MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 4071–4086 doi: 10.1111/mec.12349

MHC-disassortative mate choice and inbreeding avoidance in a solitary primate

ELISE HUCHARD,*†[1] ALICE BANIEL,*[1] SUSANNE SCHLIEHE-DIECKS*† and PETER M. KAPPELER*†

Behavioral Ecology and Sociobiology Unit, German Primate Center, Kellnerweg 4, Göttingen, Germany, †Courant Research Centre Evolution of Social Behaviour, University of Göttingen, Kellnerweg 6, Göttingen, Germany

Abstract

Sexual selection theory suggests that choice for partners carrying dissimilar genes at the major histocompatibility complex (MHC) may play a role in maintaining genetic variation in animal populations by limiting inbreeding or improving the immunity of future offspring. However, it is often difficult to establish whether the observed MHC dissimilarity among mates drives mate choice or represents a by-product of inbreeding avoidance based on MHC-independent cues. Here, we used 454-sequencing and a 10-year study of wild grey mouse lemurs (*Microcebus murinus*), small, solitary primates from western Madagascar, to compare the relative importance on the mate choice of two MHC class II genes, DRB and DQB, that are equally variable but display contrasting patterns of selection at the molecular level, with DRB under stronger diversifying selection. We further assessed the effect of the genetic relatedness and of the spatial distance among candidate mates on the detection of MHC-dependent mate choice. Our results reveal inbreeding avoidance, along with disassortative mate choice at DRB, but not at DQB. DRB-disassortative mate choice remains detectable after excluding all related dyads (characterized by a relatedness coefficient $r > 0$), but varies slightly with the spatial distance among candidate mates. These findings suggest that the observed deviations from random mate choice at MHC are driven by functionally important MHC genes (like DRB) rather than passively resulting from inbreeding avoidance and further emphasize the need for taking into account the spatial and genetic structure of the population in correlative tests of MHC-dependent mate choice.

Keywords: grey mouse lemur, inbreeding avoidance, major histocompatibility complex, mate choice, sexual selection, spatial genetic structure

Received 29 November 2012; revision received 3 April 2013; accepted 9 April 2013

Introduction

Increasing evidence suggests that genetic compatibility (or complementarity) between sexual partners may play a role in maintaining or regulating genetic variation in animal populations (Jennions 1997; Tregenza & Wedell 2000; Mays & Hill 2004). Choosing partners with maximal or optimal genetic dissimilarity may respectively help to avoid inbreeding or excessive outbreeding, and increase offspring diversity at key functional loci, such as the major histocompatibility complex (MHC). The MHC is a large cluster of highly polymorphic genes coding for the molecules involved in the adaptive (as opposed to innate) immune response. Due to its codominant expression and its function in the immune response, individuals with a high MHC diversity may be at an advantage in a population facing heterogeneous pathogenic pressures (i.e. heterozygote advantage: Doherty & Zinkernagel 1975). Alternatively, particular MHC genes may be advantageous if they confer protection against dominant pathogenic

Correspondence: Elise Huchard, Department of Zoology, University of Cambridge, Downing Street, CB2 3EJ, Cambridge UK, Fax: +441223336676; E-mail: ehuchard@gmail.com

[1]These authors contributed equally to this work.

pressures in a given environment (Bodmer 1972; Apanius *et al.* 1997). In the latter context, a given allele is only expected to be advantageous temporarily and locally, because parasites might regularly evolve resistance to the most common host alleles (i.e. frequency-dependent selection) and parasite communities typically vary in space and time (i.e. fluctuating selection) (Hedrick & Kim 1999; Spurgin & Richardson 2010; Eizaguirre *et al.* 2012a,b).

A growing number of studies on free-ranging populations have suggested a variety of targets for MHC-biased female choice. Studies encompassing a range of organisms from fish (e.g. Landry *et al.* 2001; Consuegra & de Leaniz 2008) to mammals (Schwensow *et al.* 2008a,b; Setchell *et al.* 2010), including reptiles (Olsson *et al.* 2003) and birds (Freeman-Gallant *et al.* 2003; Strandh *et al.* 2012), have reported mate choice for MHC-dissimilar partners. In contrast, other research has suggested that some animals may prefer optimally rather than maximally dissimilar partners (Reusch *et al.* 2001; Milinski *et al.* 2005; Bonneaud *et al.* 2006; Forsberg *et al.* 2007; Baratti *et al.* 2012), whereas further studies even report a reproductive disadvantage for MHC-dissimilar males (Bonneaud *et al.* 2006; Bos *et al.* 2009; Yeates *et al.* 2009). According to yet other studies, partner preferences target mates possessing maximal MHC diversity (Richardson *et al.* 2005; Bonneaud *et al.* 2006; Schwensow *et al.* 2008a) or specific MHC genotypes (von Schantz *et al.* 1996; Ditchkoff *et al.* 2001; Ekblom *et al.* 2004; Eizaguirre *et al.* 2009). Finally, contrary to all these studies, no MHC-dependent mate choice of any sort could be detected in some other species and populations (Paterson & Pemberton 1997; Westerdahl 2004; Huchard *et al.* 2010), suggesting that it may not be a ubiquitous vertebrate reproductive strategy.

The diversity of mate-choice strategies may mirror the diversity of parasite-driven selective pressures affecting MHC genes (Roberts 2009). Specifically, advantages for good vs. diverse genotypes are likely to set a trade-off between choices for good vs. compatible genes, the optimal value of which is likely to vary in predictable ways. For instance, it has been suggested that population genetic structure and diversity might influence mate-choice strategies: females might choose dissimilar partners in inbred populations and similar partners in outbred populations, as suggested by contrasting patterns among human populations (Chaix *et al.* 2008; but see: Derti *et al.* 2010; Laurent & Chaix 2012). Individuals may also flexibly adjust their mating decisions, as suggested by experimental studies in fish showing that females with low MHC diversity select dissimilar partners, whereas females with high MHC diversity select similar partners (Milinski *et al.* 2005). Alternatively, this diversity of observed outcomes might reflect a lack of methodological and operational standardization among studies in a still relatively young field of study.

Progress in MHC evolutionary ecology has long been hampered by challenges associated with MHC genotyping, largely driven by the complex structure of the MHC region (Bernatchez & Landry 2003; Piertney & Oliver 2006; Babik 2010). In addition to intense allelic polymorphism ('allelic diversity') for many loci, the extent of allelic divergence among alleles ('sequence diversity') often makes it difficult to design primers, especially when the genomic architecture of the region is unknown. It is then rarely possible to design densely spaced internal SNPs or microsatellites throughout large regions – as may be done in humans. A widespread alternative is to sequence short fragments located in highly variable region(s) containing antigen-binding sites, characterized by amino acids located in the antigen-binding pocket of the MHC molecule and directly interacting with antigens (reviewed by Huchard *et al.* 2010). But this design has several drawbacks. First, it is questionable whether similarity between individuals over short fragments accurately reflects wider MHC diversity. Second, individuals commonly differ in the number and organization of MHC loci, which makes it difficult to separate multiple coamplified sequences (often differing by a single nucleotide), as well as to assign these sequences to loci (Bernatchez & Landry 2003; Babik 2010). The recent application of next-generation sequencing in MHC studies has proven successful at overcoming some of the challenges linked with MHC genotyping (e.g. Babik *et al.* 2009; Wegner 2009; Galan *et al.* 2010; Zagalska-Neubauer *et al.* 2010; Huchard *et al.* 2012a; Sepil *et al.* 2012), allowing in-depth investigations of factors affecting MHC-dependent mate choice and its detection.

An important problem in field studies observing deviations from random mate choice at MHC genes is to establish whether MHC dissimilarity among mates does drive mate choice or may simply represent a by-product of inbreeding avoidance based on alternative, MHC-independent, cues. Inbreeding avoidance, regardless of its underlying mechanism, may generate a greater genome-wide dissimilarity among parents than among randomly matched partners, which may then be detected in any variable genomic region including the MHC (even if the MHC does not play any role in mate choice). However, discriminating both effects using correlative designs is difficult and requires large sample sizes in order to test the existence of MHC-disassortative mate choice among unrelated dyads exclusively. The need for large sample sizes limits the use of observational data which clearly identify who mated and who did not. Alternative approaches based on DNA-based parentage

assignments have been developed to test whether the MHC dissimilarity among parents exceeds those of random partners (e.g. Manning *et al.* 1992; Reusch *et al.* 2001; Bonneaud *et al.* 2006). Such an approach is very sensitive to the definition of the null hypothesis, that is, to the criteria used to define a 'random partner'. Specifically, failure to account for the spatial distance among potential mates when designing a pool of random partners may generate a bias when comparing MHC dissimilarity to genome-wide dissimilarity among mates, because the spatial population structure of functional and neutral genes may often differ (Westerdahl *et al.* 2004; Oliver *et al.* 2009; Eizaguirre *et al.* 2012a). However, few empirical efforts have been made to evaluate the consequences of this potential bias for downstream analyses.

