Molecular Ecology (2009) **18**, 1331–1340 doi: 10.1111/j.1365-294X.2009.04105.x

Diversifying selection on MHC class I in the house sparrow (*Passer domesticus*)

CLAIRE LOISEAU,* MURIELLE RICHARD,† STÉPHANE GARNIER,‡ OLIVIER CHASTEL,§ ROMAIN JULLIARD,¶ RIMA ZOOROB** and GABRIELE SORCI‡

**Laboratoire Parasitologie Evolutive, CNRS UMR 7103, †Laboratoire Fonctionnement et Evolution des écosystèmes, CNRS UMR 7625, Université Pierre et Marie Curie, 7 quai St Bernard, Case 237, F-75252 Paris Cedex 05, France, ‡Laboratoire BioGéoSciences, CNRS UMR 5561, Université de Bourgogne, 6 Boulevard Gabriel, F-21000 Dijon, France, §Centre d'Etudes Biologique de Chizé, CNRS UPR 1934, F-79360 Beauvoir-sur-Niort, France, ¶Laboratoire Conservation des espèces, restauration et suivi des populations, CNRS UMR 5173, Muséum National d'Histoire Naturelle, 55 rue Buffon, F-75005 Paris, France, **Laboratoire Génétique moléculaire et intégration des fonctions cellulaires, CNRS FRE 2937, 7 rue Guy Mocquet, BP8, F-94801 Villejuif Cedex, France*

Abstract

Genes of the major histocompatibility complex (MHC) are the most polymorphic loci known in vertebrates. Two main hypotheses have been put forward to explain the maintenance of MHC diversity: pathogen-mediated selection and MHC-based mate choice. Host–parasite interactions can maintain MHC diversity via frequency-dependent selection, heterozygote advantage, and diversifying selection (spatially and/or temporally heterogeneous selection). In this study, we wished to investigate the nature of selection acting on the MHC class I across spatially structured populations of house sparrows (*Passer domesticus*) in France. To infer the nature of the selection, we compared patterns of population differentiation based on two types of molecular markers: MHC class I and microsatellites. This allowed us to test whether the observed differentiation at MHC genes merely reflects demographic and/or stochastic processes. At the global scale, diversifying selection seems to be the main factor maintaining MHC diversity in the house sparrow. We found that (i) overall population differentiation at MHC was stronger than for microsatellites, (ii) MHC marker showed significant isolation by distance. In addition, the slope of the regression of F_{ST} on geographical distance was significantly steeper for MHC than for microsatellites due to a stronger pairwise differentiation between populations located at large geographical distances. These results are in agreement with the hypothesis that spatially heterogeneous selective pressures maintain different MHC alleles at local scales, possibly resulting in local adaptation.

Keywords: balancing selection, diversifying selection, house sparrow, MHC class I, microsatellites, *Passer domesticus*

Received 2 October 2008; revision received 13 December 2008; accepted 20 December 2008

Introduction

Genes of the major histocompatibility complex (MHC) encode molecules responsible for the recognition and presentation of foreign antigens to cells of the immune system (Klein 1986). The region of the molecule responsible for the antigen presentation is the peptide-binding region. Variation in this coding region determines the antigenic peptides an individual is able to recognize (Klein 1986). Understanding how variation of these selected genes is maintained has received much attention from evolutionary biologists (Hedrick 1972; Hedrick & Thomson 1983; Hughes & Nei 1992; Borghans *et al*. 2004; De Boer *et al*. 2004). In addition to mate-choice-based mechanisms, selective fertilizations and abortions (Apanius *et al*. 1997; Edwards & Hedrick 1998), parasite-mediated selection can explain the maintenance of MHC variation (Hedrick 1999), in the form of several, nonexclusive, mechanisms.

Heterozygous individuals may enjoy a selective advantage if different alleles confer protection against a wider spectrum of pathogens. Heterozygous individuals might, however, also perform better than homozygous when facing single pathogen infections, if two different MHC

Correspondence: Claire Loiseau, Fax (33)-144273516; E-mail: cloiseau@snv.jussieu.fr

molecules do better at binding the antigens. This process has been called overdominance and implies that heterozygous individuals should do better than individuals that are homozygous for either of the same alleles that the heterozygous carries at a given locus (Doherty & Zinkernagel 1975; Hughes & Nei 1988; Carrington *et al*. 1999; Penn *et al*. 2002; McClelland *et al*. 2003).

Another mechanism that might account for the maintenance of MHC diversity is negative frequency-dependent selection (Apanius *et al*. 1997). Rare alleles are supposed to have a selective advantage over common alleles because pathogens tend to adapt to the commonest genetic variants (Meyer-Lucht & Sommer 2005; Schad *et al*. 2005).

Finally, selection that varies in time and/or space could maintain allele diversity as shown by theoretical models and empirical data (Hedrick 1999, 2002; Landry & Bernatchez 2001; Meyer & Thomson 2001). It is important to note that these selection forces can operate in concert to maintain MHC diversity.

Recently, several studies have investigated the geographical variation of MHC genes in wildlife (Landry & Bernatchez 2001; Piertney 2003; Lukas *et al*. 2004; Aguilar & Garza 2006; Bowen *et al*. 2006; Hayashi *et al*. 2006; Ekblom *et al*. 2007; Alcaide *et al*. 2008). The results of these studies generally indicate that either the strength of the selection on MHC genes is weak relative to other microevolutionary forces, or that selection acts in a diversifying manner across populations, possibly as a result of heterogeneity in parasite abundance and diversity. However, such inference is often made without examination of how parasite load or diversity changes spatially. Therefore, characterizing variation in parasite intensity and diversity across populations is an important step in the way to understand the maintenance of MHC diversity.

