Molecular Ecology (2007) **16**, 5084–5097 doi: 10.1111/j.1365-294X.2007.03584.x

Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles

J. BRYJA,*† N. CHARBONNEL,* K. BERTHIER,* M. GALAN* and J.-F. COSSON*
**Centre de Biologie et Gestion des Populations (UMR 22), INRA, Campus International de Baillarguet, CS 30016, 34988 Montferrier sur Lez Cedex, France, †Department of Population Biology, Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, 67502 Studenec 122, Czech Republic*

Abstract

Host–pathogen interactions are of particular interest in studies of the interplay between population dynamics and natural selection. The major histocompatibility complex (MHC) genes of demographically fluctuating species are highly suitable markers for such studies, because they are involved in initiating the immune response against pathogens and display a high level of adaptive genetic variation. We investigated whether two MHC class II genes (*DQA1*, *DRB*) were subjected to contemporary selection during increases in the density of fossorial water vole (*Arvicola terrestris*) populations, by comparing the neutral genetic structure of seven populations with that estimated from MHC genes. Tests for heterozygosity excess indicated that *DQA1* was subject to intense balancing selection. No such selection operated on neutral markers. This pattern of selection became more marked with increasing abundance. In the low-abundance phase, when populations were geographically isolated, both overall differentiation and isolation-by-distance were more marked for MHC genes than for neutral markers. Model-based simulations identified *DQA1* as an outlier (i.e. under selection) in a single population, suggesting the action of local selection in fragmented populations. The differences between MHC and neutral markers gradually disappeared with increasing effective migration between sites. In the high-abundance year, *DQA1* displayed significantly lower levels of overall differentiation than the neutral markers. This gene therefore displayed stronger homogenization than observed under drift and migration alone. The observed signs of selection were much weaker for *DRB*. Spatial and temporal fluctuations in parasite pressure and locus-specific selection are probably the most plausible mechanisms underlying the observed changes in selection pattern during the demographic cycle.

Keywords: *Arvicola terrestris*, balancing selection, local adaptation, MHC, population cycles

Received 25 July 2007; revision accepted 14 September 2007

Introduction

Evolutionary biologists are increasingly focusing their interest on the interplay between population dynamics and natural selection (Saccheri & Hanski 2006). Molecular markers linked to host–pathogen interactions are of great potential use for increasing our understanding of these processes (reviewed in Charbonnel *et al*. 2006). The genes of the major histocompatibility complex (MHC), encoding proteins that bind pathogen-derived foreign peptides to T cells to initiate an immune response, provide a classic example of such markers. MHC genes may be extremely polymorphic due to their role in disease resistance; high levels of variation in these genes are therefore thought to be adaptive in natural populations (Apanius *et al*. 1997). There is growing body of empirical data from natural populations suggesting that MHC genes have been subject to positive selection over an evolutionary timescale, resulting in high ratio of nonsynonymous to synonymous

Correspondence: Josef Bryja, Czech Republic. Fax: +420 568423121; E-mail: bryja@brno.cas.cz

nucleotide substitutions and the sharing of MHC sequences between species or genera (reviewed by Bernatchez & Landry 2003; Garrigan & Hedrick 2003).

It is more difficult to demonstrate selection acting on MHC genes in contemporary generations (Piertney & Oliver 2006). One of the promising approaches is the comparison of population divergence levels based on neutral and MHC markers (Schierup *et al*. 2000). Balancing selection should increase within-population MHC diversity relative to total diversity. An incoming migrant MHC allele absent in the resident population may be selected for, increasing its chance of invasion over those for neutral alleles and therefore increasing its effective migration rate. It has been suggested that population subdivision — as measured by conventional statistics such as F_{ST} — should be much less apparent at genes under balancing selection than at neutral genes (Schierup *et al*. 2000; Muirhead 2001). Conversely, if the selective advantage of MHC alleles varies across populations due to the heterogeneity of selection pressures (e.g. local variation in parasite community structure, etc.), then particular subsets of alleles should be favoured locally (i.e. local adaptation) but have limited success in invading other populations, thus resulting in higher levels of population differentiation than would be expected under neutrality (Muirhead 2001).