Here, we take advantage of a data set generated from a long-term field study of grey mouse lemurs (*Microcebus murinus*) (Huchard *et al.* 2012a) to investigate the implications of different analytical procedures in tests of MHC-biased mate choice. The MHC class II diversity of the grey mouse lemur, a small, nocturnal and solitary forager with a polygynandrous mating system, has been well studied (Schad *et al.* 2004, 2005; Schwensow *et al.* 2010; Averdam *et al.* 2011; Huchard *et al.* 2012a). MHC class II DRB is highly polymorphic (Schad *et al.* 2004; Schwensow *et al.* 2008a) and influences nematode resistance (Schad *et al.* 2005; Schwensow *et al.* 2010). Recent work has further reported mate choice for MHC dissimilarity (Schwensow *et al.* 2008a). We extended the sample size and the MHC region examined using 454-sequencing to (i) compare the relative importance of two MHC markers, DRB vs. DQB, that are equally variable but display contrasting patterns of selection at the molecular level (with DRB under stronger diversifying selection than DQB) (Huchard *et al.* 2012a); (ii) assess the effect of the genetic relatedness and (iii) the spatial distance between candidate mates on the detection of MHC-dependent mate choice.

Methods

Study species and population

Grey mouse lemurs are seasonal breeders, and each female is sexually receptive for one or two (if the first cycle was not conceptive) night(s) per year, during which she mates polyandrously. Males roam outside their usual home range during the mating season and can successively compete for and mate with several females. As a result, the proportion of litters with mixed paternity is high (Eberle & Kappeler 2002, 2004a,b).

The study population is located within a 12500-ha forestry concession of the Centre National de Formation, d'Etude et de Recherche en Environnement et Foresterie (CNFEREF) in Kirindy Forest, located about 60 km northeast of Morondava in western Madagascar (Kappeler & Fichtel 2012). Since 1994, DNA samples and population parameters were collected during monthly routine captures using about 160 traps at a time in an area of about 9 ha within a 60-ha grid system with small footpaths every 25 metres, locally known as CS7. Additional captures in surrounding areas were conducted once or twice a year and covered an additional area of about 18 ha. Each capture lasted three consecutive nights to generate estimates of trapping success. Traps were set at dusk, collected at dawn, and animals were carried to camp during their first day of capture to be processed and were released at dusk the same day, at their trapping location. Animals processed for the first time were briefly anesthetized by applying 0.01 ml Ketanest 100 subdermally, marked with subdermal transponders, and a small ear biopsy (~2–3 mm^2) was taken, along with a set of morphometric measures and biological samples. Age was known for most individuals born in the study population (because they were first captured as juveniles), and individuals trapped for the first time as adults were estimated to be born the year before, as most males disperse during their first year of life (Schliehe-Diecks *et al.* 2012). Occasional protocols including behavioural observations relied on identifying every known individual in the study area with visible marks (shaved rings on the tail) and suggested that 80–90% of the population was known (Eberle & Kappeler 2004b). Complementary information regarding the study population and the trapping procedures are available elsewhere (Eberle & Kappeler 2004a,b). Individuals included in this study belong to the cohorts present between 2000 and 2010, and two of these cohorts (2000–2001) had already been used in a previous analysis of MHC-dependent mate choice (Schwensow *et al.* 2008a). To estimate the spatial location of individual home ranges, we calculated the arithmetic mean of all trapping events recorded for a given individual. Based on the obtained average UTM-coordinates, the closest trail intersection of the grid system was assigned as being the centre of home range of a specific individual. This position was then used to calculate distances between home-range centres of individuals.

Molecular methods

DNA was isolated from ear biopsies following standard protocols (Qiagen QIAmp DNA Mini Kit no. 51306). Maternities and paternities have been established for 1094 individuals of this population, using microsatellite markers for 13 loci with an average of 21 alleles per

locus (SD = 9.48, range = 13–43). Both paternal and maternal identities have been established for 333 offspring (S. Schliehe-Diecks & P. M. Kappeler,unpublished data). Methods used to genotype individuals and to establish maternities and paternities have been described in detail elsewhere (Eberle & Kappeler 2002, 2004b).

MHC genotyping was performed for 643 individuals at two MHC class II loci, DRB and DQB, using pyrosequencing – specifically a Roche GS Junior sequencer. These loci were chosen because a recent genomic analysis of MHC class II organization confirmed that DRB and DQB loci are the most variable parts of the class II region (Averdam *et al.* 2011). In contrast to most primates examined so far, these loci appear to be nonduplicated in grey mouse lemurs, thus showing a maximum of two alleles per locus (Schad *et al.* 2004; Averdam *et al.* 2011; Huchard *et al.* 2012a) (but see Schwensow *et al.* 2008a). Methods regarding MHC amplification, cloning and sequencing, as well as the various quality control steps performed to ensure and validate sequencing quality have been described elsewhere (Huchard *et al.* 2012a) and are only briefly summarized here.

Both loci were amplified using published primers (DQB: Averdam *et al.* 2011; DRB: Schad *et al.* 2004). In addition to these template-specific primers, the 454-sequencing system requires the addition of a 25-mer 5′-portion whose sequence is designed according to manufacturer's instructions (Roche Applied Science, Mannheim, Germany); a 10-bp tag identifying an individual can be added between the sequencing key and the template-specific sequence. To genotype 96 individuals, we thus used 10 different individual tags in the forward primer, combined with 10 different tags in the reverse primer. Equimolar amounts of 96 individual amplicons were pooled for a given locus (i.e. DRB or DQB) after purification. Here and later, we define as 'amplicon' the pool of reads amplified from a particular individual during a single PCR. DRB and DQB libraries were then pooled and sequenced together in both directions on a 454 (Roche®) Junior System. We conducted 7 successive sequencing runs in total, yielding over 800 000 sequences with an average of 110 000 sequences per run. A total of 768 DRB amplicons, including 685 individual samples plus 76 duplicates and 7 triplicates, and 672 DQB amplicons were initially included in the sequencing process. A total of 654 individuals, including 53 duplicates and four triplicates were successfully genotyped for DRB, and 643 individuals for DQB. Genotyping duplicates involved duplicating the PCR (with a different tag combination), library preparation and sequencing. Our resulting coverage for each gene ranged from 0 to 7.042 reads per amplicon, with a median of 292. Genotyping was considered successful for a given amplicon if the coverage was higher than 18 reads (see Huchard *et al.* 2012a for a justification of this threshold).

Although 454-sequencing represents a highly cost-effective method for large-scale MHC genotyping, it is also highly error-prone, generating a high rate of artefactual alleles (AA) that may be difficult to discriminate from true alleles (TA) (Babik *et al.* 2009; Wegner 2009; Galan *et al.* 2010; Zagalska-Neubauer *et al.* 2010). As a result, a stringent stepwise variant validation procedure was applied. A total of 61 (including 23 new; Genbank Accession nos for 11 new alleles retrieved in more than one individual: HE801954-HE801964) and 60 (including 53 new; Genbank Accession nos for 40 new alleles retrieved for more than one individual: HE801914-HE801953) alleles were detected at DRB and DQB genes, respectively. Note that our stepwise validation procedure retained alleles that were possessed by only one individual if these (i) displayed relatively high per-amplicon frequency (i.e. >0.05) and (ii) were not identified as sequencing artefacts through comparison with more common alleles from the same amplicon (Huchard *et al.* 2012a). By contrast, traditional approaches in MHC typing require amplification from at least two independent PCRs to ensure reliability of a new allele description. This procedure would inevitably lead to the elimination of true alleles, thus introducing a risk of biasing further population genetic analyses based on this data set through the occurrence of null alleles. Consequently, sequences that passed our stepwise validation procedure but had been retrieved from only one individual (12 DRB and 13 DQB alleles) were included in the analyses, but not submitted to Genbank. Their nucleotide sequences have been published elsewhere (Huchard *et al.* 2012a).

The quality of our final genotypes was evaluated through three independent methods (Huchard *et al.* 2012a): (i) genotyping five individuals by both cloning (see Supporting Information for methodological details regarding cloning procedure) and 454-sequencing; (ii) running duplicates; and (iii) comparing parent–offspring dyads; each displayed very high accuracy, with a perfect match between cloning and 454-sequencing results, a repeatability close to 100% between duplicate samples and an error rate estimated to be less than 2% using parent–offspring dyads (see Supporting Information for further details on the evaluation of genotyping quality). Finally, the estimated frequency of null alleles was very low (DRB: 0%, DQB: 0.6%) and neither DQB nor DRB showed a significant heterozygote excess.

Genetic analyses

Definition of MHC haplotypes. In order to investigate whether the extent of the region genotyped may affect

the detection of MHC-biased mate choice, analyses were carried out with each locus separately, as well as at the haplotype level, because DRB and DQB are in intense linkage disequilibrium (Huchard *et al.* 2012a). As a result, the cosegregation of DRB and DQB sequences among parent–offspring triads allowed the identification of 71 MHC class II haplotypes, each containing one DRB and one DQB sequence.