The aim of this study was to investigate the nature of selection acting on MHC class I by comparing patterns of genetic differentiation based on MHC and on microsatellite loci among 13 house sparrow (*Passer domesticus*) populations located at different geographical distances. The house sparrow has already been used as a model species to explore the distribution of their blood parasites (Valkiunas *et al*. 2006) and the role of MHC genes in the resistance to malaria parasites (Bonneaud *et al*. 2006; Loiseau *et al*. 2008). Previous work on different bird species have shown that (i) exposure to malaria is heterogeneous in space as shown by significant difference in parasite diversity and prevalence across populations (Bensch & Akesson 2003; Beadell *et al*. 2004; Svensson *et al*. 2007; Loiseau *et al*. in preparation) and within-populations (Wood *et al*. 2007), (ii) MHC alleles are associated with risks of harbouring malaria infection (Westerdahl *et al*. 2005; Loiseau *et al*. 2008) and, importantly, associations differ across populations (Bonneaud *et al*. 2006; Loiseau *et al*. in preparation). Interestingly, Westerdahl *et al*. (2005) have shown, in a temporal dimension, that the variation in MHC allele frequencies between cohorts is not a result of demographic events, but rather an effect of selection favouring different MHC alleles in different years. Here, in a spatial dimension, we wished to determine if the pattern of genetic differentiation at MHC is mainly driven by the effect of mutation, gene flow and random genetic drift or if selection plays also a role in the observed genetic differentiation. The comparison between the pattern of differentiation based on MHC and microsatellite loci can give three possible outcomes, each of which would provide evidence for different scenarios. First, if population differentiation based on MHC merely reflects demographic and/or stochastic factors, with no selection acting on MHC, we should observe no difference between MHC and microsatellite differentiation patterns. Second, if balancing selection is the main selective force acting on MHC, we should expect a weaker population differentiation at the MHC compared with differentiation at microsatellites. Third, if diversifying selection is the main selective force acting on MHC, we should expect a stronger differentiation based at the MHC than at microsatellites because different allelic lineages may be retained in different populations.

Methods

Sampling

The house sparrow is a common, sedentary bird (Summers-Smith 1988) in both rural and urban areas. Since 2004, a monitoring was started in different sites through France. Birds were caught with mist nets and ringed with a numbered metal ring. We collected blood samples in 13 populations (from 20 to 85 individuals per population, Fig. 1; Table 1), between April 2004 and June 2006.

Screening of MHC class I variation

Genomic DNA was extracted from blood samples using the QIAquick 96 Purification Kit (QIAGEN) according to the manufacturer's instructions. We screened individuals to assess allelic diversity at the most variable MHC class I gene family using the polymerase chain reaction (PCR)-based denaturant gradient gel electrophoresis (DGGE) method (Myers *et al*. 1987). This method allows us to examine single nucleotide polymorphism at MHC class I exon 3, corresponding to the highly variable peptide-binding site of the protein (α2 domain). In order to preferentially amplify transcribed alleles, cDNA sequences were used to design the following primers: GCA21M 5′-CGTACAGCGG-CTTGTTGGCTGTGA-3′ and fA23M 5′-GCGCTCCAGCTC-CTTCTGCCCATA-3′ (Bonneaud *et al*. 2004). Amplifications were run in a final volume of 50 μL including 15–50 ng of DNA, 0.25 μM of each primer, 200 μM of dNTPs, 5 μL of 10× buffer and 0.5 U of *Taq* DNA polymerase. The thermal

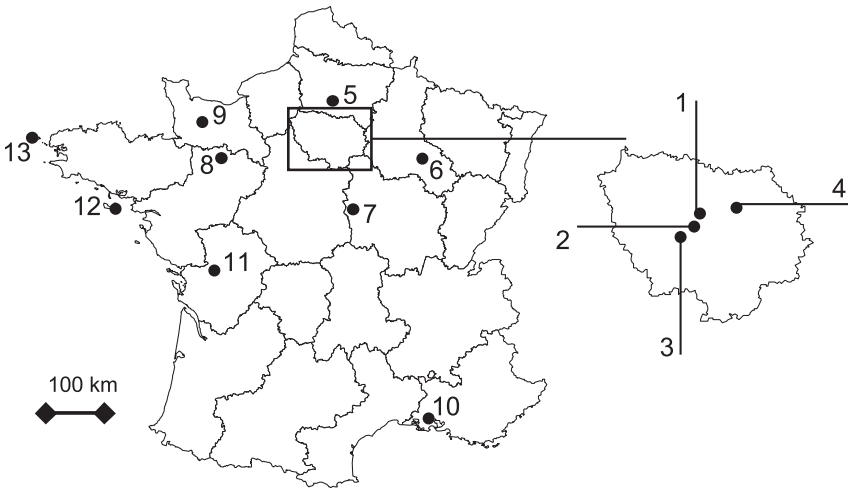

Fig. 1 Location of sampling sites in France.