Only a limited number of empirical studies comparing neutral and MHC marker variation in wild populations have been performed. In the bighorn sheep (*Ovis canadensis*), mean F_{ST} values were found to be very similar for MHC and microsatellite markers, suggesting that nonselective forces are of the utmost importance in determining the spatial pattern of differentiation at the MHC loci (Boyce *et al*. 1997; Gutierrez-Espeleta *et al*. 2000). Major effects of genetic drift and demographic factors on the variation and distribution of MHC alleles have also been found in other mammals (e.g. Seddon & Ellegren 2004; Worley *et al*. 2006) and birds (Miller & Lambert 2004). In contrast, large differences in the extent of genetic differentiation at MHC and microsatellite loci have been observed between wild Atlantic salmons from different sites within a river, supporting the hypothesis of local adaptation on a small geographical scale (Landry & Bernatchez 2001; Aguilar & Garza 2006).

The intensity and target of selection may vary in time and space due to the direct effects of variation in the prevalence of parasites or the indirect effects of variations in the fitness costs of infection with resource availability. Fluctuating selection has been explored theoretically and validated as a concept and several empirical studies have also provided support for this idea (Charbonnel & Pemberton 2005 and references therein). Suitable models for the analysis of spatial and temporal fluctuations in selection include populations of arvicoline rodents displaying regular multi-annual changes in density. Water vole populations in the Jura Mountains, on the border between France and Switzerland, display cyclic dynamics with densities varying from less than one to several hundred individuals per hectare over 5–6 years (Saucy 1994; Giraudoux *et al*. 1997). We recently reported high levels of neutral genetic diversity in these populations, despite the small numbers of animals periodically observed (Berthier *et al*. 2005; Berthier *et al*. 2006). In the most recent of these two studies, seven populations were monitored at a very fine scale over 3 years marking the transition from a patchy structure (demes) at low density to a continuous population at high density. At the end of the low-density phase, demes displayed the effects of genetic drift due to their small size and geographical isolation. This genetic drift led to a loss of local genetic diversity and spatial differentiation among demes. This situation was counterbalanced during the phase of increasing population density by the spatial expansion of demes and an increase in effective migration between differentiated demes. Despite significant changes in the genetic structure of populations with changes in density, no evidence for a departure from mutation-drift equilibrium was observed (Berthier *et al*. 2006).

In this study, we expand the data set reported by Berthier *et al*. (2006), by adding new data on the genetic polymorphism of two MHC class II genes (*DQA1* and *DRB*) in these populations. This study had two main aims. The first was to describe genetic variation in the *DRB* sequences of the water vole, with the aim of identifying a signature of historical selection. The second aim was to compare the genetic structures of neutral genes (microsatellites) with those of genes thought to be under selection (MHC genes), with the aim of identifying signals of contemporary selection acting on MHC genes. Under balancing selection, the level of differentiation at MHC genes between populations should be significantly lower than that observed for neutral markers. Alternatively, if selection pressure varies spatially, then evolution should lead to local differentiation and the MHC genes should display greater population subdivision than neutral nuclear markers. We also investigated whether the increase in effective migration due to an increase in density (Berthier *et al*. 2006) outweighed the effects of selection (either balancing or diversifying). The factors underlying the cycling of arvicoline rodent populations are unknown, but studies on these populations have recently suggested that pathogens may play a key role, as in other cyclic vertebrates (O'Brien *et al*. 2006; Redpath *et al*. 2006). If we assume this to be the case, then genes encoding proteins involved in the immune response should be under strong selection pressure, providing us with a unique opportunity to study the balance between drift, migration and selection on MHC genes in natural vertebrate populations and changes in this balance over time.

