference in the population is also very likely to be high. Therefore, it was expected that the APD would positively correlate with MHC/Ind, MHC/Pop, andθ_k. We also compared the values of APD and MHC/Pop for different populations using Kruskal–Wallis and Mann–Whitney *U* tests in *R* (package *stats*).

Next, for microsatellite diversity, mean number of alleles, mean number of private alleles, mean expected heterozygosity (H_e), and mean observed heterozygosity (H_o) for each population were calculated using GenAlEx 6.5 (Peakall & Smouse, 2006, 2012). Mean allelic richness per locus was calculated using FSTAT 2.9 (Goudet, 1995) with 1,000 resamplings. Small populations ($N < 14$) were removed from allelic richness analysis. In the British populations, microsatellite data were only available from three populations. These three populations, in Ipswich, North Scotland, and South Wales, had been demonstrated to be genetically distinct previously (Frantz et al., 2014; Pope et al., 2005). We, therefore, calculated the above microsatellite statistics separately for these, and used the mean value to represent the overall population in Great Britain in subsequent analyses. Next, to test for the effect of genetic drift, we compared measures of within-population diversity at MHC and microsatellite loci using Pearson's product–moment correlation test in *R* (package *stats*), with *p* values corrected by the Benjamini–Hochberg procedure.

Population differentiation analysis

Since mixed MHC data from more than one locus often violates assumptions of many commonly used population statistics (e.g., mixed inheritance pattern in the case of copy number variation, Hardy–Weinberg equilibrium), we therefore employed several approaches to calculate the fixation index at MHC genes (e.g., Bichet et al., 2015; Gillingham et al., 2017; Miller et al., 2010). We analyzed the data at two levels (i.e., allelic identity vs. functional level). At the level of allelic identity, we used three approaches (i.e., $F_{ST(MHC)}$, $F_{ST(MHC\ RFLP)}$, and $Rho_{ST(MHC)}$; see Supplementary Method for details) to calculate fixation index at MHC. In brief, $F_{ST(MHC)}$ was calculated by inputting allelic frequency as haplotype data, and $F_{ST(MHC\ RFLP)}$ was calculated with each allele treated as absence or presence like RFLP data, using Arlequin 3.1. The $Rho_{ST(MHC)}$ were calculated using SPAGeDi by inputting the DRB alleles from each individual as polyploid data at a single locus, following Gillingham et al. (2017). Since $F_{ST(MHC)}$ was highly correlated with both $F_{ST(MHC\ RFLP)}$ ($r_M = 0.96$, $p = <.001$) and $Rho_{ST(MHC)}$ ($r_M = 0.97$, $p = <.001$; Supplementary Material S5), we only reported the results of $F_{ST(MHC)}$ in subsequent analyses.

At the functional level, we used four indexes (i.e., $\Phi_{ST(MHC\ nucleotides)}$, $\Phi_{ST(MHC\ G\text{-}distance)}$, $\Phi_{ST(MHC\ S\text{-}distance)}$, and $F_{ST(MHC\ supertype)}$) to calculate the pairwise genetic population differentiation distance using Arlequin 3.1. $\Phi_{ST(MHC\ nucleotides)}$ considers the allelic genealogy, which is calculated by inputting both allelic frequencies and DNA sequences as haplotype data. Jukes & Cantor model (Erickson, 2010) was used to calculate the allelic distance. For both $\Phi_{ST(MHC\ G\text{-}distance)}$ and $\Phi_{ST(MHC\ S\text{-}distance)}$, a matrix of pairwise amino acid distance (i.e., based on the Grantham's distance or Sandberg's distance, respectively) among MHC alleles were used to calculate the fixation index. $F_{ST(MHC\ supertype)}$, as another measure of MHC functional differentiation, was calculated based on supertype frequency. The input setting in Arlequin 3.1 implies that the allelic frequency is calculated as the total number of individuals in a population carrying a particular allele, divided by the total allele count observed in the population. Note that this method of calculating allelic frequency may overestimate the frequency of rare alleles or underestimate the frequency of common alleles. This way of calculating allelic frequency and treating MHC data as haplotypic data in Arlequin 3.1 are commonly used when alleles cannot be assigned to their corresponding loci (Ekblom et al., 2007; Loiseau et al., 2009; Miller et al., 2010).

To investigate whether the genetic structuring at MHC is the result of previous demographic events, MDS was conducted in *R* (package *vegan*) to visualize the population differentiation pattern at MHC based on the pairwise $F_{ST(MHC)}$. The exact test of population differentiation (Goudet et al., 1996; Raymond & Rousset, 1995) was also performed using Arlequin 3.1 to test for population differentiation at MHC loci. To disentangle the effects of drift, selection, and the likely PMBS mechanisms on MHC, we tested for an IBD relationship at the MHC while considering the effects of stochastic demographic processes. If genetic drift is the dominant force driving spatial MHC variation, then we would expect the degree of population structuring at MHC to be similar to that at microsatellites. We first performed a simple Mantel test to test for the correlation between pairwise genetic distance at the MHC and pairwise geographic distance in *R* (package *vegan*). Next, both partial Mantel test and multiple regression on distance matrices (MRM) were further applied to consider the effects of stochastic demographic processes (Legendre & Fortin, 2010; Lichstein, 2007; Shirk et al., 2018). The MRM has been extensively used in landscape genetic studies (Balkenhol et al., 2009; Peterman & Pope, 2021), as a supplement to the partial Mantel test, to test the associations between several variables using matrix data (e.g., Andrew et al., 2012; Kanno et al., 2011). In the partial Mantel test, the pairwise genetic distance at microsatellite loci ($F_{ST(Microsatellites)}$ or $R_{ST(Microsatellites)}$) was input as the third matrix. The geographic coordinates for each population were obtained by averaging the coordinates of all individuals in that population. The pairwise genetic distances were first transformed into $F_{ST}/(1 - F_{ST})$, and then tested for their correlations with the natural logarithm of geographic distance (kilometers). Pairwise $F_{ST(Microsatellites)}$ and $R_{ST(Microsatellites)}$ were both calculated using Arlequin 3.1. In MRM, multiple linear regression was performed in *R* (package *ecodist*). Both pairwise genetic distance at microsatellite loci and pairwise geographic distance were input as the explanatory variable, while the genetic distance at MHC was input as the predicted variable. Due to the nonindependence nature of the pairwise data, the statistical significance of the model was inferred with 10,000 permutations. In addition, an analysis of molecular variance (AMOVA) was performed using Arlequin for both MHC and microsatellite loci.

To examine the effects of the MHC allele grouping treatment on the spatial structuring of MHC supertypes, we performed a simulation to randomly assign MHC alleles into supertypes, with the allele frequency remained unchanged and the number of alleles in each supertype stayed constant. Five thousand iterations were performed for the randomization.

Mantel tests were then conducted for the iterations to obtain a distribution of 5,000 r_M values, in order to determine the effect of grouping alleles into supertypes on the r_M values.

Results

MHC genotyping

In total, 434 individuals from 26 populations were genotyped at MHC class II DRB loci (Figure 2; Supplementary Materials S6 and S7). Eighteen alleles were identified, eight of which (*Meme-DRB*21–Meme-DRB*27*; *Meme-DRB*PS04*) were newly described. Two alleles, *Meme-DRB*01* and *Meme-DRB*02*, were the most frequent across the range of distribution. Allele *Meme-DRB*01* was shared in all populations but was rare in Israel, and allele *Meme-DRB*02* occurred in all populations with varying frequencies, but was absent from Israel. Four private alleles (*Meme-DRB*03*, *Meme-DRB*24*, *Meme-DRB*26*, and *Meme-DRB*27*) were found in Britain, Sweden, Croatia, and Israel, respectively. All identified alleles were putatively functional, except two (*Meme-DRB*PS01N* and *Meme-DRB*PS04*) that had pseudogenic features as a nucleotide deletion at the same position at the end of exon 2 caused a frameshift. The amplicon length for the putative functional alleles was 248 bp, spanning the majority of the exon 2. No sequence length variation was found among these alleles. Notably, two alleles (*Meme-DRB*01* and *Meme-DRB*08*) had different nucleotide composition but had an identical amino acid sequence (Figure 3).

The number of putatively functional alleles found per individual ranged from 2 to 5 (162 individuals with 2 alleles, 143 individuals with 3 alleles, 118 individuals with 4 alleles, and 11 individuals with 5 alleles), suggesting that at least 2 functional MHC class II DRB loci in most individuals of *M. meles* (Sin et al., 2015), with potential copy number variation that some individuals had at least 3 loci (Abduriyim et al., 2017). The individuals with five alleles included four individuals from Belgium, two from France and Luxembourg, and one from each of Israel, Serbia, and Spain, respectively.

Positive selection at MHC

There was evidence of historical selection on MHC class II DRB genes. Two codons were detected to be under positive selection ($p < .1$) using MEME, and four codons were detected to be under positive selection (posterior probability > .9; Supplementary Material S8) using FUBAR. These positively selected codons detected by both methods are within or near the putative PBS inferred from the human HLA molecule (Figure 3). Outside the PBS, seven and one codons were identified to be under purifying selection using FEL ($p < .1$) and FUBAR (posterior probability > .9), respectively. Overall, the ratio ω for PBS was greater than 1, and around 3 times higher than for non-PBS (Table 2), suggesting that positive selection was operating on the PBS of the DRB genes in *M. meles* across its distribution range.