Table 1 Sample size (N) for each marker, index of allelic richness Theta (k) for MHC marker and allelic richness (A), observed (H_O) and expected heterozygosities (H_E) for microsatellite marker are given for each population

			MHC		Microsatellites			
Population	N_{Total}	Sampling	N	Theta k	N	A	H_O	H_E
1 — Paris	83	2004–2005	52	10.75	83	28.80	0.880	0.923
2 — Cachan	59	2004	20	13.38	59	26.58	0.897	0.916
3 — Wissous	39	2004	39	9.29	37	24.81	0.943	0.914
4 — Crégy	85	2005–2006	85	8.88	58	28.97	0.889	0.926
5 — Thieux	45	2004	45	7.83	36	25.99	0.824	0.921
6 — Anglus	58	2004–2005 (winter)	58	7.66	36	27.62	0.869	0.913
7 — Cours	61	2004–2005 (winter)	61	9.15	55	28.16	0.878	0.926
8 — Crennes	61	2004–2005	55	12.85	61	25.55	0.900	0.912
9 — Rully	58	2005	58	9.22	50	27.53	0.843	0.920
10 — Arles	49	2005	49	10.57	49	28.76	0.883	0.920
11 — Chizé	62	2005	54	13.58	62	27.46	0.883	0.921
12 — Hoedic	52	2006	44	7.61	52	22.09	0.836	0.903
13 — Ouessant	57	2006	44	4.98	57	23.36	0.877	0.911

profile started with 1 min 30 of denaturation at 94 °C, followed by 35 cycles at 94 °C, 65 °C and 72 °C for 30 s each, and ended with an elongation step at 72 °C for 10 min. The PCR products were separated using a DGGE. The DGGE gel contained 7% 19:1 acrylamide/bisacrylamide, 1× TAE, formamide, and a 40–65% denaturating gradient of urea (Bonneaud *et al*. 2004). The gels were run at 60 °C in 1× TAE buffer for 20 h at 180 V. Gels were stained with ethidium bromide, and visualized under UV light illumination. All gels always included two copies of a same marker (i.e. PCR fragments made from genomic DNA from three house sparrow individuals) to enable comparisons between gels. The migration distance of the bands on the DGGE gels were identified relative to these marker bands with a high repeatability. Only one person read all the DGGE gels to limit errors. In addition, 12 individuals were double-checked in at least two gels and in five cases in three DGGE gels. The same number and the same identity of DGGE bands were always found.

Microsatellite genotyping

Individuals were genotyped using seven microsatellite markers: Pdo3, Pdo4, Pdo5, Pdo6 (Griffith *et al*. 1999), Mjg1 (Li *et al*. 1997), Ase18 (Richardson *et al*. 2000) and Fhu2 (Primmer *et al*. 1996). We performed two multiplex: one with Fhu2, Mjg1, Ase 18 and one with Pdo3, Pdo4, Pdo5 and Pdo6. Each multiplex was amplified in a final volume of 10 μL including 15–50 ng of DNA, 50–200 nM of each primer (concentrations given in Table 2), 300 μM of dNTPs, 1 μL of 10× incubation buffer (50 mM KCl, 10 mM Tris-HCl, 1.5 mM MgCl2, 0.1% TritonX-100, pH 9.0) and 0.25 U of *Taq* DNA polymerase (Qbiogene). Samples were subjected to initial denaturation at 94 °C for 10 min, followed by 10

Table 2 Characteristics of the seven microsatellite locus. *C*, primers concentrations (nM) used in PCR reaction (*F* = forward primer; *R* = reverse primer); N_a, number of alleles found over all sampling sites and the size range

Locus	Reference	C_F	C_R	N_a	Range
Ase18	Richardson *et al*. 2000	200	200	35	187–249
Mjg1	Li *et al*. 1997	50	50	29	153–183
FhU2	Primmer *et al*. 1996	100	100	29	111–148
Pdo3	Griffith *et al*. 1999	150	150	36	99–164
Pdo4	Griffith *et al*. 1999	200	150	213	222–471
Pdo5	Griffith *et al*. 1999	100	50	28	223–251
Pdo6	Griffith *et al*. 1999	100	100	187	281–482

cycles at 94 °C, 50 °C and 72 °C for 15 s each and by 20 cycles at 89 °C, 50 °C and 72 °C for 15 s each, with a final extension phase at 72 °C for 10 min. The reaction was performed in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems). The PCR fragments were fluorescently labelled using primers bound to different dyes (FAM, HEX or NED) and were analysed in an ABI 310 automated sequencer following the manufacturer's protocol using GeneMapper 3.0 software.

Estimates of genetic variation

Intrapopulation genetic diversity. Among the 13 populations, 10 populations were sampled in the same year or two different years but during the same winter, meaning that no reproduction occurred between sampling events. Three populations were sampled during two distinct years. For these three populations, we pooled data from different years because we found no genotypic differentiation at microsatellite loci between years for two populations, and because the significant differentiation between years that we found for the third sampling site (site 1) was due to a single locus (Pdo6). This finding should not be seen as evidence against the neutrality of Pdo6 since the F_{ST} outlier analysis (Beaumont & Nichols 1996; see results below) confirmed that none of the loci used in this study is under selection.

For microsatellites, intrapopulation genetic variation was estimated by observed (H_O) and expected (H_E) heterozygosities using the program Genetix (Belkhir *et al*. 2004). Allelic richness was calculated using a rarefaction index and estimated for the smallest sample size (36 individuals) using FSTAT 2.9.3 (Goudet 1995). Linkage disequilibrium between all pairs of loci and deviations from Hardy–Weinberg equilibrium for each locus were tested within each population using exact tests (Guo & Thompson 1992) as implemented in GenePop 4.0 (Rousset 2008). A global test for Hardy–Weinberg equilibrium across all loci was constructed using Fisher's method, providing that loci were previously shown to be statistically independent. Finally, we used the method put forward by Beaumont & Nichols (1996) to check whether there was any selection on each of the six microsatellite loci using the F_{ST} outlier approach. We used the programs FDIST 2 (Beaumont & Nichols 1996; Flint *et al*. 1999) and we set the parameter d (the total number of demes present in the system) to 100, using the infinite allele model of mutation. Coalescent simulations were performed to generate 20 000 paired values of F_{ST} and heterozygosity. These simulations were used to compute the 0.975, 0.5 and 0.025 quantiles of the distribution of F_{ST} as a function of heterozygosity.