Materials and methods

Populations

Genetic variability was analysed at seven sites (1.6–12.6 km apart) sampled during the phase of density increase of the population cycle of the water vole in the Franche-Comté region in eastern central France. Each site covered 1–2 ha of large, open meadows and the animals found within this area are referred to hereafter as a 'population'. At each site, 80 traps were set a minimum of 5 m apart (i.e. in different vole colonies) to avoid the capture of closely related individuals. The water vole populations were sampled in October 2001, October 2002, and October 2003. In total, 591 individuals were collected and analysed (15–37 individuals per population sample; the sample size did not depend on population density). Additional details on sampling and abundance surveys in the field have been published elsewhere (Berthier *et al*. 2006).

The water vole displays pronounced, regular, multi-annual changes in population density in the study area. The abundance of these rodents was very low between the end of the last increase in population in 1998 and the end of 2001. During the first sampling in October 2001, vole colonies were detected at only seven sites in the entire study area, despite intensive field searches. Between October 2001 and October 2003, the abundance of voles significantly increased. In October 2003, abundance indices were high and displayed little variation between sites, indicating a uniform abundance of voles across the whole sampling area (see Berthier *et al*. 2006).

Genotyping at microsatellite and MHC loci

Genomic DNA was extracted from tissue samples (toe or end of the tail) using the QIAmp Tissue Kit (QIAGEN) according to the manufacturer's instructions. All individuals were typed at 17 autosomal microsatellite loci (see details on the loci analysed and the genotyping methods used in Berthier *et al*. 2006) and for two MHC class II genes (*DQA* and *DRB* homologues). We previously showed (Bryja *et al*. 2006) that the *DQA* gene is duplicated in water voles (*DQA1* and *DQA2*). Based on genotype composition, we were able to distinguish alleles at particular loci and found that the *DQA2* gene is not present in all individuals (see details in Bryja *et al*. 2006). We therefore excluded this locus from further analyses. The second exons of the *DQA1* and *DRB* genes, encoding the protein region thought to be involved in peptide binding (the antigen-binding site, ABS) were genotyped in all individuals by polymerase chain reaction (PCR) and single-strand conformation polymorphism (SSCP) analysis. Details concerning the genotyping of *DQA1* are provided in a previous publication (Bryja *et al*. 2006). A fragment of exon 2 of the *DRB* gene (216 bp) was amplified with the primers JS1 and JS2 (Schad *et al*. 2004) in the following mixture: 2 mM $MgCl_2$, 0.3 μM of each primer, 100 μM dNTPs, 0.5 U *Taq* polymerase (QIAGEN) in the appropriate 1× PCR buffer, and 1 μL of extracted DNA. De-ionized water was added to a give a final reaction volume of 10 μL. Samples were subjected to initial denaturation at 96 °C for 2 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 57 °C for 30 s and extension at 72 °C for 1 min, with a final extension phase at 72 °C for 10 min. The PCR fragments of both genes were fluorescently labelled using primers bound to different dyes (6'FAM, HEX, or NED) and SSCP analysis was then carried out on a MEGABACE 1000 DNA Analysis System (Amersham Biosciences; see Bryja *et al*. 2005 for details). Briefly, PCR products were diluted in 10 mM Tris-HCl (pH 8.5), mixed with internal ROX-labelled marker (Amersham Biosciences), denatured, snap-cooled on ice and injected into the capillaries. We used nondenaturing 6% low-to-medium molar weight linear polyacrylamide (Genomac) as a sieving matrix and MEGABACE 1× running buffer (Amersham Biosciences) supplemented with 10% glycerol as a running buffer for capillary electrophoresis in 27 °C. The electrophoretograms were aligned and analysed with MEGABACE Genetic Profiler 1.5 software (Amersham Biosciences). This made it possible to distinguish reliably the SSCP patterns of all sequences of *DQA1* (Bryja *et al*. 2006) and *DRB* identified by cDNA cloning and sequencing.