Phylogenetic analysis

The phylogenetic tree showed that *Meme-DRB* alleles did not form a monophyletic group. Instead, they were widely

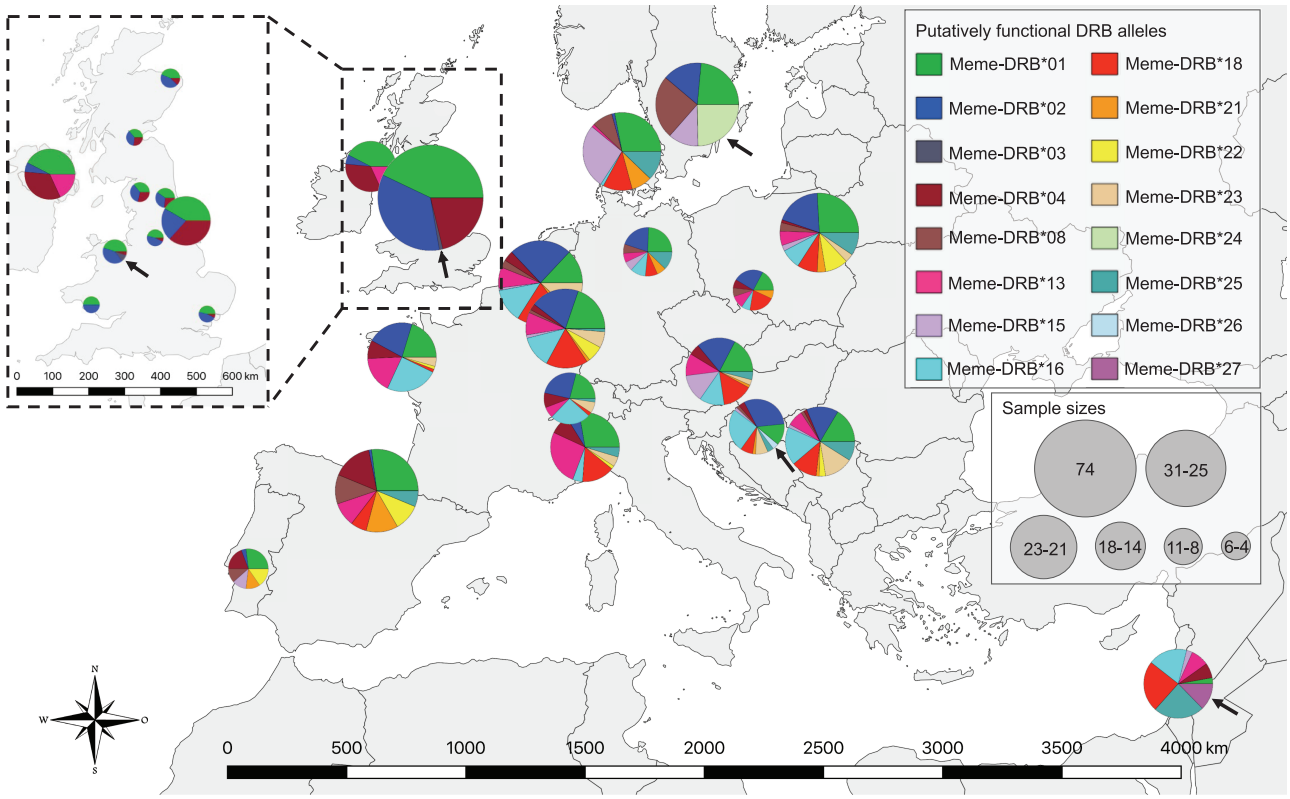

Figure 2. Sampling locations and allele distributions of major histocompatibility complex (MHC) class II DRB genes in *Meles meles*. Based on geographical criteria, samples were divided into 26 populations. Only putatively functional alleles are shown here. The pie charts indicate the allele frequencies in each population, calculated as the total number of individuals carrying the particular allele divided by the total allele count observed in the population (see Materials and Methods for details). The black arrows indicate the presence of private alleles. The distribution of MHC supertypes is provided in Supplementary Material S3.

```
                                  8          18         28         38         48         58         68         78
M. meles_Meme-DRB*01              LFLTTSECHF TNGTERVRFL DRYFYNGEEY VRFDSDVGEY RPVTELGRPI AQGWNSQKDI MEQKRANVDT YCRHNYGVGE
M. meles_Meme-DRB*02              .......... .......... .......... .......... .........D ..Y....... ..RR..A... ........V.
M. meles_Meme-DRB*03              .......... .......... .......... .......... .........D ..Y....... ..RR..E... V.......V.
M. meles_Meme-DRB*04              .MQFKG..Y. ........L. V.HI..R..F .......... .......... .......... ..RR..E... V.......V.
M. meles_Meme-DRB*05              .H......Y. .......... Q........D .........F ........RD .EY....... ..DE..A... ........F.
M. meles_Meme-DRB*06              .EQGK...Y. .......... E.HI..R..F A......... .A........ .ES..R..E. L.DA..A... ..........
M. meles_Meme-DRB*07              .EQGK...Y. .......... ...IH.R... .......... .A........ .ES..R..E. L.E...K... ..........
M. meles_Meme-DRB*08              .......... .......... .......... .......... .......... .......... .......... ..........
M. meles_Meme-DRB*09              .Y......Y. ........L. E.H......F .........F .........S ..Y......F L.DA..P... ......E...
M. meles_Meme-DRB*10              .H.AKFQ.Y. .......QL. IKAV.....F .........F .........V ..S......Y ..RM..A... ........V.
M. meles_Meme-DRB*11              .H......Y. .......... Q........L .......... .........S ..S......F L.DA..P... ......E...
M. meles_Meme-DRB*12              .L.VKP...Y C.......L. ......S... .H.N...... .......... .......... ..R..SE... V....H..F.
M. meles_Meme-DRB*13              .MQFKG..Y. ........L. V.HI..R..F .......... .......... .......... ..RR..E... V.........
M. meles_Meme-DRB*14              .H......Y. .......... Q........L .......... .........S ..S......F L.DA..P... ......E...
M. meles_Meme-DRB*15              .L.GKA.... .......... A..I..R... .......... .........T ..Y......F ..RT..A... V.........
M. meles_Meme-DRB*16              .......... .......... .......... .......... .........D ..Y....... .......... ..........
M. meles_Meme-DRB*17              .LPVKP...Y C.......L. ......S... .H.N...... .......... .......... ..R..SE... V....H..F.
M. meles_Meme-DRB*18              .......... .......... ........K. .......... .......... .......... .......... ..........
M. meles_Meme-DRB*19              .......... .......... .......... .H........ .......... .......... .......... ..........
M. meles_Meme-DRB*20              .L.VKP...Y S.......L. ......S..F .......... .......... .......... ..R..SE... V....HR.F.
M. meles_Meme-DRB*21              .......... .......... .......... .......... .......... .......... ...A..A... V.........
M. meles_Meme-DRB*22              .L.GKA.... .......... A..I..R... .......... .......... ..Y......F ..RT..A... V.........
M. meles_Meme-DRB*23              .MQFKG..Y. ........L. V..I..R..F .......... .......... .......... ..RR..E... V.......V.
M. meles_Meme-DRB*24              .MQFKG..Y. ........L. V.HI..Q..F .......... .......... .......... ..RR..E... V.......V.
M. meles_Meme-DRB*25              .L.GKA.... .......... A..I..R... ..Y....... .........T ..Y......F ..RT..A... V.......L.
M. meles_Meme-DRB*26              .......... .K........ .......... .......... .........D ..Y....... ..RR..A... ........V.
M. meles_Meme-DRB*27              .......... .......... .......... .......... .........D ..Y......L ...R..E... V.........
M. meles_Meme-DRB*PS01N           .LPVKP...Y C.......L. ......S... .H.N...... .......... .......... ..R..SE... V....H..F.
M. meles_Meme-DRB*PS02N           .L.VK....Y C......--- ---------- ---------- ---------- ---------- ---------- ----------
M. meles_Meme-DRB*PS03N           .L.VK....Y C......--- ---------- ---------- ---------- ---------- ---------- ----------
M. meles_Meme-DRB*PS04            .L.VKP...Y C.......L. ......S... .H.N...... .......... .......... ..R..SE... V....H..F.
M. anakuma_Mean-DRB*01            .H......Y. .......... Q........D .........F ........RD .EY....... ..DE..A... ........F.
M. anakuma_Mean-DRB*02            .EQGK...Y. .......... E.HI..R..F A......... .A........ .ES..R..E. L.DA..A... ..........
M. anakuma_Mean-DRB*03            .EQGK...Y. .......... ...IH.R... .......... .A........ .ES..R..E. L.E...K... ..........
M. canescens_Meca-DRB*01          .L.VKP...Y C.......L. ......S... .H.N...... .......... .......... ..R..SE... V....H..F.
M. canescens_Meca-DRB*02          .H......Y. .......... Q........D .........F ........RD .EY....... ..DE..A... ........F.
M. canescens_Meca-DRB*03          .MQFKG..Y. ........L. V.HI..R..F .......... .......... .......... ..RR..E... V.........
M. leucurus_Mele-DRB*01           .EQGK...Y. .......... ...IH.R... .......... .A........ .ES..R..E. L.E...K... ..........
M. leucurus_Mele-DRB*02           .L.GKA.... .......... A..I..R..F .......... .........D ..Y....... ..RR..E... V.........
M. leucurus_Mele-DRB*03           .EQFK..... ........L. V.SI..R... ..Y....... .........D ..Y......F ..RR..A... ........V.
M. lutreola_Mulu-DRB*5            ********** *......... E.H......F .......... .........V ..S......F V.....E... ..........
E. lutris_Enlu-DRB*01             ********** ********.. Q.....R... .......... .........D .PY....... L.DA..A... ........V.
C. lupus familiaris_DLA-DRB1*3    .EVAK...Y. .........V E..IH.R..F .........F .A.......V .ES..G..E. L..E..T... ........I.
Z. californianus_Zaca-DRB*04      .L.FKA...Y S......... V..I...... .......... .........S .EY..R...V V.....E... V.......V.
H. sapiens_HLA-DRB1*10            .EEVKF.... F.......L. E.RVH.Q... A.Y....... .A.......D .EY......L L.RR..A... ..........
                         Human PBS ^ ^ ^             ^  ^ ^      ^ ^        ^          ^    ^     ^    ^  ^    ^   ^  ^^
                               FEL-           --                                      -                 -  -     -     +
                              SLAC                                                                               -
                             FUBAR   +                                                     +          +          -     +
                              MEME                                                                             +       +
                   GARD breakpoint                                                         %
```