For the MHC class I, calculations of allele frequencies and index of allelic richness (theta *k*) were performed using Arlequin 3.1 (Excoffier *et al*. 2005), entering the sequence name and number of individuals with an allele as haplotypic data. Allele frequencies were thus estimated as the number of individuals carrying a certain allele divided by the total allele count observed in the population. Total allele count is defined as the sum of different alleles per individual. This procedure was necessary since more than two alleles were found in some individuals, because we amplified more than one locus. Indeed, like previously found (Bonneaud *et al*. 2004), we amplified at least five loci with a number of MHC alleles varying between 2 and 10. We have to note that this way of determining allele frequencies may underestimate the frequency of common alleles and overestimate the frequency of rare alleles (see Ekblom *et al*. 2007).

Population differentiation and isolation by distance. For both markers, differentiation across all populations and between population pairs was tested using a log-likelihood based exact *G*-test (Goudet *et al*. 1996) using GenePop 4.0 (Rousset 2008) for microsatellites and Arlequin 3.1 (Excoffier *et al*. 2005) for MHC class I data. In addition, both global and pairwise estimate of F_{ST} were estimated using Arlequin 3.1 (Excoffier *et al*. 2005) following Weir & Cockerham (1984). We identified 95% confidence intervals (CI) of both overall F_{ST} and pairwise F_{ST} for microsatellites by bootstrapping over loci in FSTAT 2.9.3 (Goudet 1995).

Isolation by distance was assessed by testing the correlation between $F_{ST}/(1-F_{ST})$ and the logarithm of the geographical distance (km) considering all population pairs for each type of marker (Rousset 1997). Statistical significance was evaluated using Mantel tests (with 10 000 permutations, using XLSTAT). We also performed a partial Mantel test where pairwise F_{ST} at MHC were correlated with geographical distance while keeping constant differentiation at microsatellites. This test should provide evidence in favour of a significant positive correlation between geographical distance and MHC differentiation that is independent of demographic and stochastic factors (see Ekblom *et al*. 2007 for a similar example). We also tested whether the slopes of the regression between $F_{ST}/(1-F_{ST})$ and ln(geographical distance) based on the two markers were significantly

different. We computed the 95% confidence intervals (CI) for the two slopes to check whether they were overlapping or not. To go a bit further in the comparison of the genetic differentiation with geographical distance, we assessed if all MHC pairwise F_{ST} overlapped with the 95% CI of all microsatellite pairwise F_{ST}, considering non-overlapping values as significantly different (Weir 1996). Finally, we compared levels of differentiation estimated from the two markers at different geographical distances by the mean of a single parameter. We computed the pairwise F_{ST} for each of the six microsatellites and the MHC, and assigned a rank to each pairwise F_{ST} (rank 1 to the highest value of pairwise F_{ST}, rank 7 to the lowest value). We then tested the correlation between ranks obtained for MHC-based pairwise F_{ST} and geographical distances, using a Mantel test with 10 000 permutations.

Results

Intrapopulation genetic diversity

Considering all the microsatellite loci, a total number of 528 alleles were found from 13 populations. The amount of polymorphism varied among loci, ranging from 29 to 213 alleles (Table 2). Out of 267 exacts tests performed for genotypic disequilibrium, 27 (10.1% of these tests) were significant at the 0.05 level. However, only six tests remained significant after sequential Bonferroni correction. Because these tests did not concern the same pairs of loci and the same populations, there was no evidence of linkage between loci, which were then considered statistically independent.

Hardy–Weinberg equilibrium (HWE) was tested for each locus in each population. A total of 91 tests were performed and 38 were significant at the 0.05 level. Eleven tests remained significant after Bonferroni correction: these tests concerned 4 of the 7 loci (FhU2, Pdo4, Pdo5 and Pdo6) and 9 of the 13 sites. Results strongly suggested the presence of null alleles for locus FhU2. Indeed, 4 out of the 11 significant tests after Bonferroni correction concerned this locus. In addition, we found positive and high F_{IS} values in all populations, ranging from 0.055 to 0.379 (mean = 0.186). Then, this locus was removed from all further analyses. The remaining significant tests concerned locus Pdo4, Pdo5, Pdo6 (with 3, 2, and 2 significant tests, respectively). Although we cannot fully discard the presence of null alleles in some locus–population combinations, we did not exclude these three loci from the analyses because they did not show a systematic heterozygote deficiency. Indeed, for instance, except for the remaining significant tests, Pdo4 exhibited F_{IS} values, close to zero, from –0.019 to 0.063, contrary to FhU2 for which F_{IS} values ranged from 0.055 to 0.284, suggesting a strong departure from HWE.

Allelic richness, observed and expected heterozygosity values for the six remaining microsatellite loci are given in Table 1. Values of gene diversity (H_E) were quite high and homogeneous across populations, ranging from 0.903 to 0.926. Allelic richness ranged from 22.09 to 28.97 microsatellite alleles per population and the two smallest values were observed in the two insular populations.

MHC class I alleles were screened for a total number of 664 individuals (Table 1). Sixty different alleles were found. Six alleles were found in all populations and seven alleles were specific to a single population. Allele frequencies are reported in Appendix S1, Supporting Information. We found a large difference between populations for index of allelic richness which ranged from 4.98 to 13.58 (Table 1). Overall, the number of amplified MHC alleles varies between 2 and 10. The number of alleles per individual differed among populations ($F_{12,645} = 5.60$, $P = 0.001$); however, this was due to a single population where individuals have a higher number of alleles compared to the other populations (population 7, mean ± SE: 5.08 ± 0.27). When removing this population, the number of alleles per individual no longer differed among populations ($F_{11,646} = 1.30$, $P = 0.218$).