cDNA cloning and sequencing

The second exon of *DRB* was characterized by the cloning and sequencing of particular cDNAs as described in detail in Bryja *et al*. (2006). Briefly, we selected individuals with diverse SSCP patterns, and extracted total RNA from their spleens, which had been preserved in RNAlater (Ambion). We used 1 μL of extracted RNA and poly(T) primer for reverse transcription. Exon 2 of the *DRB* gene was partially amplified from cDNA, using the primers JS1 and JS2. Purified PCR products were cloned with the pGEM-T Cloning Kit (Promega). We screened about 12 clones per individual by PCR and SSCP analyses, in which the SSCP pattern of individual clones was compared with those of PCR products from a particular individual. This made it possible to exclude clones containing PCR artefacts (clones with an SSCP pattern not observed in the SSCP pattern of the corresponding individual). Plasmid DNA was extracted with the QIAprep Spin Miniprep Kit (QIAGEN) and inserts were sequenced by DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences).

Sequence analysis

The partial *DRB* exon 2 sequences obtained (171 bp without primers) were edited and aligned in BIOEDIT Sequence

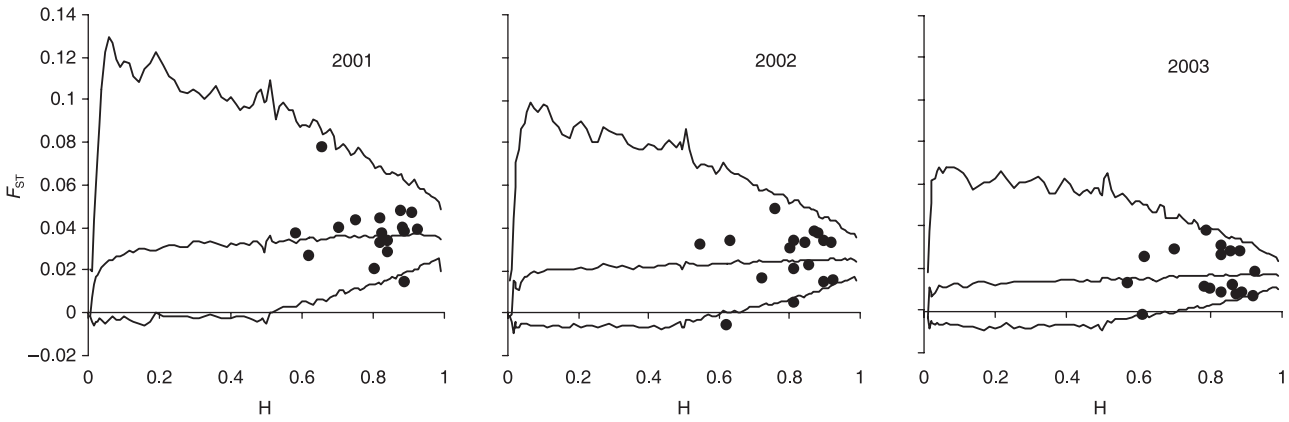

Fig. 1 F_{ST} values estimated from 17 microsatellites (black circles) plotted against heterozygosity. Lines denote the 0.975, 0.500 and 0.025 quantiles of the distribution of F_{ST} against heterozygosity estimated from 50 000 simulations in FDIST2.

Alignment Editor (Hall 1999), using CLUSTAL W Multiple Alignment. As a large proportion of the sequences obtained could potentially have corresponded to artefacts of PCR amplification (Bryja *et al*. 2005), we considered a new sequence variant to be a new allele only if it was obtained from two separate PCRs for the same individual, or by PCR from at least two different individuals. Water vole allele nomenclature is as described by Klein *et al*. (1990). We started numbering alleles from Arte-*DRB**06, because five sequences (different from ours) were recently published by Oliver & Piertney (2006). The 16 partial exon 2 sequences from the water vole *DRB* gene are available from GenBank, under Accession nos EF660500– EF660515.