Linkage disequilibrium between microsatellites and MHC. Estimation of linkage disequilibrium between each microsatellite and each MHC locus was tested using a likelihood ratio test where the likelihood of the sample evaluated under the null hypothesis of no association between loci (linkage equilibrium) is compared to the likelihood of the sample when association is allowed (Slatkin & Excoffier 1996). The significance of the procedure is found by computing the null distribution of this ratio under the hypothesis of linkage equilibrium using a permutation procedure (here with 100.000 permutations) implemented in the software ARLEQUIN 3.5.1.3 (Excoffier & Lischer 2010).

Calculation of relatedness coefficients within dyads. Microsatellite-based relatedness coefficients were calculated using the triadic IBD relatedness estimates (Wang 2007) calculated with the software COANCESTRY v 1.0.0.0 (Wang 2011) for all individuals captured between the years 1999 and 2010 using 100 reference individuals and 100 bootstraps.

Calculation of MHC dissimilarity within dyads. The genetic distance between 2 MHC alleles (amino acid sequences) or between 2 MHC haplotypes was computed using three different indices: (i) the number of sequences (or haplotypes) that are different from the partner considered, that is, 0–2 in this species; (ii) the pairwise number of amino acid differences among sequences (or haplotypes) as a measure of sequence divergence; and (iii) an index estimating the functional distance among alleles or haplotypes, based on the physico-chemical properties of the amino acids forming the antigen-binding pocket (antigen-binding sites: ABS). This latter functional index is based on the assumption that the physico-chemical properties of the ABS determine which antigens the molecule can bind and ultimately, which pathogens an individual can recognize (Doytchinova & Flower 2005). This index was calculated in two steps. In a first step, ABS occurring in the DRB and DQB exon 2 of the grey mouse lemur were identified as positively selected sites (PSS) using tests of positive selection. In the absence of information regarding the location of the actual ABS, this step relies on the assumption that ABS are under strong diversifying selection. Identification of PSS in this sample has been described in detail elsewhere (Huchard *et al.* 2012a), yielding 11 PSS in DRB and 3 PSS in DQB. In a second step, each PSS from aligned DRB (or DQB) sequences was characterized by five physico-chemical descriptor variables that describe important physico-chemical properties of the amino acid within the binding site (Sandberg *et al.* 1998) : z1 (hydrophobicity), z2 (steric bulk), z3 (polarity), z4 and z5 (electronic effects) (Schwensow *et al.* 2007; Huchard *et al.* 2008). The functional distance between two alleles was then calculated as the Euclidian distance separating the two vectors of length 5 * Number of PSS describing the two alleles and does not take into account the non-PSS. For haplotypes, the measures of distance were summed across DRB and DQB. Between 2 individuals carrying 2 alleles each, the genetic distance can be computed between four possible combinations of nucleotide sequences. For each dyad, we defined 6 measures of MHC dissimilarity summarized in Table 1. We further defined 2 measures of within-individual MHC diversity (by measuring genetic distances between both alleles of the same individual) per haplotype or per locus (Table 1).

Randomization tests

Identification of mate-choice events. Our tests of female mate choice are based on actual parentage data rather than on mating observations, which are difficult to conduct in grey mouse lemurs (Eberle & Kappeler 2004a) and which provide sample sizes that are unlikely to detect moderate or small effect sizes. As such, the MHC dissimilarity (and genetic relatedness) among parents of a given offspring was compared to the MHC dissimilarity (and genetic relatedness) between the mother and a set of randomly matched males.

Randomization procedure. For each female, we compared the MHC constitution of her successful partner (i.e. the father(s) of her offspring) to candidate mates sharing a territory with the individual, using two-tailed permutation tests. We ran the analysis with and without one-year-old females because they undergo their first mating season and are physically weaker than older females and are therefore presumably less able to choose their mate (Gomez *et al.* 2012; Huchard *et al.* 2012b). By default, a male was included in the pool of candidate males for a given female if he fulfilled two criteria: (i) he was present in the last capture preceding the relevant mating season (September trapping) or during any subsequent capture session (ensuring that he was alive during the mating season); and (ii) the average distance between potential partners is lower than the maximal

Table 1 Summary of the genetic measures used

Hypothesis tested	Genetic measure	Description	Part of the sequence
Choice for MHC dissimilarity	N.seq.dif	Total number of sequences of the partner that are different from the focal individual	All
	N.hap.dif	Number of haplotypes of the partner that are different from the focal individual	All
	Dist.aa.mean	Mean number of amino acid differences between the 4 possible combinations of haplotypes (sequences) of a pair*	All
	Dist.aa.min	Number of amino acid differences between the most similar haplotypes (sequences) of a pair	All
	Dist.func.mean	Mean functional distance† between the 4 possible haplotypes (sequences) combinations of a pair	PSS
	Dist.func.min	Functional distance between the most similar haplotypes (sequences) of a pair	PSS
Choice for MHC diversity	Dist.aa.div	Number of amino acids that differ between the 2 haplotypes (sequences) possessed by the mate	All
	Dist.func.div	Functional distance between the 2 haplotypes (sequences) possessed by the mate	PSS

*Each individual carries two haplotypes (for instance, female has F1, F2 and male has M1, M2) which results in 4 possible combinations of haplotypes (F1M1, F1M2, F2M1 and F2M2).

†Functional measures take into account only the amino acids which were found to be positively selected and are assumed to be functionally important (11 PSS for DRB and 3 PSS for DQB) (see Huchard *et al.* 2012a; for more details on tests of positive selection).

spatial distance recorded between the two parents of an offspring using parentage data from our population (n = 115 parent pairs from the central area (CS7), mean distance among parents ± SD = 96.36 ± 61.58 m, range: 0–319 m, Fig. 1). The distribution of the various genetic measures was generated under the null hypothesis of random mate choice by randomly allocating, for each mating event, a given male as successful or unsuccessful. This procedure was repeated 20.000 times. In order to test whether females maximize genetic dissimilarity with their mate, the exact two-tailed P-value was computed as the proportion of simulations displaying a greater mean than the observed mean value for successful partners plus the proportion of cases displaying a lower mean than the symmetrical (relative to the simulated mean) of the observed mean value.

Mixed-paternity litters are frequent in grey mouse lemurs (Eberle & Kappeler 2004b); in this simulated sample, a conception giving birth to a litter was treated as one mating event ('mating night' is listed in Table S1, Supporting information), and two (or more) males were considered successful for that particular event in the case of a mixed-paternity litter, whereas only one was considered as successful in a single-paternity litter. Consequently, each male was either 'successful' or 'unsuccessful' concerning a given litter, and we did not take into account the fact that some males sired more than one offspring per litter. In this statistical design, mating nights are considered statistically independent, even if some mating nights involve the same female (in subsequent years) and some males competed for several females over a mating season. There is consequently some pseudoreplication in these randomization tests, which we considered acceptable for two reasons. First, it is unlikely to generate false positives with respect to our study question. Indeed, a male possessing a rare genotype will be, on average, dissimilar to most females. This may increase the risk of false positives if it makes multiple appearances in the data set. However, permutation tests limit this risk because a given individual will appear as many times in the random as in the observed distribution. Second, the distribution of the response variable is strongly 0-inflated, with an average of 1 successful male for up to 30 unsuccessful males per mating event (see Table S1, Supporting Information), which would decrease the fit and power of alternative approaches based on binomial mixed-effect models. We ran three consecutive series of tests described below, respectively aimed at testing whether (i) the MHC locus examined; (ii) the genetic relatedness; and (iii) the spatial distance among females and random mates may affect detection of MHC-biased mate choice.

Testing the influence of the extent of the MHC region sequenced. We examined the consistency of results obtained from permutation tests based on a variable extent of the MHC class II region: namely MHC-DRB exon 2 vs. MHC-DQB exon 2 vs. MHC class II haplotypes (including MHC-DRB exon 2 + MHC-DQB exon 2).

Testing the influence of the relatedness coefficients among mates. We examined the consistency of results obtained from permutation tests based on a set of dyads with variable coefficients of genetic relatedness: $r < 0.25$, <0.125, <0.0625, <0.03125, $r = 0$, in an attempt to test

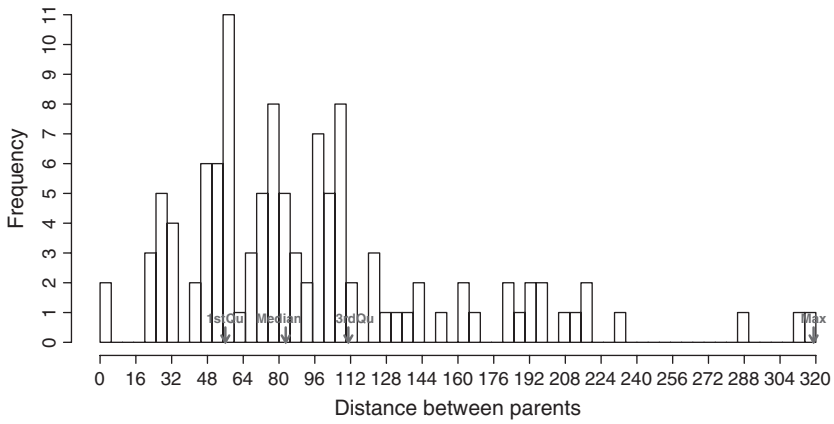

Fig. 1 Distribution of spatial distance among parental dyads in the study population, including data from 115 pairs from the central area. Grey arrows indicate the median and quartiles.

whether MHC-biased mate choice is still detectable after discarding the most related dyads.