Population differentiation and isolation by distance

For microsatellite loci and MHC, the global test of differentiation among samples was highly significant ($P < 0.0001$). Overall F_{ST} for the microsatellites and the MHC were 0.0085 (0.0071–0.0098: 95% CI), and 0.0120, respectively. In addition, differentiation test between all pairs of samples showed that all pairs and 76 out of 78 were significantly differentiated for microsatellite loci and MHC, respectively. Pairwise F_{ST} values for microsatellite loci ranged from 0.0015 to 0.0207, and for MHC from 0.0007 to 0.0415. In addition, over the 13 populations, no F_{ST} outlier was detected for each of the six microsatellite loci (Fig. 2).

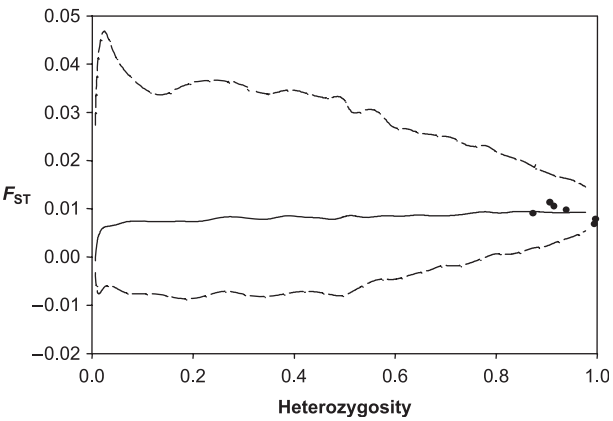

Fig. 2 Results of the F_{ST} outlier analysis. Solid line indicates the median value of F_{ST} and the dashed lines, the upper and lower 95% quantiles from 20 000 simulations. Solid circles are observed F_{ST} values from the six microsatellites loci.

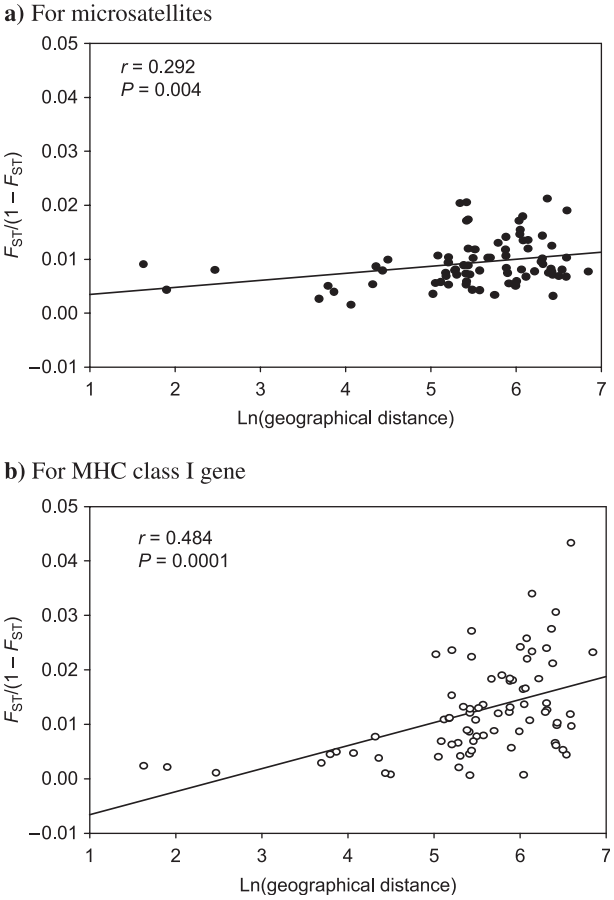

Fig. 3 Correlation between pairwise $F_{ST}/(1 - F_{ST})$ and ln(geographical distance) in kilometres for 13 populations, (a) for the six microsatellite loci, and (b) for the exon 3 of the MHC class I gene family. Significance of correlation was established with a Mantel test (100 000 permutations).

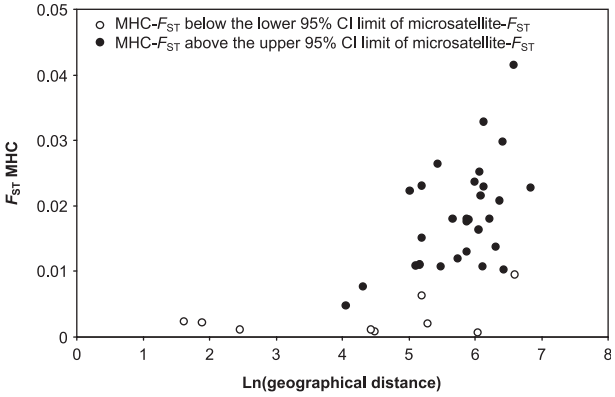

Fig. 4 Correlation between levels of pairwise population differentiation for the MHC class I gene family and ln(geographical distance). Open symbols represent MHC-based F_{ST} values below the lower 95% confidence interval limit of the microsatellite-based F_{ST} estimates. Black symbols represent MHC-based F_{ST} values above the upper 95% confidence interval limit of the microsatellite-based F_{ST} estimates.