Molecular evolutionary analyses were carried out using MEGA 3.0 (Kumar *et al*. 2004) to calculate the relative frequencies of synonymous (dS) and nonsynonymous (dN) substitution by the method of Nei & Gojobori (1986) with Jukes & Cantor (1969) correction for multiple substitutions. Standard errors were estimated by bootstrapping with 1000 replicates. ABS sites were identified, based on data for the human *DRB* molecule (Brown *et al*. 1993), at amino acids 26, 28, 30, 32, 37, 38, 47, 56, 58, 60, 61, 65, 68, 70, 71, 74, 78. The probability that dN = dS was determined with a *Z*-test (Nei & Kumar 2000) for ABS and non-ABS sites, separately.

More sophisticated analyses of selection on MHC genes over the evolutionary history of the species were performed by likelihood-ratio modelling, using CodeML, from the PAML 3.15 program suite (Yang 1997). We used maximum-likelihood models accounting for heterogeneity in site partitions using different ω (= dN/dS) parameters for the partitions. We used random-site models (Yang *et al*. 2000) assuming several heterogeneous site classes with different ω parameters without prior knowledge of the class to which each site belonged (see also Bryja *et al*. 2006). We used the M0, M1a, M2a, M3, M7, and M8 models (Yang *et al*. 2000; details concerning models M1a and M2a in Yang *et al*. 2005). We carried out three likelihood-ratio tests comparing nested models (M0 vs. M3, M1a vs. M2a, and M7 vs. M8), in which the alternative models (M2a, M3, and M8) suggest the presence of sites with ω> 1. All three tests may therefore be considered tests for positive selection (Yang *et al*. 2000). Posterior probabilities for site classes were calculated by the Bayes empirical Bayes (BEB) method in models M2a and M8. If the posterior means of ω for some site classes are > 1, the sites concerned are probably under diversifying selection (Yang *et al*. 2005).

Population genetic data analysis

A model-based approach was used to determine whether the 17 microsatellite loci used were really neutral —, i.e. not affected by selection. We used FDIST 2 (Beaumont & Balding 2004) to perform coalescent simulations, using a symmetric island model of population structure and an infinite alleles mutation model. Sample sizes were set to the median and median estimates of F_{ST} for each data set corresponding to a particular year were used to simulate the data. Coalescent simulations were performed to generate 50 000 paired values of F_{ST} and heterozygosity, and these results were used to calculate the 0.975, 0.500 and 0.025 quantiles of the distribution of F_{ST} as a function of heterozygosity (Beaumont & Nichols 1996). Loci with unusually high or low F_{ST} estimates are regarded as potentially under selection. We identified one microsatellite locus in 2001 and two loci in 2002 and 2003 as significant outliers. However, the values for these loci were very close to 0.025 quantile of the F_{ST} distribution, and we therefore considered them neutral for all analyses (Fig. 1).

The method described by Beaumont & Nichols (1996) detects outlier loci with unusually high global F_{ST} values, but pairwise comparisons of populations are more likely to be efficient in conditions of selection acting at a local scale (Vitalis *et al*. 2001). We therefore carried out pairwise

analyses between all the populations for the same year, using DETSEL 1.0 (Vitalis *et al*. 2002). For each pair of populations (*i*, *j*), and for all loci, we calculated F_i and F_j and generated the expected joint distribution of F_i and F_j by performing 10 000 coalescent simulations for a given set of nuisance parameter values ($N_0 = 500$, $\mu = 0.0015$, $N_e = 10\,000$, $T_0 = 10$). A scatter plot of the observed data points (i.e. pairs of F_i and F_j values) was superimposed over an outline of the 95% probability region for identifying outlier loci (Vitalis *et al*. 2001). These pairwise analyses frequently detected three microsatellite loci as lying outside the calculated distribution: AT23 (in 31 of 63 pairwise analyses), AT24 (in 15), and AT3 (in 20). Based on these tests, we decided to exclude these loci from subsequent comparisons of neutral and MHC marker variability.