Figure 3. Putative amino acid alignment of major histocompatibility complex (MHC) class II DRB exon 2 in *Meles meles*. All available sequences of *Meme-DRB* genes (HQ90892.1–HQ908104.1; LC180320.1–LC180335.1; LC180338.1–LC180339.1) in GenBank were downloaded for the alignment with the newly identified *Meme-DRB* genes. The accession numbers for other DRB sequences include LC180283.1–LC180285.1 (*Meles anakuma*), LC180294.1–LC180296.1 (*Meles canescens*), LC180301.1–LC180303.1 (*Meles leucurus*), EU263554.1 (*Mustela lutreola*), EU121855.1 (*Enhydra lutris*), NM001014768 (*C. lupus familiaris*), AY491459 (*Z. californianus*), and AF225565.1 (*H. sapiens*). The number at the top indicates the position of amino acids in β1 domain. The peptide-binding sites (^) were inferred from human HLA molecules (Reche & Reinherz, 2003). Codon-based selection analyses were performed using FEL, SLAC, FUBAR, and MEME implemented in HyPhy (default settings). "+" indicates positively selected sites while "–" indicates negatively selected sites. GARD implemented in HyPhy was used to test the recombination event prior selection analysis. The recombination breakpoint is indicated by "+." The two pseudogenes (*Meme-DRB*PS01N* and *Meme-DRB*PS04*) have frameshifts after position 87 (not shown in here). "." = identical amino acid. "-" = gap. "*" = missing data.

interspersed with DRB sequences from species of the same or closely related families (e.g., *M. canescens*, *M. leucurus*, *M. anakuma*, *M. itatsi*, *M. sibirica*, *M. lutreola*, *E. lutris*, *T. taxus*, *G. gulo*, and *Z. californianus*; Figure 4). In addition, no specific clustering of *Meme-DRB* alleles was observed in relation to the geographic location.

Within-population diversity

When comparing all populations, we showed that APD (measure of inter-individual difference) was significantly correlated with MHC/Ind (estimator of heterozygosity), MHC/Pop (measure of allelic richness), and θ_k(measure of allelic richness; APD vs. MHC/Ind, r = 0.715, df = 17, p = .002; APD vs. MHC/Pop, r = 0.607, df = 17, p = .015; APD vs.θ_k, r = 0.834, df = 17, p < .001; Table 3). The MHC/Pop also significantly correlated with θ_k(r = 0.714, df = 17, p = .002). However, the π (measure of nucleotide diversity) did not correlate with any of the MHC measures (e.g., π vs. APD, r = 0.249, df = 17, p = .383; Table 3).

Different degrees of MHC diversity were observed among populations, as indicated by both MHC/Ind and APD (MHC/Ind, χ^2 = 67.4, df = 17, p < .001; APD, χ^2 = 2162.2, df = 17, p < .001). The geographically isolated populations in northern Europe (i.e., Great Britain, Ireland, and Sweden) had significantly lower allelic richness as indicated byθ_k, with only four to five alleles identified per population (Table 1). The highest number of alleles identified for a population in our dataset was 12 (i.e., Luxembourg, Eastern Poland, and Serbia). Only island populations (Great Britain and Ireland) showed significantly lower values of APD and MHC/Ind. This suggests that MHC diversity in island populations was low relative to mainland populations. By contrast, when the amino acid sequence of the MHC alleles was considered, we found that most alleles identified in the isolated populations (i.e., Great Britain, Ireland, and Sweden) belonged to different supertypes (Figure 2 and Supplementary Material S3).

For the association between MHC measures and microsatellites, we found that both MHC allelic richness and

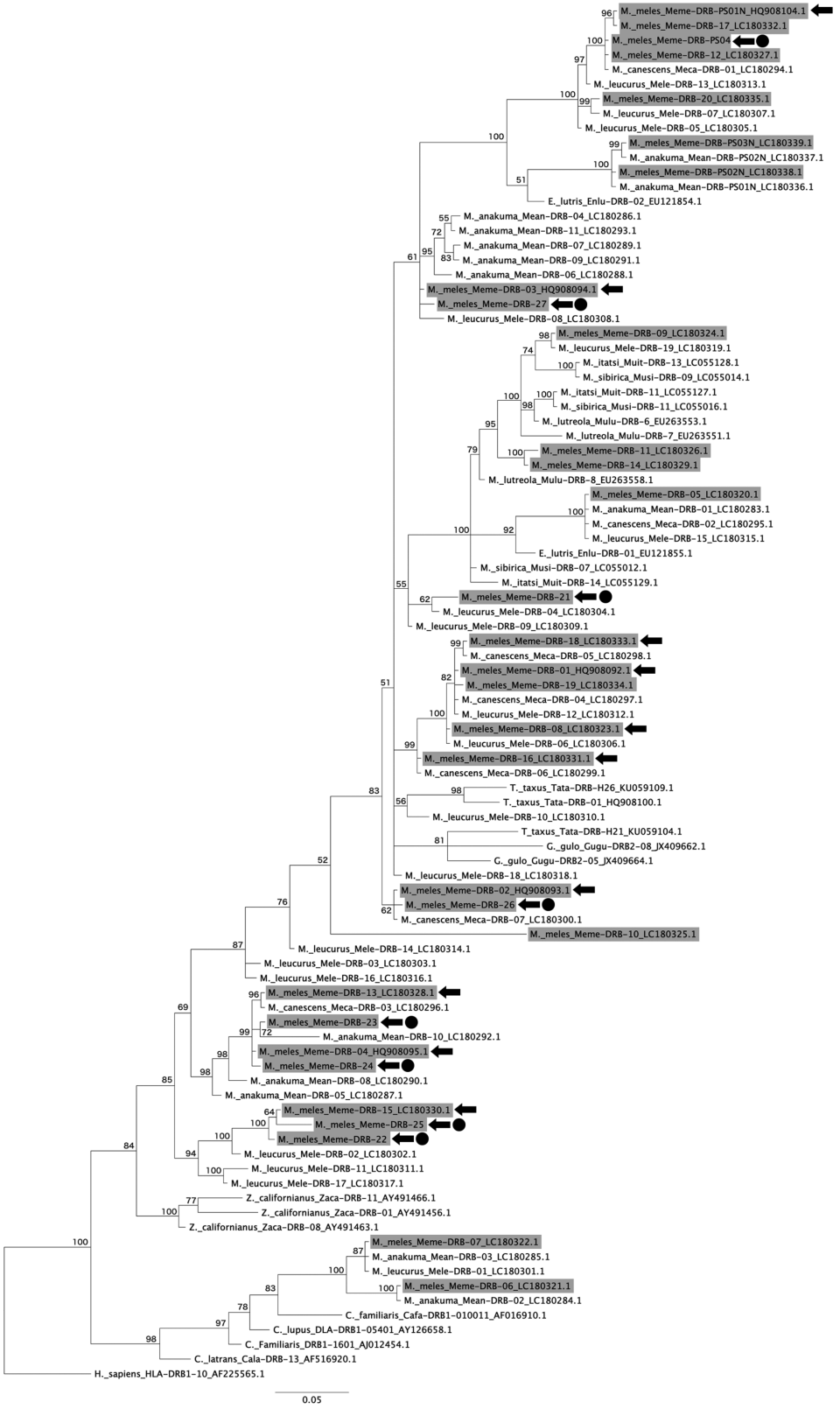

Figure 4. Phylogenetic tree of major histocompatibility complex (MHC) class II DRB exon 2 in *Meles meles*. The numbers near the tree nodes indicates the posterior probabilities. Alleles from *M. meles* are shaded with gray boxes. The black arrows indicate the DRB alleles found in this study while the black dots indicate the newly identified alleles. DRB sequences from other species were also included, with the GenBank accession number indicated after the allele name. These species include *Meles canescens,* Eurasian badger; *Meles leucurus*, Asian badger; *Meles anakuma*, Japanese badger; *Enhydra lutris*, sea otter; *Mustela itatsi*, Japanese weasel; *Mustela sibirica*, Siberian weasel; *Mustela lutreola*, European mink; *Taxidea taxus*, American badger; *Gulo gulo*, wolverine; *Zalophus californianus*, California sea lion; *Canis familiaris*, dog; *Canis lupus*, wolf; *Canis latrans*, coyote; and *Homo sapiens*, human.

Table 2. Rate (± standard error) of non-synonymous (d_N) and synonymous (d_S) substitutions for peptide-binding sites (PBS), non-PBS, and combined (PBS + non-PBS) at MHC class II DRB genes in *Meles meles*.

	No. of codons	d_N	d_S	ω
PBS	18	0.727 ± 0.285	0.979 ± 0.224	1.348
Non-PBS	62	0.276 ± 0.099	0.134 ± 0.056	0.487
Combined	80	0.377 ± 0.101	0.324 ± 0.077	0.860

Note. ω = ratio of non-synonymous to synonymous nucleotide substitution.

inter-individual difference were significantly correlated with both microsatellite heterozygosity and allelic richness (e.g., θ_k vs. H_e, r = 0.764, df = 17, p = .001; APD vs. microsatellite allelic richness, r = 0.648, df = 15, p = .014, etc.; Table 3). However, our estimates of MHC heterozygosity and nucleotide diversity did not correlate significantly with either microsatellite heterozygosity or allelic richness (e.g., MHC/Ind vs. H_e, r = 0.330, df = 17, p = .297; π vs. microsatellite allelic richness, r = –0.012, df = 17, p = .964, etc.; Table 3).

MHC population differentiation

Most of the MHC variation in our data set were observed within populations (AMOVA, Supplementary Material S9), but significant differences of allelic frequencies were still detected among populations. The pairwise $F_{ST(MHC)}$ showed that 115 of the 153 population comparisons were significantly differentiated, with the pairwise $F_{ST(MHC)}$ values for all population comparisons ranging from –0.010 to 0.244 (Supplementary Material S11). The negative $F_{ST(MHC)}$ value could be due to sampling error or uneven sample size, and can be treated as zero. There was a trend for isolated and remote populations (including Great Britain, Ireland, Israel, and Sweden) as well as populations at the periphery of Europe (including Spain, Portugal, and Denmark) to be significantly differentiated from all other populations, except for the two populations in the Iberian Peninsula (i.e., Spain and Portugal), which were not significantly differentiated from each other. In contrast, the populations in central, eastern, and southeastern Europe (i.e., Austria, Belgium, Croatia, France, Germany, Luxembourg, Poland, Serbia, and Switzerland; Supplementary Material S11) were not well differentiated from each other. Only the population in Italy, geographically separated by the Alps, was significantly differentiated from all other populations. This pattern of population differentiation was further supported by both MDS (Figure 5) and the exact test of population differentiation, in which 132 of the 153 population comparisons showed significant differentiation (Supplementary Material S10).