Testing the influence of the spatial distance among females and random mates. The probability of mating of a random dyad from the population strongly depends on the spatial distance separating the home ranges of both partners. We thus examined the consistency of tests for which the set of candidate mates was built based on different thresholds regarding the average spatial distance between potential partners: (i) the first quartile of this distribution (distance among partners <56 m); (ii) the median (distance among partners <83 m); (iii) the third quartile (distance among partners <111 m); and (iv) the maximum (distance among partners <319 m). This set of criteria ensures that only males within a realistic distance of the focal female are considered as candidate mates. We also ran a fifth model without any threshold, thus including any live male in the study area (central CS7 area) with the spatial distance among partners ranging from 0 to 1.044 m.

Sample sizes used for all tests are summarized in Table S1 (Supporting information). All calculations were carried out with R version 2.14.0 (R Development Core Team, 2011). Data used to carry out these tests have been deposited in DRYAD (doi:10.5061/dryad.647 h6).

Results

Patterns of genetic variability

We found high genetic variability in the study population, with a total of 61 DRB (163 bp) and 60 DQB (163 or 169 bp) alleles distributed within 71 haplotypes. Each nucleotide sequence had a unique amino acid sequence, and the absence of stop codons suggested that all nucleotide sequences could encode functional proteins. Allelic frequencies varied widely, from 0.1% to 21% (median 2%) for both DQB and DRB (see Huchard *et al.* 2012a for the full distribution of allelic frequencies). *Mimu*-DRB and *Mimu*-DQB sequences showed wide-ranging levels of divergence with 56 (34%) variable sites in the DRB sequences and 45 (27%) in the DQB sequences. There was an average of 14.6 ± 2.2 (DRB) and 15.2 ± 2.4 (DQB) nucleotide differences and of 10.6 ± 2.0 (DRB) and 10.4 ± 1.8 (DQB) amino acid differences between sequences. Individuals carried 1–2 alleles at each locus, and homozygosity was rare, with only 25 homozygous individuals for DRB and 20 for DQB. Among these, only 11 DRB-heterozygous males were included in the final mate-choice data set, precluding a test of mate choice for MHC heterozygosity. Finally, coefficients of genetic relatedness r between females and males varied from 0 to 0.71 in the data set (median: 0).

Impact of MHC locus on detecting MHC-dependent mate choice

The comparison between successful and random males living within the maximal spatial range recorded among parental dyads revealed that females older than 1 year favour mates with dissimilar MHC-DRB genes, in terms of both amino acid sequence and functional similarity (genetic measures: Dist.aa.min and Dist.func.-min, Table 2, Fig. 2). There was no effect of the number of alleles shared among partners or of male individual MHC diversity on female choice. However, MHC-disassortative mate choice was not observed at DQB genes nor at the haplotype level (See Tables 2 and S4, Supporting information). Furthermore, the mean relatedness coefficients among parental dyads was lower than those among random dyads (See Table 2 and Fig. 2), suggesting that females avoid mating with close relatives. However, including young females (in their first mating season) in the data set weakened effects of MHC-DRB dissimilarity and of relatedness on mate choice, suggesting that these females do not express mate choice (See Table 2, Fig. 2).

Impact of genetic relatedness among partners on detecting MHC-dependent mate choice

MHC-DRB and MHC-DQB dissimilarity among dyads correlated negatively, albeit weakly, with genetic relatedness indices (See Table S2, Supporting information). The two MHC genes were in linkage equilibrium with each microsatellite locus (See Table S3, Supporting information), suggesting that the relationship linking MHC dissimilarity to genetic relatedness was not driven by a physical linkage between MHC and some microsatellites. We discarded the most related dyads from the data set to investigate whether deviations from

Table 2 Results of the tests investigating choice for genetic dissimilarity or diversity involving variable markers (MHC-DRB, MHC-DQB and microsatellites) for a data set including only females older than one year and a data set including all females

		Females older than one year			All females		
Genetic marker	Index	Simulated mean [IC 95%]	Observed mean	*P*-value	Simulated mean [IC 95%]	Observed mean	*P*-value
MHC-DRB exon 2	N.seq.dif	1.839 [1.747; 1.920]	1.867	1.000	1.834 [1.758; 1.909]	1.838	1.000
	Dist.aa.mean	10.361 [9.973; 10.737]	10.623	0.179	10.283 [9.944; 10.621]	10.467	0.284
	Dist.aa.min	6.742 [5.987; 7.480]	7.493	*0.050*	6.661 [6.000; 7.303]	7.081	0.212
	Dist.func.mean	14.183 [13.714; 14.630]	14.525	0.142	14.156 [13.747; 14.552]	14.410	0.217
	Dist.func.min	9.976 [8.846; 11.038]	11.182	*0.031*	9.910 [8.938; 10.840]	10.652	0.129
	Dist.aa.div	10.518 [9.813; 11.200]	10.133	0.283	10.546 [9.929; 11.141]	10.343	0.515
	Dist.func.div	14.764 [13.889; 15.558]	14.569	0.648	14.677 [13.902; 15.403]	14.666	0.978
MHC-DQB exon 2	N.seq.dif	1.839 [1.747; 1.920]	1.840	1.000	1.834 [1.758; 1.909]	1.818	0.690
	Dist.aa.mean	10.596 [10.223; 10.957]	10.463	0.479	10.646 [10.311; 10.972]	10.414	0.165
	Dist.aa.min	6.904 [6.080; 7.693]	7.227	0.442	6.982 [6.263; 7.677]	7.030	0.898
	Dist.func.mean	7.249 [6.906; 7.593]	7.105	0.410	7.313 [7.013; 7.606]	7.131	0.224
	Dist.func.min	4.145 [3.567; 4.727]	4.388	0.415	4.178 [3.677; 4.673]	4.218	0.874
	Dist.aa.div	10.903 [10.187; 11.560]	11.093	0.592	10.886 [10.263; 11.475]	11.000	0.723
	Dist.func.div	7.360 [6.731; 7.977]	7.014	0.277	7.375 [6.817; 7.910]	7.293	0.771
Microsatellites	Relatedness	0.072 [0.053; 0.095]	0.049	*0.033*	0.074 [0.056; 0.094]	0.055	*0.050*

Those results are based on tests in which candidate males were ranging closer to the female than the maximal distance recorded between parental dyads. Significant *P*-values are indicated in italic. Results of analyses conducted at the haplotype level are presented in the Supporting Information, see Table S4 (Supporting information).

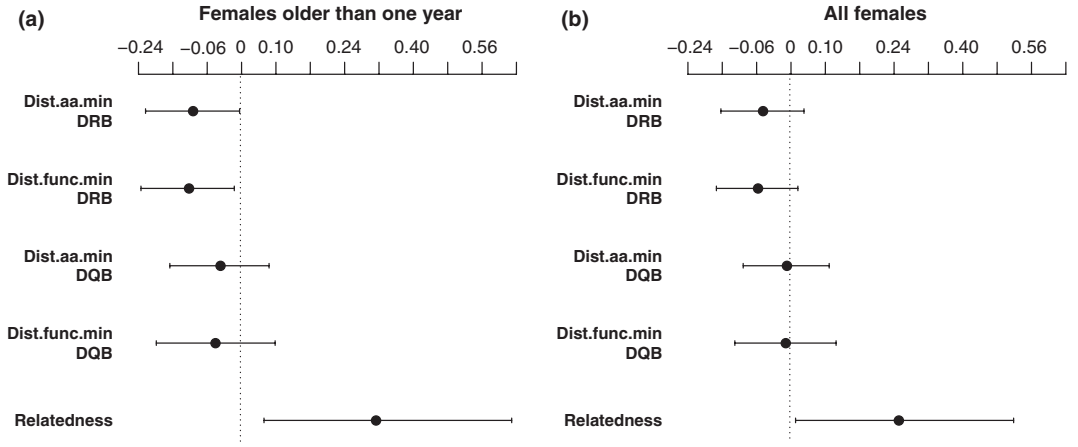

Fig. 2 Selected results of mate-choice tests based on different MHC genes and on data sets with a different age structure. Those results are based on tests in which candidate males were ranging closer to the female than the maximal distance recorded between parental dyads (i.e. set of randomly matched males named Max). Panel a: tests of mate choice for MHC dissimilarity and inbreeding avoidance when including only females older than one year. Panel b: tests of mate choice for MHC dissimilarity and inbreeding avoidance when including all females. For each genetic measure, the deviation between random and observed mate choice is represented by a point and its 95% confidence interval is indicated by a thin line, while the dashed line indicates 0. A variable is statistically significant ($P < 0.05$) when its confidence interval does not include 0. Coefficients of variation (standard error divided by the mean) are shown for the simulated distribution of each genetic measure - in order to scale the variance to the mean - so that effect sizes are comparable across measures.

random mate choice at DRB may represent a by-product of inbreeding avoidance (possibly mediated by MHC-independent effects). The MHC-DRB signal observed on mate choice in the sample of older females proved remarkably robust to this procedure and was still detectable when only those dyads for which $r = 0$ was included (See Table 3).