There was a statistically significant isolation by distance for both microsatellites and MHC (microsatellites: $r = 0.29$, $P = 0.004$, Fig. 3a; MHC: $r = 0.48$, $P < 0.0001$, Fig. 3b). However, the significant isolation by distance based on microsatellite loci was due to the two insular populations (sites 12 and 13); when removing these two populations, the isolation by distance was no longer significant ($r = 0.19$, $P = 0.165$). On the contrary, removing the two insular populations from the analysis did not alter the result for the MHC ($r = 0.43$, $P < 0.0001$). In addition, pairwise F_{ST} for the MHC were positively correlated with microsatellite pairwise F_{ST} ($r = 0.28$, $P = 0.018$). However, in spite of this correlation, the correlation between F_{ST} based on the MHC and geographical distances was very slightly affected by the concomitant pattern observed for microsatellites, because a partial Mantel test keeping constant microsatellite differentiation provided a very similar result (partial Mantel test; $r = 0.42$, $P = 0.0001$). The slope of the regression of MHC F_{ST} on geographical distance [$b = 0.0042$; 95% CI (0.0025; 0.0059)] was significantly steeper than the slope of the regression of microsatellite F_{ST} on geographical distance [$b = 0.0013$; 95% CI (0.0003; 0.0023)].

In addition, we found nine pairwise comparisons where the MHC-based F_{ST} were below the lower limit of the 95% confidence interval of the microsatellite-based F_{ST}, 39 pairwise comparisons where the MHC-based F_{ST} were within the 95% CI of the microsatellite-based F_{ST}, and 30 pairwise comparisons where the MHC-based F_{ST} were above the upper limit of the 95% confidence interval of the microsatellite-based F_{ST} (Fig. 4).

The correlation between the rank of the F_{ST} based on MHC and the geographical distance was negative ($r = -0.41$) and statistically significant (Mantel test, $P = 0.00002$). This result means that differentiation at MHC was relatively stronger (compared with microsatellites) for populations located at large geographical distances.

Discussion

The aim of this study was to explore the nature of the selection acting on the MHC class I of 13 populations of house sparrows, supposedly experiencing variable exposure to parasites. The finding of a positive correlation between pairwise estimates of F_{ST} for the MHC and neutral markers indicates that genetic drift and migration do play a role in the observed MHC variation in the house sparrow. However, our results suggest that neutral processes are not sufficient to explain spatial variation in MHC class I among the studied populations. In agreement with the prediction

of diversifying selection, we found a stronger pattern of isolation by distance for MHC class I than for neutral markers (i.e. the slope of the regression of F_{ST} on distance was significantly steeper for MHC than for microsatellites).

Comparing population differentiation at MHC and neutral genes provides a powerful tool to assess the nature of the selection acting on MHC genes. Accordingly, in the last years, a few studies have adopted this approach to explore selection on MHC in fish, birds and mammals (Miller & Withler 1997; Landry & Bernatchez 2001; Miller *et al*. 2001; Piertney 2003; Aguilar & Garza 2006; Bryja *et al*. 2007; Ekblom *et al*. 2007). However, drawing a firm conclusion from a direct comparison between MHC and microsatellites can be misleading for several reasons. First, the assumption of neutrality may be violated for some microsatellite loci, for instance, because of linkage with selected genes (Slatkin 1995; Kohn *et al*. 2000; Vasemägi *et al*. 2005; Bryja *et al*. 2007; Larsson *et al*. 2007). We thus checked whether there was any indication of selection on each of the six microsatellite loci using the F_{ST} outlier approach put forward by Beaumont & Nichols (1996), and found that no microsatellite locus gave an F_{ST} value outside the 95% limits of the randomly generated F_{ST} distribution. Second, MHC and microsatellite genes differ in their pattern of molecular evolution and the amount of variability. Comparing highly polymorphic markers with markers exhibiting lower levels of variability can bias estimates of population differentiation. A recent standardized genetic differentiation measure of F_{ST} (Hedrick 2005) has been published to control for marker variability. Unfortunately, the standardized F_{ST} needs a locus-specific approach. We definitely have to extend this standardized measure to our type of MHC data to allow the direct comparison between MHC and microsatellites. We also have to notice that MHC alleles from a co-amplification of several loci may bring a potential issue concerning the proportion of null alleles. Therefore, to develop a locus-specific typing protocol would be the next step for further investigations with an approach cloning-sequencing to be accurate on MHC typing. Finally, we used a partial Mantel test to estimate the correlation between population pairwise differentiation at MHC and geographical distance while keeping constant differentiation at microsatellites. This method has recently been put forward in a work on great snipes (*Gallinago media*) by Ekblom *et al*. (2007). This test showed a strong pattern of isolation by distance for the MHC, independently of microsatellites, suggesting that increasing amount of differentiation with distance is not the sole consequence of neutral (drift) and/or demographic (migration) factors.

We show here that the comparison of populations at different geographical distances provided evidence in support of the two main hypotheses (i.e. balancing selection and diversifying selection) that have been put forward to explain the maintenance of MHC diversity. At small spatial scale, where selection pressures are presumably similar (homogenous selection pressures might arise from similar parasite abundance, diversity, genetic strains), balancing selection should reduce the amount of between-population diversity compared to the within-population diversity (Shierup *et al*. 2000; Bernatchez & Landry 2003). This, of course, should result in a weaker MHC population differentiation compared to neutral markers. This is indeed what we found for the populations located at the smallest geographical distance. On the contrary, populations located at large geographical distance are unlikely to experience similar parasite-exerted selection pressures. Genetic differentiation of parasite population is likely to increase with geographical distance, suggesting that local parasite strains might present different antigenic epitopes (Hill 1991). Environmental differences between sites might also be responsible for local variation in vector and parasite abundance or virulence (Bensch & Akesson 2003; Wood *et al*. 2007). Indeed, we found evidence for heterogeneity in parasite prevalence between populations located in different habitats (ranging from 4.5 to 81.6%, Loiseau *et al*. in preparation). All this should therefore translate into spatially heterogeneous selection favouring specific allelic lineages in different host populations, resulting in stronger population differentiation as compared to neutral markers. This pattern of MHC-based population differentiation has been recently reported in a few studies in fish, birds and mammals (Landry & Bernatchez 2001; Miller *et al*. 2001; Aguilar & Garza 2006; Bryja *et al*. 2007; Ekblom *et al*. 2007). Salmons (*Salmo salar*) have been particularly well studied with this respect because parasites can have a very local distribution (Bakke & Harris 1998), with heavily infected isolated populations, surrounded by healthy, noninfected ones. In agreement with this finding, salmon populations do show a stronger pattern of isolation by distance for MHC than for microsatellites (Landry & Bernatchez 2001). Similarly, Ekblom *et al*. (2007) reported significant genetic differentiation for the MHC of great snipes between two large regions, with no differentiation among populations within each of the regions. These results were interpreted as the consequence of habitat differences between regions (e.g. dry mountains vs. flood plains), that may pose variable selection pressures.