When taking amino acid distance into account using pairwise $\Phi_{ST(MHC\ G\text{-}distance)}$ and $\Phi_{ST(MHC\ S\text{-}distance)}$, the total number of population comparisons showing significant differentiation reduced to 81 out of 153 in both statistics (Supplementary Materials S13 and S14). No obvious pattern of population differentiation was observed in relation to the spatial locations of the populations. Pairwise $F_{ST(MHC\ supertype)}$, which groups alleles into immunological supertypes, also demonstrated a similar pattern (Supplementary Material S15). The disagreement between these fixation indexes (measure at the level of allelic identity vs. measures at the functional level) in the pattern of population differentiation can be further observed in the IBD tests. A simple Mantel test showed a significant correlation between pairwise geographic distance and pairwise genetic distance $F_{ST(MHC)}$ (r_M = 0.453, p < .001; Figure 6B), but not $\Phi_{ST(MHC\ G\text{-}distance)}$, $\Phi_{ST(MHC\ S\text{-}distance)}$, nor $F_{ST(MHC\ supertype)}$ (e.g., $\Phi_{ST(MHC\ S\text{-}distance)}$, r_M = 0.113, p = .199; $F_{ST(MHC\ supertype)}$, r_M = 0.066, p = .277; Figure 6, Supplementary Material S18). In addition, the population differentiations at the functional levels were significantly lower than allelic level (e.g., $F_{ST(MHC)}$, mean = 0.065; $F_{ST(MHC\ supertype)}$, mean = 0.044; Wilcoxon ran sum test, p = < .001, Supplementary Materials S16 and S17). Since nucleotide distance also contains functional information (Supplementary Material S4), $\Phi_{st(MHC\ nucleotides)}$ was correlated with other indexes at the functional level. The results based on the nucleotide distance were similar to those based on the amino acid distance (i.e., $\Phi_{ST(MHC\ G\text{-}distance)}$ and $\Phi_{ST(MHC\ S\text{-}distance)}$; Supplementary Materials S17–S20).

Next, we also detected a significant signal of IBD at microsatellite loci ($F_{ST(Microsatellite)}$, r_M = 0.692, p < .001, Figure 6A). We therefore tested whether population differentiation at MHC observed using $F_{ST(MHC)}$ was a by-product of neutral differentiation. After controlling for pairwise genetic distances among microsatellites, the partial Mantel test showed that pairwise $F_{ST(MHC)}$ was no longer correlated with geographic distance (controlling for $F_{ST(Microsatellite)}$, partial r_M = 0.024, p = .415; Supplementary Material S18). This conclusion was further supported by the MRM test (Supplementary Material S19). In line with this, we also detected a significant correlation between the pairwise genetic distances at MHC and microsatellite loci ($F_{ST(MHC)}$ vs. $F_{ST(Microsatellite)}$, r_M = 0.636, p < .001; Supplementary Material S16).

The randomization test showed that grouping alleles into supertypes slightly lowered the r_M value in general, with the mean r_M of the randomization being at 0.28 (Supplementary Material S22) compared to r_M = 0.45 for $F_{ST(MHC)}$. The observed r_M was at a marginally significant value of 5.92% of the simulated data, indicating a reduced spatial structuring at MHC supertypes. Attention should be paid that the lowest 1.5% of the randomized r_M values were below 0, i.e., a negative relationship between MHC distance and geographic distance, which is unlikely to occur in natural populations. $F_{ST(MHC\ supertype)}$ was highly correlated with other indexes with functional information (i.e., $\Phi_{ST(MHC\ nucleotides)}$, $\Phi_{ST(MHC\ G\text{-}distance)}$, and $\Phi_{ST(MHC\ S\text{-}distance)}$; Supplementary Material S18), and the pattern of population differentiation based on $F_{ST(MHC\ supertype)}$ was also similar to those based on other indexes at the functional level (Supplementary Material S20). In addition, most of the populations (>72%) showed a significantly reduced mean pairwise supertype population differentiation compared to the simulated dataset (Supplementary Material S21).

Discussion

We demonstrate that MHC diversity in *M. meles* across its geographic range is shaped by both selection and drift. When

Table 3. Correlations among populations between alternative diversity measures for the MHC and microsatellites using Pearson's product–moment correlation.

	MHC					π	Microsatellites		
	MHC/Ind	Total alleles	MHC/Pop	APD	θ_k		No. of alleles	H_e	Allelic richness
MHC/Ind		**0.640**	0.416	**0.715**	**0.571**	0.292	0.330	0.293	0.444
Total alleles	**0.011**		0.386	**0.805**	**0.909**	0.078	**0.819**	**0.660**	**0.800**
MHC/Pop	0.144	0.170		**0.607**	**0.714**	−0.133	0.299	0.503	**0.865**
APD	**<0.001**	**<0.001**	**0.016**		**0.834**	0.249	**0.571**	**0.593**	**0.648**
θ_k	**0.025**	**<0.001**	**0.003**	**<0.001**		−0.065	**0.793**	**0.764**	**0.902**
π	0.326	0.798	0.690	0.399	0.798		−0.058	−0.136	−0.012
No. of alleles	0.251	**<0.001**	0.287	**0.024**	**<0.001**	0.868		**0.815**	**0.977**
H_e	0.287	**0.007**	0.055	**0.019**	**0.001**	0.663	**<0.001**		**0.891**
Allelic richness	0.127	**0.001**	**<0.001**	**0.015**	**<0.001**	0.964	**<0.001**	**<0.001**	

Note. The correlation *r* values are above the diagonal and corresponding corrected *p* values are below the diagonal. The Benjamini–Hochberg procedure was applied to the measures among MHC, among microsatellites, and between MHC and microsatellites, respectively. Significant correlations ($p < .05$) are shown in bold. For MHC, mean alleles per individual (MHC/Ind), total alleles (Total alleles), total alleles divided by sample size (MHC/Pop), average percentage difference in the number of alleles among individuals (APD), θ_k(theta k, index of allelic diversity), and π (mean number of pairwise differences between all pairs of alleles in the population) are reported. For microsatellites, mean number of alleles (No. of alleles), mean expected heterozygosity (H_e), and mean allelic richness (Allelic richness) are reported.

only MHC allelic identity but not allelic content (i.e., amino acid distance and supertype) was considered, MHC diversity was shown to be shaped by genetic drift, since similar patterns were observed for both MHC and microsatellites. However, when we took into account the functional measures of MHC, the pattern of population differentiation at the MHC was different from that based on microsatellites. Instead, less spatial structuring at MHC was detected compared to microsatellites. This indicates that MHC genes were also under selection, in addition to being affected by genetic drift. This pattern of population structuring is consistent with the balancing selection on supertype hypothesis.

Genetic drift and demographic processes affect MHC evolution

When only MHC allelic identity was considered, there are three major lines of evidence supporting the role of genetic drift and demographic processes in shaping the spatial MHC diversity in *M. meles*. First, we found strong positive correlations between pairwise $F_{ST(MHC)}$ and $F_{ST(Microsatellites)}$, as well as between allele richness of MHC and microsatellites. Such correlations suggest that MHC genes evolve following the same pattern observed in the microsatellites, which are likely subject to drift. Such correlations between MHC and neutral markers have been reported in a wide range of species (Aguilar & Garza, 2006; Biedrzycka & Radwan, 2008; Loiseau et al., 2009; Miller et al., 2010; Zeisset & Beebee, 2014, etc.). Second, the pairwise $F_{ST(MHC)}$ and MDS showed that, at the MHC, the populations at the periphery of Europe (including Great Britain, Ireland, Israel, Italy, Sweden, Spain, Portugal, and Denmark) were significantly differentiated from the central and eastern populations (including Austria, Belgium, Croatia, France, Germany, Luxembourg, Poland, Serbia, and Switzerland), concordant with the patterns previously inferred from the same panel of microsatellites using a similar dataset (Frantz et al., 2014). Third, the northern, geographically isolated populations (i.e., Great Britain, Ireland, and Sweden) showed signatures of a founder effect at MHC genes, e.g., reduced MHC diversity and similar MHC allele compositions between Britain and Ireland. The reduced MHC diversity in these isolated populations is consistent with previous conclusions that genetic drift outweighs balancing selection when populations become small and isolated (Campos et al., 2006; Miller et al., 2010; Miller & Lambert, 2004; Strand et al., 2012; Zhai et al., 2017). Also, these populations were previously shown to have small population sizes due to successive founder effects during historical postglacial expansion processes and/or the genetic bottleneck effects during recent human activities (Bevanger & Lindström, 1995; Frantz et al., 2014; Pope et al., 2005; van der Zee et al., 1992). Despite significant population differentiation being detected between Britain and Ireland using $F_{ST(MHC)}$, the similarity observed in the MHC allele composition of these two populations is in line with the hypothesis that Britain could be a potential source for the colonization of the fauna in Ireland (Frantz et al., 2014; McDevitt et al., 2011). Furthermore, when all the above evidence is taken together, we can see that the global distribution of MHC diversity of *M. meles* in Europe is largely congruent with the outcomes of previous demographic processes, for example, *M. meles* expanding out of the glacial refugia in Iberia and southeast Europe toward northern and eastern regions (e.g., Sweden) during the post-glacial expansion (Frantz et al., 2014). Similar effects of demographic processes on large spatial MHC diversity can be found in species with