Impact of spatial distance between partners on detecting MHC-dependent mate choice

Using different criteria for including males as 'random' mates in the analyses affected tests of disassortative mate choice. Notably, including only close neighbours who ranged closer to the focal female than the median distance separating parents (computed from our long-term parentage records) produced significant results for different DRB dissimilarity measures (Dist.aa.min, Dist.func.min: Table S5, Supporting information and Fig. 3). Similarly, including all males who ranged as far as, or further than, the maximal distance recorded among confirmed parents produced significant results for the same DRB dissimilarity measures (Dist.aa.min, Dist.func.min: Table S5, Supporting information and Fig. 3). These results failed to reach significance when considering only those males living within a spatial distance that was comparable or only slightly greater than the median distance separating parents (although a strong trend was observed). In contrast, deviations from random mate choice for genetic relatedness were consistently observed throughout the range of criteria tested (See Table S5, Supporting information, Fig. 3e). In order

Table 3 Results of the tests investigating choice for genetic dissimilarity involving a variable threshold of maximal relatedness between dyads (e.g. the data set named '$r = 0$' included only dyads that are not related, data set '$r < 0.03125$' included only dyads whose relatedness is less than 0.03125, and data set 'all r' included all dyads regardless of their relatedness)

Genetic marker	Genetic measure	Conditions set on relatedness coefficients among partners (r)	Simulated mean [IC 95%]	Observed mean	P-value
MHC-DRB exon 2	Dist.aa.min	$r = 0$	6.854 [5.724; 7.931]	8.069	*0.031*
		$r < 0.03125$	6.784 [5.730; 7.784]	7.730	0.069
		$r < 0.0625$	6.935 [6.077; 7.769]	7.731	0.067
		$r < 0.1250$	6.953 [6.191; 7.691]	7.750	*0.038*
		$r < 0.2500$	6.844 [6.095; 7.554]	7.486	0.089
		All r	6.742 [5.987; 7.480]	7.493	*0.050*
	Dist.func.min	$r = 0$	10.274 [8.637; 11.79]	12.047	*0.027*
		$r < 0.03125$	10.117 [8.566; 11.506]	11.723	*0.033*
		$r < 0.0625$	10.254 [8.99; 11.436]	11.523	*0.042*
		$r < 0.1250$	10.207 [9.087; 11.279]	11.441	*0.027*
		$r < 0.2500$	10.133 [9.029; 11.174]	11.201	0.051
		All r	9.976 [8.846; 11.038]	11.182	*0.031*
MHC-DQB exon 2	Dist.aa.min	$r = 0$	6.802 [5.586; 7.966]	7.517	0.249
		$r < 0.03125$	7.000 [5.892; 8.027]	7.324	0.559
		$r < 0.0625$	7.133 [6.212; 8.038]	7.442	0.505
		$r < 0.1250$	7.155 [6.324; 7.926]	7.412	0.544
		$r < 0.2500$	6.996 [6.176; 7.770]	7.324	0.427
		All r	6.904 [6.080; 7.693]	7.227	0.442
	Dist.func.min	$r = 0$	4.089 [3.248; 4.909]	4.895	0.058
		$r < 0.03125$	4.206 [3.425; 4.969]	4.758	0.162
		$r < 0.0625$	4.274 [3.615; 4.932]	4.634	0.284
		$r < 0.1250$	4.305 [3.712; 4.893]	4.542	0.430
		$r < 0.2500$	4.210 [3.631; 4.780]	4.447	0.423
		All r	4.145 [3.567; 4.727]	4.388	0.415
Microsatellites	Relatedness	$r = 0$	NA	NA	NA
		$r < 0.03125$	0.003 [0.001;0.005]	0.004	0.171
		$r < 0.0625$	0.012 [0.007;0.017]	0.018	*0.041*
		$r < 0.1250$	0.033 [0.025;0.042]	0.035	0.726
		$r < 0.2500$	0.057[0.043;0.072]	0.045	0.102
		All r	0.072 [0.053; 0.095]	0.049	*0.033*

Tests presented below include only females older than one year and are based on tests in which candidate males were ranging closer to the female than the maximal distance recorded between parental dyads. Significant P-values are indicated in italics.

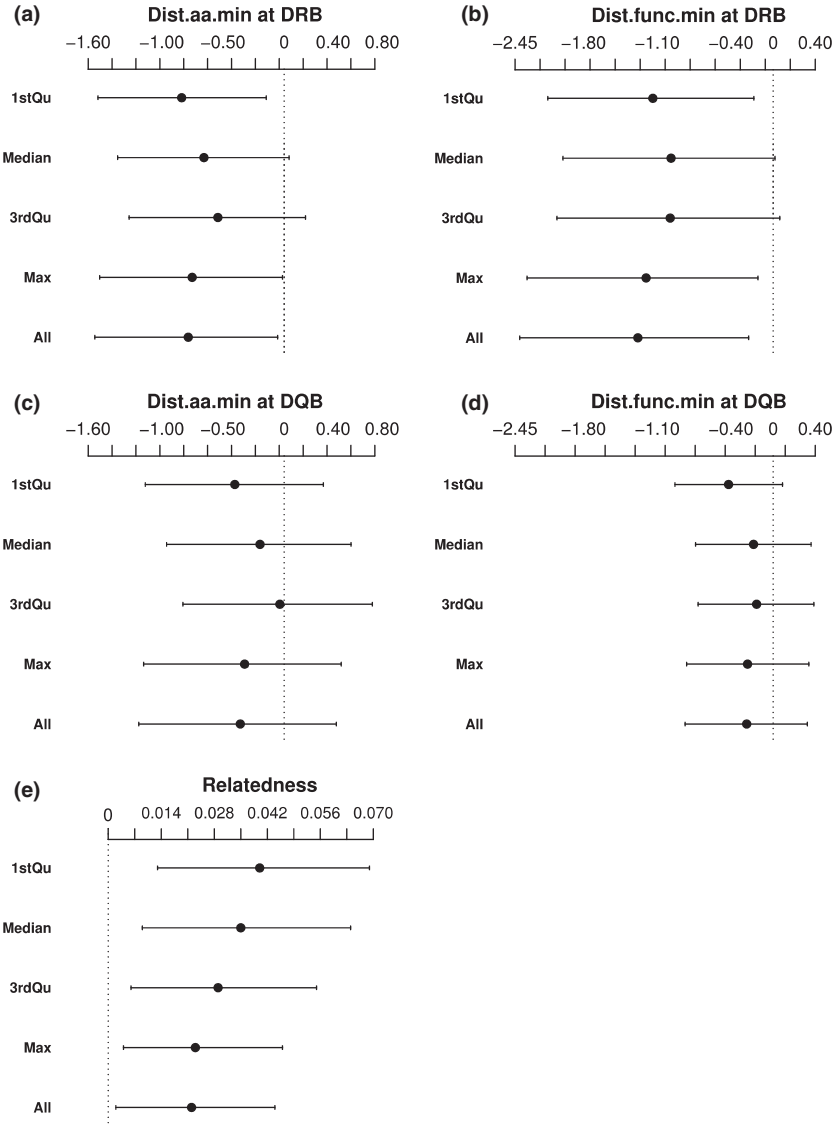

Fig. 3 Selected results of mate-choice tests using different sets of randomly matched males. Those results include only females older than one year and include tests of mate choice for MHC-DRB dissimilarity with genetic measures Dist.aa.min (Panel a.) and Dist.func.min (Panel b.), for MHC-DQB dissimilarity with genetic measures Dist.aa.min (Panel c.) and Dist.func.min (Panel d.) and a test of inbreeding avoidance (Panel e.). For each test, the set of candidate mates was built based on different thresholds for the average spatial distance between potential partners: (i) the first quartile of the distribution of the spatial distance separating actual parents shown in Fig. 1 (distance among partners <56 m); (ii) the median (distance among partners <83 m); (iii) the third quartile (distance among partner <111 m); and (iv) the maximum (distance among partners <319 m). This set of criteria ensures that only males ranging within a realistic distance of the focal female are considered as candidate mates. We also ran a fifth model without any threshold, thus including any live male in the study area (central CS7 area) with the spatial distance among partners ranging from 0 to 1.044 m. For each genetic measure, the deviation between random and observed mate choice is represented by a point and its 95% confidence interval is indicated by a thin line, while the dashed line indicates 0. A variable is statistically significant ($P < 0.05$) when its confidence interval does not include 0.

to investigate whether the impact of the set of candidate mates on the results was a consequence of variations in the number of candidate mates across data sets – which may affect the power of the analysis – or of variations in their genetic constitution, we ran additional simulations. These extra tests are presented in Supporting Information (See Table S6 and Fig. S1, Supporting information) and suggest that the influence of variable sets of candidate mates is independent of the number of candidate mates thus more likely to arise from the spatial genetic structure of the population.