Genotypic linkage disequilibrium between 14 neutral microsatellite loci, *DQA1*, and *DRB* was tested for each sampling site and for each year, by exact tests, using the Markov chain methods in GENEPOP 3.3 (Raymond & Rousset 1995). We corrected for multiple testing using the false discovery rate (FDR) approach, as described by Benjamini & Hochberg (1995).

Deviations from Hardy–Weinberg proportions were quantified for each population, site, year, and marker by the unbiased Wright inbreeding coefficient F_{IS}, estimated as described by Weir & Cockerham (1984) using GENETIX 4.03 (Belkhir *et al*. 2001). We assessed the significance of the deviation from zero of the F_{IS} estimate by 1000 permutations in the same program. The effects of the type of the marker (microsatellite, *DQA1* and *DRB*), and locus (nested in the type of the marker) on the arcsin-transformed F_{IS} estimates were analysed using general linear models with a normal error structure in JMP 3.2 (SAS Institute). The Tukey-Kramer HSD test was used for *post hoc* comparisons of parameter means when main effects were significant.

Excess heterozygosity with respect to that expected at mutation-drift equilibrium for the number of alleles present may indicate a genetic bottleneck or balancing selection (Cornuet & Luikart 1996). We estimated the deviation of gene diversity from mutation-drift equilibrium at the local scale, using BOTTLENECK 1.2.02 (Piry *et al*. 1999), with a generalized mutation model (GSM) used for microsatellites (as the most appropriate; Estoup *et al*. 2001) and an infinite alleles mutation model (IAM) used for *DQA1* and *DRB*.

The genetic differentiation between sampling sites was quantified, for each sampling session and marker type, by calculating pairwise estimators of F_{ST}, as described by Weir & Cockerham (1984), in FSTAT. We identified 95% confidence intervals of 'neutral' F_{ST} by bootstrapping over 14 microsatellite loci in FSTAT. F_{ST} estimates derived from genes under balancing selection should be relatively independent of population size, whereas microsatellites display a significant decrease in F_{ST} with increasing density (Berthier *et al*. 2006). We therefore investigated the decrease in F_{ST} over time, based on 14 microsatellites, *DQA1* and *DRB*, between pairs of sampling sessions, using one-sided permutation tests implemented in FSTAT. Isolation by distance was analysed, for each sampling session and marker type, by regressing pairwise estimates of $F_{ST}/(1 - F_{ST})$ against ln-distance between sampling sites (Rousset 1997). Mantel tests were used to test the correlation between matrices of genetic differentiation and Euclidean distances between sampling sites, by 5000 permutations in GENEPOP. We also calculated confidence intervals for slope of the regression line, based on 14 microsatellites, by bootstrapping over loci to obtain approximate bootstrap confidence intervals (ABC intervals, Di Ciccio & Efron 1996).

There are two possible sources of bias in comparisons of genetic differentiation between populations based on microsatellites and MHC genes. First, null alleles (NA) are known to overestimate the genetic differentiation between populations. We corrected for this effect, using a method implemented in FREENA (available from http://www.montpellier.inra.fr/URLB) for estimating F_{ST} at a locus thought to present NA. This method efficiently corrects F_{ST} estimates for the positive bias introduced by the presence of NA (Chapuis & Estoup 2007). Second, it is often difficult to interpret genetic differentiation values because of their dependence on the level of genetic variation at particular loci. We thus used a standardized measure of genetic differentiation, as proposed by Hedrick (2005), bringing values into the same range (0–1) for all levels of genetic variation. The standardized estimator was calculated by dividing the estimated value by the maximum value obtained (using RECODEDATA freely available from http://www.bentleydrummer.nl/software). F_{ST} estimates corrected for the presence of NA F_{STcorr} and for the level of genetic variation F'_{ST} were used for subsequent analyses of isolation by distance, as described above.