Diversifying selection is thought to promote local adaptation. The co-evolutionary nature of host–parasite relationships should produce MHC–pathogen associations that are population-specific. The occurrence of local adaptation in host–parasite system has received much theoretical attention in the last decade, with several life-history traits (i.e. relative dispersal tendencies of hosts and parasites, generation time; Gandon & Michalakis 2002; Morgan *et al*. 2005) that have been identified as key factors for the outcome of the interaction (parasite or host local adaptation). Several experimental studies have also addressed this issue with mixed results (review in Greischar & Koskella

2007). However, to our knowledge, there is no single study that has identified the genes possibly involved in host adaptation to parasitism and how these genes vary in space. Recently, Bonneaud *et al*. (2006) studied the association between MHC alleles and malaria prevalence in two house sparrow populations located at large geographical distance. In agreement with the hypothesis of population-specific function of MHC alleles, they found that two alleles were associated with a significant decrease in the likelihood of harbouring the infection. However, the allele that was functionally associated with malaria prevalence in one population was neutral in the second population and vice-versa. The screening of malaria infection of the 13 house sparrow populations studied here will provide a stronger test of the idea of population-specific MHC–parasite association, as suggested by the local adaptation hypothesis.

Acknowledgements

We thank Véronique Morin, Anne Oudin and Thomas Ruby for their assistance during the screening of MHC alleles and Pierre Federici and Arnaud Gutierrez during the screening of microsatellites. We thank all the volunteers from the CRBPO (National Museum of Natural History of Paris) for their help in the field and for providing samples: F. Baroteau, J. Birard, L. Brucy, G. Chaussi, J. Marion, F. Martayan, G. Massez, P. Ollivier, V. Ternois, and we also thank A. Lendvai and M. Giraudeau for captures on the field (Chizé) and B. Faivre for his help on Ouessant and Hoedic. We thank C. Gilbert for discussions and comments on earlier versions of the manuscript and we are also grateful to the editor Staffan Bensch and three anonymous referees who provided very useful comments on how to improve the manuscript. Financial support was provided by the Région Ile de France to C.L., R.J. and G.S. and the Agence Nationale de la Recherche ('EVO-INF-ECOL', NT05-2_44272) to G.S. and S.G.

References

Aguilar A, Garza JC (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.

Alcaide M, Edwards SV, Negro JJ, Serrano D, Tella JL (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.

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

Bakke TA, Harris PD (1998) Diseases and parasites in wild Atlantic salmon (*Salmo salar*) populations. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 247–266.

Beadell JS, Gering E, Austin J *et al*. (2004) Prevalence and differential host-specificity of two avian blood parasite genera in the Australo-Papuan region. *Molecular Ecology*, **13**, 3829–3844.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) *Genetix 4.05, Logiciel sous Windows pour la Génétique des Populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bensch S, Akesson S (2003) Temporal and spatial variation of hematozoans in Scandinavian willow warblers. *Journal of Parasitology*, **89**, 388–391.

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.

Bonneaud C, Sorci G, Morin V, Westerdahl H, Zoorob R, Wittzell H (2004) Diversity of MHC class I and IIB genes in house sparrows (*Passer domesticus*). *Immunogenetics*, **55**, 855–865.

Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G (2006) Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, **60**, 383–389.

Borghans JAM, Beltman JB, De Boer RJ (2004) MHC polymorphism under host–pathogen coevolution. *Immunogenetics*, **55**, 732–739.

Bowen L, Aldridge BM, DeLong R *et al*. (2006) MHC gene configuration variation in geographically disparate populations of California sea lions (*Zalophus californianus*). *Molecular Ecology*, **15**, 529–533.

Bryja J, Charbonnel N, Berthier K, Galan M, Cosson JF (2007) Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of vole. *Molecular Ecology*, **16**, 5084–5097.

Carrington M, Nelson GW, Martin MP *et al*. (1999) HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*, **283**, 1748–1752.

De Boer RJ, Borghans JAM, van Boven M, Kesmir C, Weissing FJ (2004) Heterozygote advantage fails to explain the high degree of polymorphism of the MHC. *Immunogenetics*, **55**, 725–731.

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

Edwards SV, Hedrick PW (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends in Ecology & Evolution*, **13**, 305–311.

Ekblom R, Saether SA, Jacobsson P *et al*. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology*, **16**, 1439–1451.

Excoffier L, Laval G, Schneider S (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Flint J, Bond J, Rees DC *et al*. (1999) Minisatellite mutational processes reduce F_{ST} estimates. *Human Genetics*, **105**, 567–576.

Gandon S, Michalakis Y (2002) Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*, **15**, 451–462.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486. Available from http://www.unil.ch/popgen/softwares/fstat.htm.