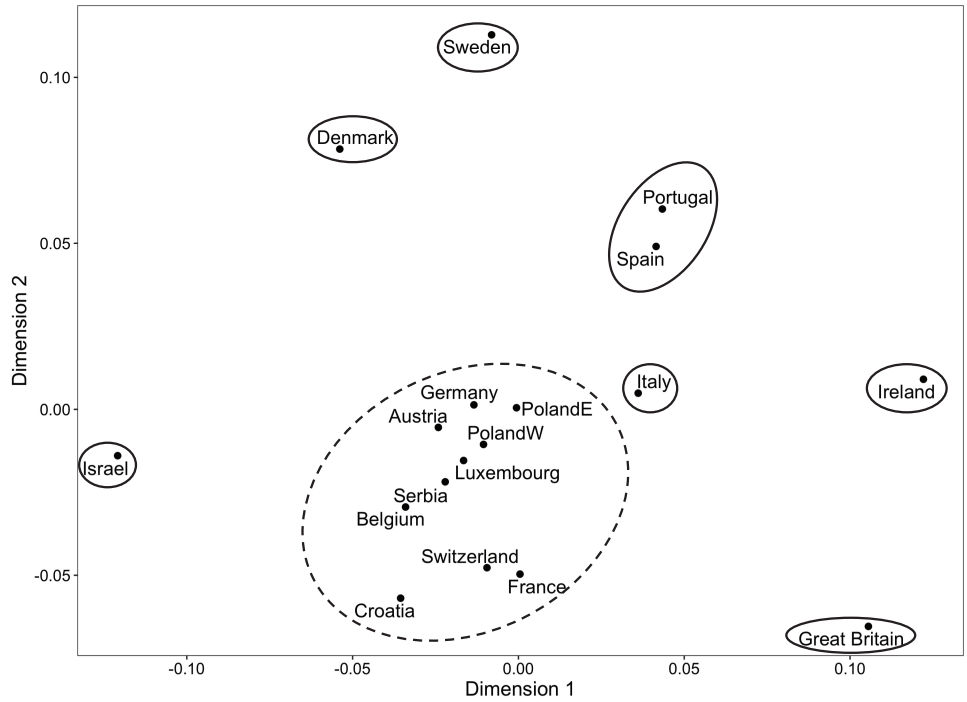

Figure 5. Multidimensional scaling (MDS) plot based on $F_{ST(MHC)}$ among 18 populations of *Meles meles* in Europe. Each dot represents a population. Populations in Denmark, Great Britain, Ireland, Israel, Italy, Portugal+Spain, and Sweden (black circles) were shown to be significantly differentiated from each other, and from the remaining populations, as indicated by pairwise $F_{ST(MHC)}$ (Table 4). No clear patterning was observed among the populations in Austria, Belgium, Croatia, France, Germany, Luxembourg, Poland (E), Poland (W), Serbia, and Switzerland (dashed circle). Stress score = 0.125. Stress score is a goodness-of-fit statistic for MDS, where the value should be between 0 to 1 with near zero indicating a better fit of the model.

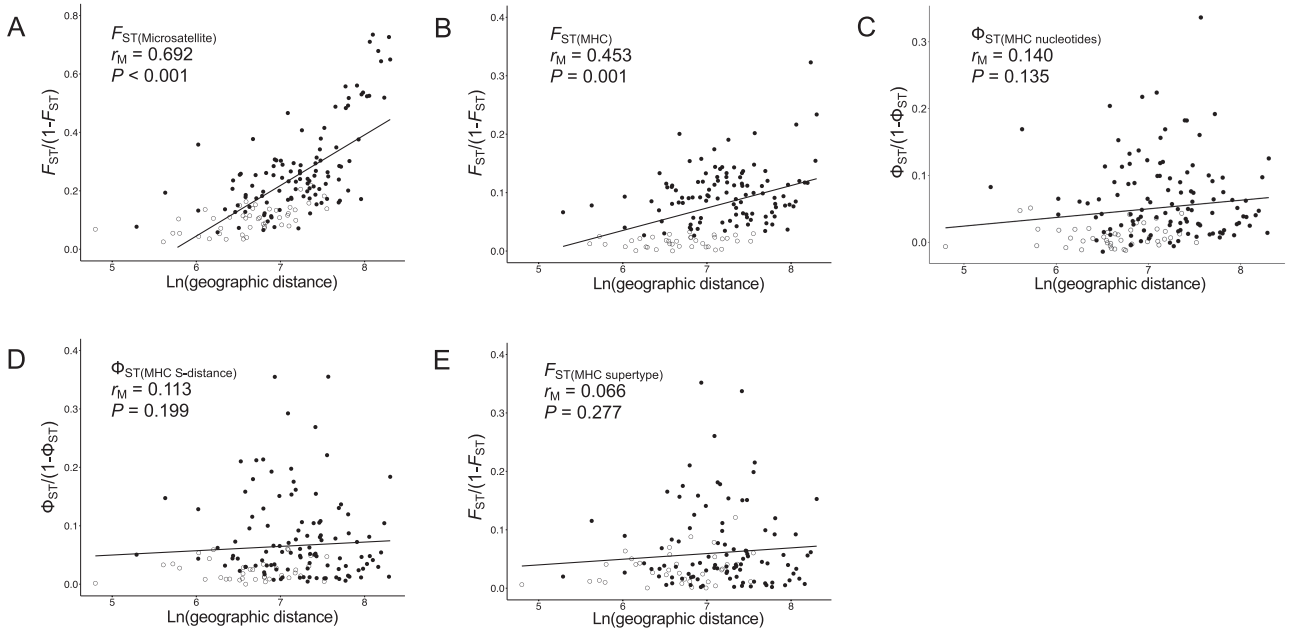

Figure 6. Isolation-by-distance at MHC class II DRB genes and microsatellite loci among 18 populations of *Meles meles*. Mantel tests were performed to test for the correlations between alternative pairwise genetic distances ($F_{ST}/(1-F_{ST})$) (A–E: $F_{ST(Microsatellite)}$, $F_{ST(MHC)}$, $\Phi_{ST(MHC\ nucleotides)}$, $\Phi_{ST(MHC\ S\text{-}distance)}$, and $F_{ST(MHC\ supertype)}$) and the logarithm of geographic distance (km). The empty circles indicate the comparisons among central, eastern and southeastern populations (Austria, Belgium, Croatia, France, Germany, Luxembourg, Poland, Serbia, and Switzerland), except Italy. The black solid circles indicate the remaining pairwise population comparisons. The black lines represent linear regression lines for all population comparisons. *p* values were calculated from 10,000 permutations. MHC, major histocompatibility complex.

limited dispersal ability or philopatric behavior (Babik et al., 2008; Berggren et al., 2005; Talarico et al., 2019; Zeisset & Beebee, 2014). The studied species have similar distribution ranges to *M. meles*.

Balancing selection maintains MHC functional diversity

Despite the evidence for neutral drift, support for balancing selection can still be found when the sequences of the MHC

alleles were taken into account. First, the nucleotide diversity of MHC did not correlate with heterozygosity or allelic richness of microsatellites, even for the isolated populations in Britain, Ireland, and Sweden, where balancing selection was expected to be weaker than genetic drift. This finding agrees with recent observations that balancing selection maintains MHC functional diversity, even under neutral drift (Heimeier et al., 2018; Sagonas et al., 2019). However, caution needs to be taken in interpreting the finding that most of the alleles retained in the bottlenecked populations (e.g., Britain, Ireland, and Sweden) belong to different allelic lineages, because this could simply be the result of random drift, as shown by a previous simulation study in bottlenecked populations (Ejsmond & Radwan, 2011). Second, the number of instances of significant pairwise population differentiation calculated using functional measures (e.g., $\Phi_{ST(MHC\ S\text{-}distance)}$ and $F_{ST(MHC\ supertype)}$) were almost half of that based on $F_{ST(MHC)}$, suggesting that most of the populations shared functionally equivalent sets of MHC alleles. The simple Mantel tests using distance-based matrix of MHC (Φ_{ST}) and supertype-based matrix (i.e., $F_{ST(MHC\ supertype)}$) further confirmed this pattern, with no geographic structuring detected for MHC at the functional level. The partial Mantel tests also supported the conclusion that the structuring observed in pairwise $F_{ST(MHC)}$ comparisons was indeed due to neutral differentiation. All this evidence suggests the presence of balancing selection in maintaining the functional diversity of MHC genes across the distribution range of *M. meles*. By contrast, if diversifying selection is the dominant force, we would expect different supertypes to be maintained in different populations, which could lead to IBD. Our results here are consistent with previous findings that balancing selection, rather than diversifying selection, maintains MHC polymorphism at a large spatial scale (Bichet et al., 2015; Malé et al., 2012; Sagonas et al., 2019).

The drive of balancing selection on MHC genes is generally thought to be pathogens (Spurgin & Richardson, 2010). Heterozygote advantage, and rare-allele advantage and/or fluctuating selection, were previously proposed as the selection mechanisms acting on MHC genes in *M. meles* (Sin et al., 2014). By comparing the population structure at MHC with that at microsatellites, we showed that the structuring at the MHC was less than for microsatellite loci, supporting the effect of heterozygote advantage and/or rare-allele advantage favoring evolutionarily old alleles that were previously common in the population (Figure 6A vs. Figure 6D and E). If a population is subject to heterozygote advantage, then within-population diversity can exceed between-population diversity (Spurgin & Richardson, 2010). Similarly, if selection maintains old alleles in multiple populations lacking novel alleles, then between-population diversity will be less than expected from neutral diversity (Spurgin & Richardson, 2010). Indeed, pronounced population structure at MHC was often detected in previous studies, and it was concluded that adaptation to the local pathogenic environments is the driving cause (Ekblom et al., 2007; Landry & Bernatchez, 2002; Miller et al., 2001; e.g., Cammen et al., 2011; Eizaguirre et al., 2012; Li et al., 2016). Fewer studies have demonstrated weaker population structure at MHC than at neutral loci (Sommer, 2003; van Oosterhout et al., 2006; e.g., Bichet et al., 2015; Del Real-Monroy & Ortega, 2017; Herdegen-Radwan et al., 2021; Rico et al., 2015), whereas some studies have concluded that there was no significant difference between MHC and neutral population structures (e.g., Biedrzycka & Radwan, 2008; Gutierrez-Espeleta et al., 2001; Spurgin & Richardson, 2010). This difference in findings among studies probably reflects a difference in underlying ecological factors (e.g., demographic history, habitat variation, life history, mating system, pathogens, etc.) that affect the strength of each selective force.