Discussion

This study investigated MHC-biased female reproductive strategies in a natural population of grey mouse lemurs. Our results provide robust evidence for inbreeding avoidance along with weaker evidence for a role of MHC dissimilarity (but not diversity) in mate choice. Further investigations revealed that MHC-disassortative mate choice is only detectable at MHC-DRB (not at MHC-DQB) and independent from inbreeding avoidance. Moreover, the probability of detecting mate choice for neutral or functional genetic dissimilarity is affected by the age and spatial structure of the sample examined.

Comparison with previous work on the study population

Previous work on this population suggested that females choose males with a higher MHC-DRB dissimilarity (measured by the number of different amino acids among the two most similar sequences possessed by

the male and the female, as measured by 'Dist.aa.min' in Table 1 of this study) (Schwensow *et al.* 2008a). Likewise, our study detected effects of MHC-DRB dissimilarity on mate choice using an extended sample (because only offspring cohorts born in 2000–2001 were common across studies) and a fairly similar statistical design comparing genetic similarity among parents and randomly matched partners using randomization tests.

Several differences were nevertheless noticeable between the two studies. The effects of MHC-DRB on mate choice appeared weaker in our study. It may be due to stochastic variations in sample composition or to the fact that we used bilateral rather than unilateral tests. Methods used for MHC genotyping also differed across studies. We used 454-sequencing (Huchard *et al.* 2012a), whereas Schwensow *et al.* (2008a) used a more traditional method (Babik 2010), namely single-stranded conformation polymorphism ('SSCP' electrophoresis). Schwensow *et al.* (2008a) report interindividual variation in the number of DRB loci, which was not detected in this (Huchard *et al.* 2012a) and other studies (Schad *et al.* 2004; Averdam *et al.* 2011) and contribute to may explain why they detected female choice for MHC-diverse males while we did not. 454-sequencing typically generates a high frequency of false, artefactual alleles (Babik *et al.* 2009) which requires the establishment of stringent and complex quality control steps to filter out these false alleles (Huchard *et al.* 2012a). Although these procedures may generally increase the risk of filtering out true alleles, it is unlikely in this case because the quality of our genotyping was estimated through several independent methods, which all suggested very high accuracy (with ~100% repeatability across duplicates and less than 2% mismatches between parent and offspring genotypes) as well as a very low frequency of null alleles (<0.5%) (Huchard *et al.* 2012a). In addition, most (38 out of 43) DRB alleles identified by Schwensow *et al.* (2008a) were retrieved in our sample.

Contrasts between the two studies were particularly pronounced regarding the effect of relatedness on mate choice, which was only found to be significant in our study, despite performing similar genotyping methods. Two major explanations may account for this lack of consistency. First, Schwensow *et al.* (2008a) included every female, regardless of their ages, in their analyses, whereas our results suggest that including young females weakens the effect of relatedness on mate choice. Young females, who are smaller and less experienced, may be less able to express precopulatory mating preferences (Gomez *et al.* 2012; Huchard *et al.* 2012b). Alternatively, this age effect may reflect variation in inbreeding risk across a female's lifetime: it may be higher between fathers and daughters, who are not necessarily familiar with each other and live in close spatial proximity (because females do not disperse), than between mothers and sons, who are familiar with each other and may live further apart (because most males disperse before reproducing). Mortality is high in this population (Kraus *et al.* 2008), so that the coresidency between fathers and daughters probably rarely lasts much longer than the first mating season of a female. Inbreeding events, if they occur, may thus be concentrated during this period and obscure the signal of inbreeding avoidance detected after excluding these young females.

Second, Schwensow *et al.*'s methods do not report controlling for the spatial distance among heterosexual dyads to assign a given male to the set of random partners but rather included all males living in the study area, similar to our data set named 'All'. Our results suggest that the combination of these factors is sufficient to account for diverging results across studies. Indeed, our analyses show that the effect of relatedness on mate choice fails to reach significance when including all females and when controlling for the spatial distance among potential mates ($P > 0.05$, data not shown).

Effect of the MHC locus examined

Deviations from random mate choice observed at DRB are not observed at DQB or at the haplotype level. It is possible that the lower number of PSS at DQB ($n = 3$) than DRB ($n = 11$) decreases the power of the analyses based on indices of functional DQB dissimilarity. However, results based on indices of functional MHC dissimilarity do not differ qualitatively from results based on indices of MHC dissimilarity calculated throughout the whole fragment at both DRB and DQB. The fragments used may also be too short to provide an accurate reflection of dissimilarity across the MHC region: while we would expect more related individuals to display a higher MHC similarity, the covariance between MHC dissimilarity and genetic relatedness is low (about 10%: Table S2, Supporting information). Finally, female reproductive decisions may be affected by dissimilarity at DRB, but not at DQB because DRB is under stronger natural selection than DQB with more amino acids being positively selected in DRB than in DQB (Huchard *et al.* 2012a), and may play a greater role in the mechanisms driving MHC-based mate choice such as the production of odour cues mediating precopulatory choice or molecular interactions mediating postcopulatory choice. Although potential contrasts in the function of DRB vs. DQB molecules are unknown in mouse lemurs, their level of allelic diversity (number of alleles in the population) and of amino acid divergence among alleles are comparable, suggesting that DRB-specific effects do not simply result from a greater variability of DRB.

Overall, these results suggest that female choice specifically targets MHC loci that play an important role in disease resistance, in order to improve offspring immune performances.

Effects of excluding related dyads on MHC-dependent mate choice

The relatively weak effects of MHC-DRB on mate choice, along with evidence for inbreeding avoidance, may suggest that the signal observed at MHC-DRB represents a by-product of inbreeding avoidance based on MHC-independent cues – especially because both disappear when integrating young females in the data set. However, the DRB signal on mate choice proved remarkably robust to the exclusion of related dyads from the data set, suggesting that MHC-disassortative mate choice is independent of inbreeding avoidance. Using a sample of unrelated partners to discriminate MHC-disassortative mate choice from inbreeding avoidance may be more informative than simply assessing whether paternity outcomes are better predicted by MHC similarity (measured over short fragments) or by coefficients of relatedness (estimated over a limited number of microsatellites). Both measures are different in nature, likely to show heterogeneous variances, and expected to measure the actual relatedness of the considered dyad with a substantial inaccuracy (see Csillery *et al.* 2006 for a test of the relationship between marker-based and pedigree-based estimates of relatedness). The criteria used to discard related dyads nevertheless relied on these presumably inaccurate measures of genetic relatedness, so that the apparent independence between inbreeding avoidance and MHC-dependent mate choice should be interpreted cautiously. Overall, these results suggest that MHC-disassortative mate choice is independent of inbreeding avoidance in this population and reinforces the idea that it may aim at increasing offspring diversity at specific, functionally important loci, rather than throughout the whole genome.

Effect of a variable spatial distance among randomly matched partners

Our results further indicate that it is important to include as potential mates only those males who range close enough from a given female to have a chance of encountering her, especially in nongregarious species where the set of candidate mates available to the female is unknown, and in terrestrial species where population densities and average spatial distances among mates may frequently constrain mating opportunities. Here, the spatial group 'Max' including only those males who range closer than the maximal distance recorded among parents in our long-term data set is probably the most relevant to test mate choice, as it includes all those males who range close enough to mate with the focal female. It remains unclear why MHC dissimilarity among mates (relatively to dissimilarity among randomly matched partners) is higher when including only close or distant neighbours in the set of random partners, compared to cases including medium-distanced males. Behavioural data show that mating success depends on the spatial proximity among partners, with successful males ranging slightly closer than nonsuccessful males (mates: median = 71 m, range = 28–151 m; nonmates: median = 103 m, range: 41–235 m, $n = 20$ females) (Eberle & Kappeler 2004b), so that our spatial group '3rdQ' may discard a number of candidate males who could approach the female. The effect of the spatial distance on the detection of MHC-dissimilar mate choice may be due to random sampling fluctuations or variations in the spatial genetic structure.

MHC variation is likely to be either more or less spatially structured than neutral variation depending on the variation of its selective pressures (i.e. parasite communities) across space and time (Hedrick & Kim 1999; Garrigan & Hedrick 2003; Piertney & Oliver 2006; Spurgin & Richardson 2010). For example, MHC genes may show higher levels of spatial structure than neutral genes if parasite communities fluctuate substantially locally (on a scale dictated by the dispersal distance of the host species) and temporally (on a scale dictated by the generation time of the host species) (Westerdahl *et al.* 2004; Oliver *et al.* 2009; Evans *et al.* 2011; Eizaguirre *et al.* 2012a,b). As a result, levels of MHC dissimilarity between two random individuals may often vary with their spatial distance and this may affect results of tests of MHC-biased mate choice. For example, controlling for the proximity of potential mates when analysing snake mate-choice decisions in relation to MHC affected the detection of both inbreeding avoidance and MHC-dependent mate choice (Miller *et al.* 2009). In summary, controlling for the spatial distance among potential mates in test of MHC-dependent mate choice is essential to limit the risk of artefactual results suggesting that MHC dissimilarity, but not genetic relatedness, affects mate choice (because the spatial structure of functional and neutral genes may vary in different ways). Results of this kind are considered important in the field of MHC evolutionary ecology, as they are interpreted as MHC-specific effects on mate choice occurring beyond and over inbreeding avoidance.

Where to go next?