Goudet J, Raymond M, de Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1931–1938.

Greischar MA, Koskella B (2007) A synthesis of experimental work on parasite local adaptation. *Ecology Letters*, **10**, 418–434.

Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999) Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biological Journal of the Linnean Society*, **68**, 303–316.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hayashi K, Yoshida H, Nishida S *et al*. (2006) Genetic variation of the MHC DQB locus in the finless porpoise (*Neophocaena phocaenoides*). *Zoological Science*, **23**, 147–153.

Hedrick PW (1972) Maintenance of genetic variation with a frequency-dependent selection model as compared to the overdominant model. *Genetics*, **72**, 771–775.

Hedrick PW (1999) Balancing selection and MHC. *Genetica*, **104**, 207–214.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution*, **56**, 1902–1908.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Hedrick PW, Thomson G (1983) Evidence for balancing selection at HLA. *Genetics*, **104**, 449–456.

Hill AVS (1991) HLA associations with malaria in Africa: some implications for MHC evolution. In: *Molecular Evolution of the Major Histocompatibility Complex* (eds Klein J, Klein D), pp. 403–419. Springer, Berlin .

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

Hughes AL, Nei M (1992) Maintenance of MHC polymorphism. *Nature*, **355**, 402–403.

Klein J (1986) *The Natural History of the Major Histocompatibility Complex*. John Wiley & Sons, New York.

Kohn MH, Pelz H-J, Wayne RK (2000) Natural selection mapping of the warfarin-resistance gene. *Proceedings of the National Academy of Sciences, USA*, **97**, 7911–7915.

Landry C, Bernatchez L (2001) 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.

Larsson LC, Laikre L, Palm S, André C, Carvalho GR, Ryman N (2007) Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. *Molecular Ecology*, **16**, 1135–1147.

Li SH, Huang YJ, Brown JL (1997) Isolation of tetranucleotide microsatellites from the Mexican jay *Aphelocoma ultramarine*. *Molecular Ecology*, **6**, 499–501.

Loiseau C, Zoorob R, Garnier S *et al*. (2008) Antagonistic effects of a MHC class I allele on malaria-infected house sparrows. *Ecology Letters*, **11**, 258–265.

Lukas D, Bradley BJ, Nsubuga AM, Doran-Sheehy D, Robbins MM, Vigilant L (2004) Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. *Molecular Ecology*, **13**, 3389–3402.

McClelland EE, Penn DJ, Potts WK (2003) Major histocompatibility complex heterozygote superiority during co-infection. *Infection and Immunity*, **71**, 2079–2086.

Meyer D, Thomson G (2001) How selection shapes variation of the human major histocompatibility complex: a review. *Annals of Human Genetics*, **65**, 1–26.

Meyer-Lucht Y, Sommer S (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). *Molecular Ecology*, **14**, 2233–2243.

Miller KM, Withler RE (1997) MHC diversity in Pacific salmon: Population structure and trans-species allelism. *Hereditas*, **127**, 83–95.

Miller KM, Kaukinen KH, Beacham TD, Withler RE (2001) Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. *Genetica*, **111**, 237–257.

Morgan AD, Gandon S, Buckling A (2005) The effect of migration on local adaptation in a coevolving host–parasite system. *Nature*, **437**, 253–256.

Myers RM, Maniatis T, Lerman LS (1987) Detection and localization of single base changes by denaturant gradient gel electrophoresis. *Methods in Enzymology*, **155**, 501–527.

Penn DJ, Damjanovich K, Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences, USA*, **99**, 11260–11264.

Piertney SB (2003) Major histocompatibility complex B-LB gene variation in red grouse *Lagopus lagopus scoticu*s. *Wildlife Biology*, **9**, 251–259.

Primmer CR, Møller AP, Ellegren H (1996) New microsatellites from the pied flycatcher *Ficedula hypoleuca* and the swallow *Hirundo rustica* genomes. *Hereditas*, **124**, 281–283.

Richardson DS, Jury FL, Dawson DA, Salgueiro P, Komdeur J, Burke T (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology*, **9**, 2226–2231.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2008) GenePop'007: a complete reimplementation of the GenePop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

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.

Shierup MH, Vekemans X, Charlesworth D (2000) The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genetical Research*, **76**, 51–62.

Slatkin M (1995) Hitchhiking and associative overdominance at a microsatellite locus. *Molecular Biology and Evolution*, **12**, 473–480.

Summers-Smith JD (1988) The sparrows. *T. &* A. D. Pyser Ltd, Calton.

Svensson LME, Ruegg KC, Sekercioglu CH, Sehgal RNM (2007) Widespread and structured distributions of blood parasite haplotypes across a migratory divide of the Swainson's trush (*Catharus ustulatus*). *Journal of Parasitology*, **93**, 1488–1495.

Valkiunas G, Iezhova TA, Bolshakov CV, Kosarev V (2006) Blood parasites of the house sparrow *Passer domesticus* from northwestern Russia, with remarks on trends of global geographical distribution in this bird. *Journal of Natural History*, **40**, 1709–1718.

Vasemägi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, **22**, 1067–1076.

Weir BS (1996) *Genetic Data Analysis II*. Sinauer Associates, Sunderland, Massachusetts.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **43**, 1349–1368.

Westerdahl H, Waldenstrom J, Hansson B, Hasselquist D, von Schantz T, Bensch S (2005) Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1511–1518.

Wood MJ, Cosgrove CL, Wilkin TA, Knowles SCL, Day KP, Sheldon BC (2007) Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. *Molecular Ecology*, **16**, 3263–3273.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Allele frequencies of the different band DGGE Mhc class I in the 13 populations

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.