MHC-based mate choice is another mechanism that could affect MHC diversity (Bentkowski & Radwan, 2020; Eizaguirre et al., 2009; Slade et al., 2019). MHC-assortative mate choice was previously demonstrated in *M. meles* (Sin et al., 2015). Such a mate choice behavior could complicate the interpretation of our results. For example, MHC homozygosity is expected to increase via assortative mate choice while heterozygote advantage predicts an increase in heterozygosity (Slade et al., 2019; Spurgin & Richardson, 2010), although the major role of heterozygote advantage in shaping spatial MHC diversity is still controversial (Borghans et al., 2004; Ejsmond & Radwan, 2015; Radwan et al., 2020; Sutton et al., 2011). Indeed, the predicted outcome of MHC-assortative mate choice seems to be in line with our observation that very few private MHC alleles were identified in different populations, probably because assortative mate choice tends to prevent newly mutated allele from increasing in frequency if sexual selection is stronger than PMBS. However, as MHC-assortative mate choice was only demonstrated in a British population (Sin et al., 2015), how it might influence continent-wide pattern of MHC diversity in *M. meles* is still an open question (Havlíček et al., 2020; Kamiya et al., 2014; Winternitz et al., 2017).

Balancing selection on supertype hypothesis

Balancing selection on supertype hypothesis proposes that MHC alleles should be grouped into immunological supertypes (Herdegen-Radwan et al., 2021; Lighten et al., 2017), which will allow us to understand the evolution of MHC diversity at the functional level. In this study, we reveal a weaker population structuring of MHC supertypes than that based on individual MHC alleles. This finding of high supertype sharing despite high population differentiation at the MHC allelic level has also been reported in other species (Lillie et al., 2015). There is growing evidence showing that balancing selection is better detected at the functional level (Sepil et al., 2013; Trachtenberg et al., 2003; Vlček et al., 2016), but such an idea was not consolidated into a well-defined hypothesis previously, and Lighten et al. (2017) proposed that spatial signal of selection on MHC is stronger at MHC supertype level. Currently, the balancing selection on supertype hypothesis still received little empirical and theoretical support (Ejsmond et al., 2018; Herdegen-Radwan et al., 2021; Lighten et al., 2017). However, caution is required when analyzing the effect of MHC supertypes, since the grouping of alleles into supertypes may generate statistical artifacts, e.g., lowering of the population differentiation. Future studies should follow earlier work (Herdegen-Radwan et al., 2021; Lighten et al., 2017) in including a simulation of random allele grouping into supertypes to test the effect of grouping. In addition, analyses based on functional distance between MHC alleles should also be performed to directly test the divergent allele advantage (Wakeland et al., 1990) in addition to the balancing selection on supertype hypothesis. We therefore highlight the importance of analyzing MHC diversity at both allelic and functional levels, and of using various statistical approaches that can

capture different aspects of diversity within MHC genes. These methods allow us to distinguish the effects between selection and neutral processes on MHC genes more specifically. The idea that balancing selection mainly operates at MHC supertypes has so far received little attention (Sepil et al., 2013; Trachtenberg et al., 2003; Vlček et al., 2016; but see Ejsmond et al., 2018; Lighten et al., 2017). This might explain why different patterns were observed in earlier studies of MHC diversity.

Overall, our study demonstrates how both balancing selection and genetic drift simultaneously could affect the evolution of MHC genes (Babik et al., 2008; Sagonas et al., 2019; Vlček et al., 2016). Neutral genetic drift mainly acts on allele identity while selection targets the functional diversity of MHC genes. The current spatial pattern of MHC diversity in *M. meles* was also influenced by previous demographic processes. Our continent-wide sampling and analytic approach to study both MHC genes and microsatellites further provides evidence for heterozygote advantage and/or rare-allele advantage. The complexity of MHC evolution demonstrated here can only be revealed when we conducted the analysis at different levels (i.e., allelic vs. functional). We advocate future studies of MHC evolution in wild populations to use a similar analytical approach to differentiate neutral processes and selection, and the selection mechanisms at different levels of the MHC diversity.

Supplementary material

Supplementary material is available online at *Evolution* (https://academic.oup.com/evolut/qpac014)

Data availability

MHC DNA allele sequences are deposited in GenBank (accession numbers: MW911247-MW911254). MHC and microsatellite genotyping data and sampling location of each individual are available via Dryad (doi.org/10.5061/dryad.z34tmpghz).

Author contributions

S.Y.W.S. and A.C.F. designed research; A.C.F., S.Y.W.S., and E. G. collected field samples; S.Y.W.S. and A.C.F. conducted the laboratory work in T.B.'s laboratory; D.K.L. and S.Y.W.S. analyzed data; D.K.L. wrote the paper and all authors contributed to revised versions.

Funding statement

None declared.

Conflict of interest: The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgments

We would like to acknowledge the following people for sample contribution: E.O. Ågren, A. Balestrieri, H. Bocklisch, V. Boyaval, C. Braun, P. Breyne, H. Brøseth, D. Ćirović, M. Colyn, J. Davison, E. Do Linh San, M. Elmeros, M. Ghirardi, K. Hindelang, J. Kochan, B.I. Kostka, R. Kowalczyk, J. Lanszki, A.B. Madsen, G. Molina-Vacas, F. Müller, R. Myslajek, R. Oliveira, V. Racheva, E. Randi, L. Remonti, A. Ruiz-Gonzalez, M. Schaul, L. Schley, N. Šprem, M. Statham, N. Stier, A. Stubbe, M. Stubbe, F. Suchentrunk, A. Trua, P. Turner, K. Van Den Berge, K. Weidemann, K. Zalewski. Sampling in Serbia was supported by the Serbian Ministry of Education, Science and Technological Development. We thank Jacek Radwan for helpful comments.

References

Abduriyim, S., Nishita, Y., Kosintsev, P. A., Raichev, E., Väinölä, R., Kryukov, A. P., Abramov, A. V., Kaneko, Y., & Masuda, R. (2017). Diversity and evolution of MHC class II DRB gene in the Eurasian badger genus Meles (Mammalia: Mustelidae). *Biological Journal of the Linnean Society*, 122, 258–273.

Abduriyim, S., Nishita, Y., Kosintsev, P. A., Raichev, E., Väinölä R., Kryukov, A. P., Abramov, A. V., Kaneko, Y., & Masuda, R. (2019). Evolution of MHC class I genes in Eurasian badgers, genus *Meles* (Carnivora, Mustelidae). *Heredity*, 122, 205–218.

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

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., & Wayne, R. K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 3490.

Alcaide, M., Edwards, S. V., Negro, J. J., Serrano, D., & Tella, J. L. (2008). Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Molecular Ecology*, 17, 2652–2665.

Andrew, R. L., Ostevik, K. L., Ebert, D. P., & Rieseberg, L. H. (2012). Adaptation with gene flow across the landscape in a dune sunflower. *Molecular Ecology*, 21, 2078–2091.

Annavi, G., Newman, C., Dugdale, H. L., Buesching, C. D., Sin, Y. W., Burke, T., & Macdonald, D. W. (2014). Neighbouring-group composition and within-group relatedness drive extra-group paternity rate in the European badger (*Meles meles*). *Journal of Evolutionary Biology*, 27, 2191–2203.

Apanius, V., Penn, D., Slev, P. R., Ruff, L. R., & Potts, W. K. (1997). The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology*, 17, 179–224.

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

Balkenhol, N., Waits, L. P., & Dezzani, R. J. (2009). Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography*, 32, 818–830.

Bentkowski, P., & Radwan, J. (2020). Mating preferences can drive expansion or contraction of major histocompatibility complex gene family. *Proceedings of the Royal Society B: Biological Sciences*, 287, 20192706.

Berggren, K. T., Ellegren, H., Hewitt, G. M., & Seddon, J. M. (2005). Understanding the phylogeographic patterns of European hedgehogs, *Erinaceus concolor* and *E. europaeus* using the MHC. *Heredity*, 95, 84–90.

Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, 16, 363–377.

Bevanger, K., & Lindström E. R. (1995). Distributional history of the European badger *Meles meles* in Scandinavia during the 20th century. *Annales Zoologici Fennici*, 32, 5–9.

Bichet, C., Moodley, Y., Penn, D. J., Sorci, G., & Garnier, S. (2015). Genetic structure in insular and mainland populations of house sparrows (*Passer domesticus*) and their hemosporidian parasites. *Ecology and Evolution*, 5, 1639–1652.

Biedrzycka, A., & Radwan, J. (2008). Population fragmentation and major histocompatibility complex variation in the spotted suslik, *Spermophilus suslicus*. *Molecular Ecology*, 17, 4801–4811.

Bilham, K., Sin, Y. W., Newman, C., Buesching, C. D., & Macdonald, D. W. (2013). An example of life history antecedence in the European badger (*Meles meles*): Rapid development of juvenile antioxidant capacity, from plasma vitamin E analogue. *Ethology Ecology & Evolution*, 25, 330–350.

Borghans, J. A. M., Beltman, J. B., & De Boer, R. J. (2004). MHC polymorphism under host-pathogen coevolution. *Immunogenetics*, 55, 732–739.

Buesching, C.D., Tinnesand, H.V., Sin, Y.W., Rosell, F., Burke, T., & Macdonald, D.W. 2016. Coding of group odor in the subcaudal gland secretion of the European badger meles meles: Chemical Composition and Pouch Microbiota. In: B. Schulte, T. Goodwin, & M. Ferkin (Eds.) *Chemical Signals in Vertebrates*, Vol. 13. Springer, Cham.

Cammen, K., Hoffman, J. I., Knapp, L. A., Harwood, J., & Amos, W. (2011). Geographic variation of the major histocompatibility complex in Eastern Atlantic grey seals (*Halichoerus grypus*). *Molecular Ecology*, 20, 740–752.