Taken together, these results indicate that studies of MHC-biased mate choice in wild vertebrate populations

face a number of methodological challenges. Discriminating MHC-biased mate choice from inbreeding avoidance using correlative approaches will greatly benefit from genotyping large samples of individuals, especially in long-term individually based studies where relatedness coefficients among individuals are estimated from a multigenerational pedigree built using parentage analyses (Clutton-Brock & Sheldon 2010), rather than from allele-sharing at a handful of neutral markers (see above). Next-generation sequencing massively increases the scope of MHC projects in the wild, by substantially decreasing the costs of genotyping, both in terms of time and money, compared to traditional methods (Babik *et al.* 2009; Babik 2010). As such, genotyping an entire population becomes realistic (Zagalska-Neubauer *et al.* 2010; Sepil *et al.* 2012) and may help to derive quantitative estimates of the (small) effect sizes of MHC on mate choice in order to calculate its fitness benefits, to discriminate MHC-disassortative mate choice and mate choice for genome-wide dissimilarity, as well as to test for allele-specific effects, all questions which are critical to test evolutionary theories proposed to explain the extraordinary polymorphism observed at MHC genes (Setchell & Huchard 2010).

Future studies may also usefully target a larger number of highly variable MHC markers, which may provide a more accurate assessment of genetic dissimilarity throughout the MHC, as well as help to compare the relative importance of several MHC markers on mate choice. Discriminating MHC and inbreeding avoidance may ultimately require the design of a set of densely spaced SNPs throughout the variable parts of the MHC region, to be compared to the variation observed at SNPs distributed throughout the genome, as can be done in some human studies (Chaix *et al.* 2008; Derti *et al.* 2010; Laurent & Chaix 2012; Laurent *et al.* 2012). Although this approach might be costly and difficult in species for which the genome is not fully sequenced, the development of population genomics makes this perspective realistic.

Biases resulting from a lack of control for the genetic structure of the population may be overcome in several ways. First, and ideally, knowing the identity of the candidate mates allows tests of MHC-biased mate choice without having to define 'potential partners', and some studies of wild populations have shown that this is sometimes possible (Westerdahl 2004). Even though observations of a large number of mating events may be difficult, recent progress in designing technological tools monitoring the spatial proximity between individuals might help acquiring these data (Ryder *et al.* 2012). Second, when documenting the identity of candidate mates is impossible, analytical designs may employ haplotype-based approaches, as these facilitate the use of population genetics tools which offer more possibilities to control for the genetic structure of the population. Controlling for the spatial distance among potential mates may also be possible using multivariate statistical models – such as generalized linear mixed-effect models – although the distribution of the response variable (successful vs. unsuccessful paternity) is expected to be highly 0-inflated in most mating systems (typically where operational sex ratios are highly male-biased, as it is the case in this study). Finally, experimental approaches represent the 'gold standard' approach and can sometimes be carried out in the wild or using individuals born in the wild (e.g. Olsson *et al.* 2003; Milinski *et al.* 2005; Consuegra & de Leaniz 2008).

Conclusions

Our findings suggest that observed deviations from random mate choice at MHC are driven by functionally important MHC genes (like DRB) rather than passively resulting from inbreeding avoidance. It also illustrates the challenges of demonstrating MHC-specific effects on mate choice using correlative studies in the wild by revealing how the research design – which MHC markers are used and whether analyses control for the population age and spatial structure – may affect the detection of inbreeding avoidance and MHC-dependent mate choice, as well as the estimation of their relative importance. Finally, it emphasizes the opportunities offered by next-generation sequencing to overcome such challenges through the genotyping of entire populations, preferably multiple generations of recognizable individuals with a documented life history.

Acknowledgements

We are very grateful to all people who have contributed to the collection of mouse lemur tissue samples between 2000 and 2010, especially Manfred Eberle and the Kirindy field assistants. We acknowledge the Département de Biologie Animale, Université d'Antananarivo, the CAFF of the Direction des Eaux et Forêts and the CNFEREF Morondava for authorizing this study, and more generally for permission to work in Kirindy Forest. We have adhered to the Guidelines for the Treatment of Animals in Behavioral Research and Teaching (Animal Behaviour 2006, 71: 245-253) and the legal requirements of the country (Madagascar) in which the fieldwork was carried out. We are very grateful to Christina Oberdieck for help with laboratory work, and to Lutz Walter, Christian Roos, Christina Albrecht and Markus Brameier for precious technical advice and help regarding 454-sequencing, and to three anonymous reviewers for insightful comments on a previous version of this manuscript. This research was funded by a Research Grant from the Deutsche Forschungsgemeinschaft (DFG) (HU 1820/1-1).

References

Apanius V, Penn D, Slev P, Ruff LR, Potts WK (1997) The nature of selection on the major histocompatibility complex. *Critical Review of Immunology*, **17**, 179–224.

Averdam A, Kuschal C, Otto N *et al.* (2011) Sequence analysis of the grey mouse lemur (*Microcebus murinus*) MHC class II DQ and DR region. *Immunogenetics*, **63**, 85–93.

Babik W (2010) Methods for MHC genotyping in non-model vertebrates. *Molecular Ecology Resources*, **10**, 237–251.

Babik W, Taberlet P, Ejsmond MJ, Radwan J (2009) New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. *Molecular Ecology Resources*, **9**, 713–719.

Baratti M, Dessi-Fulgheri F, Ambrosini R *et al.* (2012) MHC genotype predicts mate choice in the ring-necked pheasant *Phasianus colchicus*. *Journal of Evolutionary Biology*, **25**, 1531–1542.

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.

Bodmer WF (1972) Evolutionary significance of the HL-A system. *Nature*, **237**, 139–183.

Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G (2006) Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society of London. Series B*, **273**, 1111–1116.

Bos DH, Williams RN, Gopurenko D, Bulut Z, Dewoody JA (2009) Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Molecular Ecology*, **18**, 3307–3315.

Chaix R, Cao C, Donelly P (2008) Is mate choice in humans MHC-Dependent? *PloS Genetics*, **4**, 1–5.

Clutton-Brock T, Sheldon BC (2010) Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology & Evolution*, **25**, 562–573.

Consuegra S, de Leaniz CG (2008) MHC-mediated mate choice increases parasite resistance in salmon. *Proceedings of the Royal Society of London. Series B*, **275**, 1397–1403.

Csillery K, Johnson T, Beraldi D *et al.* (2006) Performance of marker-based relatedness estimators in natural populations if outbred vertebrates. *Genetics*, **173**, 2091–2101.

Derti A, Cenik C, Kraft P, Roth FP (2010) Absence of evidence for MHC-dependent mate selection within HapMap populations. *PloS Genetics*, **6**, e1000925.

Ditchkoff SS, Lochmiller RL, Masters RE, Hoofer SR, Van Den Bussche RA (2001) Major histocompatibility complex-associated variation in secondary sexual traits of white-tailed deer (*Odocoileus virginianus*): evidence for good-genes advertisement. *Evolution*, **55**, 616–625.

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

Doytchinova IA, Flower DR (2005) In silico identification of supertypes for class II MHCs. *Journal of Immunology*, **174**, 7085.

Eberle M, Kappeler PM (2002) Mouse lemurs in space and time: a test of the socioecological model. *Behavioral Ecology and Sociobiology*, **51**, 131–139.

Eberle M, Kappeler PM (2004a) Selected polyandry: female choice and intersexual conflict in a small nocturnal solitary primate (*Microcebus murinus*). *Behavioral Ecology and Sociobiology*, **57**, 91–100.

Eberle M, Kappeler PM (2004b) Sex in the dark: determinants and consequences of mixed male mating tactics in *Microcebus murinus*, a small solitary nocturnal primate. *Behavioral Ecology and Sociobiology*, **57**, 77–90.

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

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

Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012b) Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications*, **3**, doi:10.1038/ncomms1632.

Ekblom R, Saether A, Grahn M *et al.* (2004) Major Histocompatibility complex variation and mate choice in a lekking bird, the great snipe (*Gallinago media*). *Molecular Ecology*, **13**, 3821–3828.

Evans ML, Dionne M, Miller KM, Bernatchez L (2011) Mate choice for major histocompatibility complex genetic divergence as a bet-hedging strategy in the Atlantic salmon (*Salmo salar*). *Proceedings of the Royal Society of London. Series B*, **22**, 379–386.

Excoffier L, Lischer HEL (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.

Forsberg LA, Dannewitz J, Petersson E, Grahn M (2007) Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout – females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology*, **20**, 1859–1869.

Freeman-Gallant CR, Meguerdichian M, Wheelwright NT, Sollecito SV (2003) Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Molecular Ecology*, **12**, 3077–3083.

Galan M, Guivier E, Caraux G, Charbonnel N, Cosson J-F (2010) A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. *BMC Genomics*, **11**, 296.

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

Gomez D, Huchard E, Henry P-Y, Perret M (2012) Mutual mate choice in a female-dominant and sexually dimorphic primate. *American Journal of Physical Anthropology*, **147**, 370–379.

Hedrick PW, Kim T (1999) Genetics of complex polymorphism: parasites and maintenance of the Major Histocompatibility Complex variation. In: *Evolutionary Genetics: From Molecules to Morphology* (eds. Singh R, Krimbas C), pp. 204–233. Cambridge University Press, Cambridge.