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

Carpenter, P. J., Dawson, D. A., Greig, C., Parham, A., Cheeseman, C. L., & Burke, T. (2003). Isolation of 39 polymorphic microsatellite loci and the development of a fluorescently labelled marker set for the Eurasian badger (*Meles meles*) (Carnivora: Mustelidae). *Molecular Ecology Notes*, 3, 610–615.

Corner, L. A. L., Murphy, D., & Gormley, E. (2011). *Mycobacterium bovis* infection in the Eurasian Badger (*Meles meles*): The disease, pathogenesis, epidemiology and control. *Journal of Comparative Pathology*, 144, 1–24.

Del Real-Monroy, M., & Ortega, J. (2017). Spatial distribution of microsatellite and MHC-DRB exon 2 gene variability in the Jamaican fruit bat (*Artibeus jamaicensis*) in Mexico. *Mammalian Biology*, 84, 1–11.

Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, 256, 50–52.

Domingo-Roura, X., Macdonald, D. W., Roy, M. S., Marmi, J., Terradas, J., Woodroffe, R., Burke, T., & Wayne, R. K. (2003). Confirmation of low genetic diversity and multiple breeding females in a social group of Eurasian badgers from microsatellite and field data. *Molecular Ecology*, 12, 533–539.

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

Edwards, S. V., & Hedrick, P. W. (1998). Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends in Ecology & Evolution*, 13, 305–311.

Eimes, J. A., Bollmer, J. L., Whittingham, L. A., Johnson, J. A., Van Oosterhout, C., & Dunn, P. O. (2011). Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. *Journal of Evolutionary Biology*, 24, 1847–1856.

Eizaguirre, C., Lenz, T. L., Kalbe, M., & Milinski, M. (2012). Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, 15, 723–731.

Eizaguirre, C., Yeates, S. E., Lenz, T. L., Kalbe, M., & Milinski, M. (2009). MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology*, 18, 3316–3329.

Ejsmond, M. J., Phillips, K. P., Babik, W., & Radwan, J. (2018). The role of MHC supertypes in promoting trans-species polymorphism remains an open question. *Nature Communications*, 9, 4362.

Ejsmond, M. J., & Radwan, J. (2011). MHC diversity in bottlenecked populations: A simulation model. *Conservation Genetics*, 12, 129–137.

Ejsmond, M. J., & Radwan, J. (2015). Red queen processes drive positive selection on major histocompatibility complex (MHC) genes. *PLoS Computational Biology*, 11, e1004627.

Ekblom, R., Saether, S. A., Jacobsson, P., Fiske, P., Sahlman, T., Grahn, M., Kålås J. A., & Höglund J. (2007). Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology*, 16, 1439–1451.

Erickson, K. (2010). The jukes-cantor model of molecular evolution. *PRIMUS*, 20, 438–445.

Evans, M. L., Neff, B. D., & Heath, D. D. (2010). MHC genetic structure and divergence across populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Heredity*, 104, 449–459.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.

Frantz, A. C., McDevitt, A. D., Pope, L. C., Kochan, J., Davison, J., Clements, C. F., Elmeros, M., Molina-Vacas, G., Ruiz-Gonzalez, A., Balestrieri, A., Berge, K. V. D., Breyne, P., Do Linh San, E.E., Ågren, E. O.E. O., Suchentrunk, F., Schley, L., Kowalczyk, R., Kostka, B. I., Ćirović D.D., Sprem, N., … Burke, T. (2014). Revisiting the phylogeography and demography of European badgers (Meles meles) based on broad sampling, multiple markers and simulations. *Heredity*, 113, 443–453.

Fumagalli, M. (2013). Assessing the effect of sequencing depth and sample size in population genetics inferences. *PLoS One*, 8, e79667.

Gillingham, M. A. F., Béchet A., Courtiol, A., Rendón-Martos M., Amat, J. A., Samraoui, B., Onmuş O., Sommer, S., & Cézilly F. (2017). Very high MHC Class IIB diversity without spatial differentiation in the mediterranean population of greater Flamingos. *BMC Evolutionary Biology*, 17, 56.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, 86, 485–486.

Goudet, J., Raymond, M., Meeüs T. de, & Rousset, F. (1996). Testing differentiation in diploid populations. *Genetics*, 144, 1933.

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. *Science*, 185, 862–864.

Griffiths, H. I., & Thomas, D. H. (1993). The status of the Badger *Meles meles* (L., 1758) (Carnivora, Mustelidae) in Europe. *Mammal Review*, 23, 17–58.

Guerrero, J., Byrne, A. W., Lavery, J., Presho, E., Kelly, G., Courcier, E. A., O'Keeffe, J., Fogarty, U., O'Meara, D. B., Ensing, D., McCormick, C., Biek, R., Skuce, R. A., & Allen, A. R. (2018). The population and landscape genetics of the European badger (*Meles meles*) in Ireland. *Ecology and Evolution*, 8, 10233–10246.

Gutierrez-Espeleta, G. A., Hedrick, P. W., Kalinowski, S. T., Garrigan, D., & Boyce, W. M. (2001). Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation?. *Heredity*, 86, 439–450.

Havlíček, J., Winternitz, J., & Roberts, S. C. (2020). Major histocompatibility complex-associated odour preferences and human mate choice: Near and far horizons. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375, 20190260.

Haynes, W. 2013. Benjamini–Hochberg method. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of systems biology* (pp. 78–78). Springer New York.

Heimeier, D., Alexander, A., Hamner, R. M., Pichler, F., & Baker, C. S. (2018). The influence of selection on MHC DQA and DQB haplotypes in the endemic New Zealand Hector's and Māui dolphins. *Journal of Heredity*, 109, 744–756.

Herdegen, M., Babik, W., & Radwan, J. (2014). Selective pressures on MHC class II genes in the guppy (*Poecilia reticulata*) as inferred by hierarchical analysis of population structure. *Journal of Evolutionary Biology*, 27, 2347–2359.

Herdegen-Radwan, M., Phillips, K. P., Babik, W., Mohammed, R. S., & Radwan, J. (2021). Balancing selection versus allele and supertype turnover in MHC class II genes in guppies. *Heredity*, 126, 548–560.

Hill, A. V. S. 1991. *HLA associations with malaria in Africa: Some implications for MHC evolution*. Springer Berlin Heidelberg.

Hughes, A. L., & Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335, 167–170.

Jeffery, K. J., & Bangham, C. R. (2000). Do infectious diseases drive MHC diversity?. *Microbes and Infection*, 2, 1335–1341.

Jukes, T. H. and C. R. Cantor. 1969. CHAPTER 24 - Evolution of protein molecules. In H. N. Munro (Ed.), *Mammalian protein metabolism* (pp. 21–132). Academic Press.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A., & Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: The role of diversity and dissimilarity. *Molecular Ecology*, 23, 5151–5163.

Kanno, Y., Vokoun, J. C., & Letcher, B. H. (2011). Fine-scale population structure and riverscape genetics of brook trout (*Salvelinus fontinalis*) distributed continuously along headwater channel networks. *Molecular Ecology*, 20, 3711–3729.

Kosakovsky Pond, S. L., & Frost, S. D. W. (2005). Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, 22, 1208–1222.

Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2019). HyPhy 2.5—A customizable platform for evolutionary hypothesis testing using phylogenies. *Molecular Biology and Evolution*, 37, 295–299.

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution*, 23, 1891–1901.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.

Landry, C., & Bernatchez, L. (2002). Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, 10, 2525–2539.

Legendre, P., & Fortin, M. -J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, 10, 831–844.

Li, J., Shen, H., Wang, H., Zhao, M., Luo, Z., & Wu, H. (2016). Diversifying selection is the dominant factor affecting the geographical variation of MHC class II genes in the Omei tree frog. *Journal of Zoology*, 300, 197–204.

Lichstein, J. W. (2007). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology*, 188, 117–131.

Lighten, J., Oosterhout, C. van, Paterson, I. G., McMullan, M., & Bentzen, P. (2014). Ultra-deep Illumina sequencing accurately identifies MHC class IIb alleles and provides evidence for copy number variation in the guppy (*Poecilia reticulata*). *Molecular Ecology Resources*, 14, 753–767.

Lighten, J., Papadopulos, A. S. T., Mohammed, R. S., Ward, B. J., Paterson, I. G., Baillie, L., Bradbury, I. R., Hendry, A. P., Bentzen, P., & Oosterhout, C. van (2017). Evolutionary genetics of immunological supertypes reveals two faces of the Red Queen. *Nature Communications*, 8, 1294.

Lillie, M., Grueber, C. E., Sutton, J. T., Howitt, R., Bishop, P. J., Gleeson, D., & Belov, K. (2015). Selection on MHC class II supertypes in the New Zealand endemic Hochstetter's frog. *BMC Evolutionary Biology*, 15, 63.

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

Loiseau, C., Richard, M., Garnier, S., Chastel, O., Julliard, R., Zoorob, R., & Sorci, G. (2009). Diversifying selection on MHC class I in the house sparrow (*Passer domesticus*). *Molecular Ecology*, 18, 1331–1340.

Lynch, M. (1990). The similarity index and DNA fingerprinting. *Molecular Biology and Evolution*, 7, 478–484.

Lynch, M. (1991). Analysis of population genetic structure by DNA fingerprinting. In T. Burke, G. Dolf, A. J. Jeffreys, & R. Wolff (Eds.), *DNA fingerprinting: Approaches and applications* (pp. 113–126). Birkhäuser Basel.

Macdonald, D. W., Newman, C., Dean, J., Buesching, C. D., & Johnson, P. J. (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.

Malé, P. G., Martin, J. -F., Galan, M., Deffontaine, V., Bryja, J., Cosson, J. -F., Michaux, J., & Charbonnel, N. (2012). Discongruence of Mhc and cytochrome b phylogeographical patterns in *Myodes glareolus* (Rodentia: Cricetidae). *Biological Journal of the Linnean Society*, 105, 881–899.

Martinsohn, J. T., Sousa, A. B., Guethlein, L. A., & Howard, J. C. (1999). The gene conversion hypothesis of MHC evolution: A review. *Immunogenetics*, 50, 168–200.

Matsumura, M., Fremont, D. H., Peterson, P. A., & Wilson, I. A. (1992). Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science*, 257, 927–934.

McDevitt, A. D., Vega, R., Rambau, R. V., Yannic, G., Herman, J. S., Hayden, T. J., & Searle, J. B. (2011). Colonization of Ireland: Revisiting 'the pygmy shrew syndrome' using mitochondrial, Y chromosomal and microsatellite markers. *Heredity*, 107, 548–557.

Miller, H. C., Allendorf, F., & Daugherty, C. H. (2010). Genetic diversity and differentiation at MHC genes in island populations of tuatara (*Sphenodon* spp.). *Molecular Ecology*, 19, 3894–3908.

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16, 1215–1215.

Miller, K. M., Kaukinen, K. H., Beacham, T. D., & Withler, R. E. (2001). Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. *Genetica*, 111, 237–257.

Miller, H. C., & Lambert, D. M. (2004). Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Molecular Ecology*, 13, 3709–3721.

Minias, P., Pikus, E., & Anderwald, D. (2019). Allelic diversity and selection at the MHC class I and class II in a bottlenecked bird of prey, the White-tailed Eagle. *BMC Evolutionary Biology*, 19, 2.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., & Scheffler, K. (2013). FUBAR: A fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolution*, 30, 1196–1205.

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8, e1002764.

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

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1–10.

Peakall, R. O. D., & Smouse, P. E. (2006). genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)*, 28, 2537–2539.

Penn, D. J., & Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *The American Naturalist*, 153, 145–164.

Peterman, W. E., & Pope, N. S. (2021). The use and misuse of regression models in landscape genetic analyses. *Molecular Ecology*, 30, 37–47.

Piertney, S. B., & Oliver, M. K. (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity*, 96, 7–21.

Pope, L. C., Domingo-Roura, X., Erven, K., & Burke, T. (2005). Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology*, 15, 371–386.

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253–1256.

Pruett, C. L., & Winker, K. (2008). The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. *Journal of Avian Biology*, 39, 252–256.

Radwan, J., Babik, W., Kaufman, J., Lenz, T. L., & Winternitz, J. (2020). Advances in the evolutionary understanding of MHC polymorphism. *Trends in Genetics*, 36, 298–311.

Raymond, M., & Rousset, F. (1995). An exact test for population differentiation. *Evolution*, 49, 1280–1283.

Reche, P. A., & Reinherz, E. L. (2003). Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. *Journal of Molecular Biology*, 331, 623–641.

Rico, Y., Morris-Pocock, J., Zigouris, J., Nocera, J. J., & Kyle, C. J. (2015). Lack of spatial immunogenetic structure among Wolverine (*Gulo gulo*) populations suggestive of broad scale balancing selection. *PLoS One*, 10, e0140170.

Rogers, L. M., Cheeseman, C. L., Mallinson, P. J., & Clifton-Hadley, R. (1997). The demography of a high-density badger (*Meles meles*) population in the west of England. *Journal of Zoology*, 242, 705–728.

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

Roper, T. J. 2010. *Badger*. Collins.

Rosalino, L. M., Macdonald, D. W., & Santos-Reis, M. (2004). Spatial structure and land-cover use in a low-density Mediterranean population of Eurasian badgers. *Canadian Journal of Zoology*, 82, 1493–1502.

Sagonas, K., Runemark, A., Antoniou, A., Lymberakis, P., Pafilis, P., Valakos, E. D., Poulakakis, N., & Hansson, B. (2019). Selection, drift, and introgression shape MHC polymorphism in lizards. *Heredity*, 122, 468–484.

Sandberg, M., Eriksson, L., Jonsson, J., Sjöström M., & Wold, S. (1998). New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *Journal of Medicinal Chemistry*, 41, 2481–2491.

Sepil, I., Lachish, S., Hinks, A. E., & Sheldon, B. C. (2013). Mhc supertypes confer both qualitative and quantitative resistance to avian malaria infections in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20130134.

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2018). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, 18, 55–67.

Siddle, H. V., Marzec, J., Cheng, Y., Jones, M., & Belov, K. (2010). MHC gene copy number variation in Tasmanian devils: Implications for the spread of a contagious cancer. *Proceedings of the Royal Society B: Biological Sciences*, 277, 2001–2006.

Sidney, J., Grey, H. M., Kubo, R. T., & Sette, A. (1996). Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunology Today*, 17, 261–266.

Sin, Y. W., Annavi, G., Dugdale, H. L., Newman, C., Burke, T., & Macdonald, D. W. (2014). Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (*Meles meles*). *Molecular Ecology*, 23, 5072–5088.

Sin, Y. W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D. W., & Dugdale, H. L. (2015). MHC class II-assortative mate choice in European badgers (*Meles meles*). *Molecular Ecology*, 24, 3138–3150.

Sin, Y. W., Dugdale, H. L., Newman, C., Macdonald, D. W., & Burke, T. (2012a). Evolution of MHC class I genes in the European badger *(Meles meles*). *Ecology and Evolution*, 2, 1644–1662.

Sin, Y. W., Dugdale, H. L., Newman, C., Macdonald, D. W., & Burke, T. (2012b). MHC class II genes in the European badger (*Meles meles*): Characterization, patterns of variation, and transcription analysis. *Immunogenetics*, 64, 313–327.

Sin, Y. W., Buesching, C. D., Burke, T., & Macdonald, D. W. (2012c). Molecular characterization of the microbial communities in the subcaudal gland secretion of the European badger (*Meles meles*). *FEMS Microbiology Ecology*, 81, 648–659.

Sin, Y. W., Newman, C., Dugdale, H. L., Buesching, C., Mannarelli, M. -E., Annavi, G., Burke, T., & Macdonald, D. W. (2016). No compensatory relationship between the innate and adaptive immune system in wild-living European badgers. *PLoS One*, 11, e0163773.

Slade, J. W. G., Watson, M. J., & MacDougall-Shackleton, E. A. (2019). "Balancing" balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection. *Ecology and Evolution*, 9, 5146–5157.

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

Spurgin, L. G., & Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*, 277, 979–988.

Strand, T. M., Segelbacher, G., Quintela, M., Xiao, L., Axelsson, T., & Höglund J. (2012). Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? *Ecology and Evolution*, 2, 341–353.

Subramanian, S. (2016). The effects of sample size on population genomic analyses – implications for the tests of neutrality. *BMC Genomics*, 17, 123.

Sutton, J. T., Nakagawa, S., Robertson, B. C., & Jamieson, I. G. (2011). Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology*, 20, 4408–4420.

Swain, S. L. (1983). T cell subsets and the recognition of MHC class. *Immunological Reviews*, 74, 129–142.

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

Talarico, L., Babik, W., Marta, S., & Mattoccia, M. (2019). Genetic drift shaped MHC IIB diversity of an endangered anuran species within the Italian glacial refugium. *Journal of Zoology*, 307, 61–70.

Trachtenberg, E., Korber, B., Sollars, C., Kepler, T. B., Hraber, P. T., Hayes, E., Funkhouser, R., Fugate, M., Theiler, J., Hsu, Y. S., Kunstman, K., Wu, S., Phair, J., Erlich, H., & Wolinsky, S. (2003). Advantage of rare HLA supertype in HIV disease progression. *Nature Medicine*, 9, 928–935.

van der Zee, F. F., Wiertz, J., Ter Braak, C. J. F., van Apeldoorn, R. C., & Vink, J. (1992). Landscape change as a possible cause of the badger *Meles meles* L. decline in The Netherlands. *Biological Conservation*, 61, 17–22.

van Oosterhout, C., Joyce, D. A., Cummings, S. M., Blais, J., Barson, N. J., Ramnarine, I. W., Mohammed, R. S., Persad, N., & Cable, J. (2006). Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). *Evolution*, 60, 2562–2574.

Vlček, J., Hoeck, P. E. A., Keller, L. F., Wayhart, J. P., Dolinová I., & Štefka J. (2016). Balancing selection and genetic drift create unusual patterns of MHCIIβ variation in Galápagos mockingbirds. *Molecular Ecology*, 25, 4757–4772.

Wakeland, E. K., Boehme, S., She, J. X., Lu, C. C., McIndoe, R. A., Cheng, I., Ye, Y., & Potts, W. K. (1990). Ancestral polymorphisms of MHC class II genes: Divergent allele advantage. *Immunologic Research*, 9, 115–122.

Walsh, P. S., Metzger, D. A., & Higuchi, R. (2013). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, 10, 506–513.

Wilson, G., Harris, S., McLaren, G. (1998). *Changes in the British badger population, 1988–1997*. London: People's Trust for Endangered Species.

Winternitz, J., Abbate, J. L., Huchard, E., Havlíček J., & Garamszegi, L. Z. (2017). Patterns of MHC-dependent mate selection in humans and nonhuman primates: A meta-analysis. *Molecular Ecology*, 26, 668–688.

Wong, A. T. C., Lam, D. K., Poon, E. S. K., Chan, D. T. C., & Sin, S. Y. W. (2022). Intra-specific copy number variation of MHC class II genes in the Siamese fighting fish. *Immunogenetics*, 74, 327–346.

Woodroffe, R., Macdonald, D. W., & da Silva, J. (1995). Dispersal and philopatry in the European badger, *Meles meles*. *Journal of Zoology*, 237, 227–239.

Zeisset, I., & Beebee, T. J. C. (2014). Drift rather than selection dominates MHC class II allelic diversity patterns at the biogeographical range scale in natterjack toads *Bufo calamita*. *PLoS One*, 9, e100176.

Zhai, T., Yang, H. -Q., Zhang, R. -C., Fang, L. -M., Zhong, G. -H., & Fang, S. -G. (2017). Effects of population bottleneck and balancing selection on the Chinese alligator are revealed by locus-specific characterization of MHC genes. *Scientific Reports*, 7, 5549.

Zhang, P., Huang, K., Zhang, B., Dunn, D. W., Chen, D., Li, F., Qi, X., Guo, S., & Li, B. (2018). High polymorphism in MHC-DRB genes in golden snub-nosed monkeys reveals balancing selection in small, isolated populations. *BMC Evolutionary Biology*, 18, 29.