Huchard E, Weill M, Cowlishaw G, Raymond M, Knapp LA (2008) Polymorphism, haplotype composition, and selection in the *Mhc-DRB* of wild baboons. *Immunogenetics*, **60**, 585–598.

Huchard E, Knapp LA, Wang J, Raymond M, Cowlishaw G (2010) MHC, mate choice and heterozygote advantage in a wild social primate. *Molecular Ecology*, **19**, 2545–2561.

Huchard E, Albrecht C, Schliehe-Diecks S *et al.* (2012a) Large-scale MHC class II genotyping of a wild lemur population by next generation sequencing. *Immunogenetics*, **64**, 895–913.

Huchard E, Canale CI, Le Gros C *et al.* (2012b) Convenience polyandry or convenience polygyny? Costly sex under female control in a promiscuous primate. *Proceedings of the Royal Society of London. Series B*, **279**, 1371–1379.

Jennions MD (1997) Female promiscuity and genetic incompatibility. *Trends in Ecology and Evolution*, **12**, 251–253.

Kappeler PM, Fichtel C (2012) A 15-year perspective on the social organization and life history of sifaka in Kirindy Forest. In: *Long-term field studies of primates*(eds Kappeler PM & Watts DP), pp. 101–121. Springer, Heidelberg.

Kraus C, Eberle M, Kappeler PM (2008) The costs of risky male behavior: sex differences in seasonal survival in a small sexually monomorphic primate. *Proceedings of the Royal Society of London. Series B*, **275**, 1635–1644.

Landry C, Garant D, Duchesne P, Bernatchez L (2001) 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proceedings of the Royal Society of London. Series B*, **268**, 1279–1285.

Laurent R, Chaix R (2012) MHC-dependent mate choice in humans: why genomic patterns from the HapMap European American dataset support the hypothesis. *BioEssays*, **34**, 267–271.

Laurent R, Toupance B, Chaix R (2012) Non-random mate choice in humans: insights from a genome scan. *Molecular Ecology*, **21**, 587–596.

Manning CJ, Wakeland EK, Potts WK (1992) Communal nesting patterns in mice implicate MHC genes in kin recognition. *Nature*, **360**, 581–583.

Mays H, Hill GE (2004) Choosing mates: good genes versus genes that are a good fit. *Trends in Ecology and Evolution*, **19**, 554–559.

Milinski M, Griffiths S, Wegner KM *et al.* (2005) Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 4414–4418.

Miller HC, Moore JA, Nelson NJ, Daugherty CH (2009) Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proceedings of the Royal Society of London. Series B*, **276**, 1695–1704.

Oliver MK, Lambin X, Cornulier T, Piertney SB (2009) Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. *Molecular Ecology*, **18**, 80–92.

Olsson M, Madsen T, Nordby J *et al.* (2003) Major histocompatibility complex and mate choice in sand lizards. *Proceedings of the Royal Society of London. Series B*, **270**, S254–S256.

Paterson S, Pemberton JM (1997) No evidence for major histocompatibility complex-dependent mating patterns in a free-living ruminant population. *Proceedings of the Royal Society of London. Series B*, **264**, 1813–1819.

Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity*, **96**, 7–21.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

Reusch TBH, Haberli MA, Aeschlimann PB, Milinski M (2001) Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*, **414**, 300–302.

Richardson DS, Komdeur J, Burke T, von Schantz T (2005) MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society of London. Series B*, **272**, 759–767.

Roberts SC (2009) Complexity and context of MHC-correlated mating preferences in wild populations. *Molecular Ecology*, **18**, 3121–3123.

Ryder TB, Horton BM, van den Tillaart M, Morales JDD, Moore IT (2012) Proximity data-loggers increase the quantity and quality of social network data. *Biology Letters*, **8**, 917–920.

Sandberg M, Eriksson L, Jonsson J, Sjostrom 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.

Schad J, Sommer S, Ganzhorn JU (2004) MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. *Conservation Genetics*, **5**, 299–309.

Schad J, Ganzhorn JU, Sommer S (2005) Parasite burden and constitution of major histocompatibility complex in the Malagasy mouse lemur, *Microcebus murinus*. *Evolution*, **59**, 439–450.

von Schantz T, Wittzell H, Göransson G, Grahn M, Persson K (1996) MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proceedings of the Royal Society of London. Series B*, **263**, 265–271.

Schliehe-Diecks S, Eberle M, Kappeler PM (2012) Walk the line - dispersal movements of gray mouse lemurs (*Microcebus murinus*). *Behavioral Ecology and Sociobiology*, **66**, 1175–1185.

Schwensow N, Fietz J, Dausmann KH, Sommer S (2007) Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity*, **99**, 265–277.

Schwensow N, Eberle M, Sommer S (2008a) Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proceedings of the Royal Society of London. Series B*, **275**, 555–564.

Schwensow N, Fietz J, Dausmann K, Sommer S (2008b) MHC-associated mating strategies and the importance of overall genetic diversity in an obligate pair-living primate. *Evolutionary Ecology*, **22**, 617–636.

Schwensow N, Eberle M, Sommer S (2010) Are there ubiquitous parasite-driven Major Histocompatibility Complex selection mechanisms in gray mouse lemurs? *International Journal of Primatology*, **31**, 519–537.

Sepil I, Moghadam HK, Huchard E, Sheldon BC (2012) Characterization and 454 pyrosequencing of Major Histocompatibility Complex class I genes in the great tit reveal complexity in a passerine system. *BMC Evolutionary Biology*, **12**, 68.

Setchell JM, Huchard E (2010) The hidden benefits of sex: evidence for MHC-associated mate choice in primate societies. *BioEssays*, **32**, 940–948.

Setchell JM, Charpentier MJE, Abbott KM, Wickings EJ, Knapp LA (2010) Opposites attract: MHC-associated mate choice in a polygynous primate. *Journal of Evolutionary Biology*, **23**, 136–148.

Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity*, **76**, 377–383.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society of London. Series B*, **277**, 979–988.

Strandh M, Westerdahl H, Pontarp M *et al.* (2012) Major Histocompatibility complex class II, but not class I, predicts mate choice in a bird with highly developed olfaction. *Proceedings of the Royal Society of London. Series B*, **279**, 4457–4463.

Tregenza T, Wedell N (2000) Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology*, **9**, 1013–1017.

Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research*, **89**, 135–153.

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, **11**, 141–145.

Wegner KM (2009) Massive parallel MHC genotyping: titanium that shines. *Molecular Ecology*, **18**, 1818–1820.

Westerdahl H (2004) No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology*, **13**, 2465–2470.

Westerdahl H, Hansson B, Bensch S, Hasselquist D (2004) Between-year variation of MHC allele frequencies in great reed warblers: selection or drift? *Journal of Evolutionary Biology*, **17**, 485–492.

Yeates SE, Einum S, Fleming IA *et al.* (2009) Atlantic salmon eggs favour sperm in competition that have similar major histocompatibility alleles. *Proceedings of the Royal Society of London. Series B*, **276**, 559–566.

Zagalska-Neubauer M, Babik W, Stuglik M *et al.* (2010) 454 sequencing reveals extreme complexity of the class II Major Histocompatibility Complex in the collared flycatcher. *BMC Evolutionary Biology*, **10**, 395.

E.H. and P.M.K. designed the study, E.H. and S.S.D. collected the data and performed the labwork, E.H. and A.B. analysed the data and all authors contributed to the writing-up.

Data accessibility

The main data set supporting the analyses presented is available in DRYAD (doi:10.5061/dryad.647 h6, entitled 'Mimu_Main dataset') as well as the spatial coordinates of individual home ranges (doi:10.5061/dryad.647 h6, entitled 'Mimu_Spatial data'), the microsatellite genotypes (doi:10.5061/dryad.647 h6, entitled 'Mimu_Microsatellite data') and the MHC-DRB and MHC-DQB exon 2 genotypes (doi:10.5061/dryad.647 h6, entitled 'Mimu_MHC data'). MHC allele sequences identified in the course of this project have been deposited in Genbank (Accession nos HE801914-HE801964) while the full list of alleles present in this sample is given in Huchard *et al.* 2012a (see Appendix).

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Methods.

Table S1 Sample size included in various permutation tests.

Table S2 Pearson's product moment correlation coefficients between the different MHC-dissimilarity variables at DRB (indicated by the letter R), DQB (indicated by the letter Q), haplotype (indicated by the letter H) and relatedness assessed using microsatellite markers.

Table S3 Results of tests of linkage disequilibrium between MHC loci (DQB and DRB) and microsatellites.

Table S4 Results of the tests investigating choice for maximal genetic dissimilarity or diversity at the haplotype level for a dataset including only females older than one year and a dataset including all females.

Table S5 Results of the mate choice tests for genetic dissimilarity involving a variable set of candidate mates.

Table S6 Effect of the number of candidate mates on the power of the analysis investigating the impact of variable sets of candidate mates.

Figure S1 Effect of the number of candidate mates on the power of the analysis investigating the impact of variable sets of candidate mates.