Finally, we evaluated the presence and intensity of contemporary selection on MHC genes, using coalescent simulations in FDIST 2 and DETSEL, as described above. Both analyses included data on 14 neutral microsatellites, *DQA1* and *DRB* and were performed separately for each year. Loci detected as outliers in these analyses are usually considered to be affected by selection.

Results

Variation of DRB *sequences and historical selection*

By cloning and sequencing cDNAs from seven populations of water voles from the canton of Nozeroy, we identified 16 transcribed alleles of exon 2 of the *DRB* gene that could be translated to give unique amino-acid sequences (Fig. 2). In 16 sequences, 46 of the 171 nucleotide positions were variable, with 38 of these positions minimally informative.

```
                * * * *    **        *        * * **   *  * **  *   *
Arte-DRB*06  QRVRFLVRDIYNREEVVRFDSDVGEFRAVTELGRPDAEYFNSLKDFLEQKRAEVDTY
Arte-DRB*07  ........Y......F.........H.........GI..N...R.EL..RT..A...V
Arte-DRB*08  ....Y....F.....Y.....E...Y.........RS...W..Q..V..R....I..V
Arte-DRB*09  ......D.YFF....Y.................GI..NL..Q..I..RL......V
Arte-DRB*10  ......D.Y.H.Q..F.........H.........GI...W..Q..I..R...A....
Arte-DRB*11  ....Y...V.H....Y.........Y.........GI...L......M.....A....
Arte-DRB*12  ......D.Y.H.Q..FM........H.........GI...W..Q..I..R...A....
Arte-DRB*13  ....L.D.YFF.Q..YL.....................Y......M.DR..A....
Arte-DRB*14  ....Y.......L.............Y.............W.....A..R...A...V
Arte-DRB*15  ....L.D.V......Y.................RS...R......I..L..S....
Arte-DRB*16  ......E.YF.....Y.........Y.E......GI..NL..R.EL.......I...
Arte-DRB*17  ....Y.......L............Y.........GI...W.....A..R...A...V
Arte-DRB*18  ......D.YFF....Y.........Y.........RS...W..Q.EL..NR......V
Arte-DRB*19  ....L.....................................................I..F
Arte-DRB*20  ......D.Y.H.Q..F.........H.........GI...W..Q.EL..R...A....
Arte-DRB*21  ....L.D.YFF.Q..Y.....................Y......M.DR..A....
```

Fig. 2 Alignment of 16 expressed amino-acid sequences (corresponding to amino-acid residues 22–78 from Brown *et al*. 1993) of the partial *DRB* exon 2 from water voles. Dots indicate identity with the nucleotide sequence of Arte-*DRB**06. Asterisks identify sites in the putative peptide-binding region of the human *DRB* molecule (according to Brown *et al*. 1993).

Table 1 Results of maximum-likelihood models for exon 2 of the DRB gene in water voles

Model code	P	Likelihood	Parameter estimates	Positively selected sites
M0 (one ratio)	1	–858.410	w = 1.782	None
M1a (nearly neutral)	1	–800.013	p0 = 0.598 (p1 = 0.402)	Not allowed
M2a (positive selection)	3	–782.798	p0 = 0.524, p1 = 0.123 (p2 = 0.353), w2 = 8.274	**26**, **28**, *32*, **37**, *47*, **56**, **57**, **61**, *64*, **67**, **70**, **71**, *74*, **75**, **78**
M3 (discrete)	5	–767.079	p0 = 0.605, p1 = 0.360 (p2 = 0.035), w0 = 0.092, w1 = 9.651, w2 = 18.297	Not analysed
M7 (beta)	2	–800.186	p = 0.016, q = 0.021	Not allowed
M8 (beta and omega)	4	–782.741	p0 = 0.642 (p1 = 0.358), p = 0.005, q = 0.020, w = 8.934	**26**, **28**, *30*, *31*, **32**, *34*, **37**, **47**, **56**, **57**, *60*, **61**, **64**, **67**, *68*, **70**, **71**, **74**, **75**, **78**

An alignment of 16 *DRB* sequences (171 bp) from water vole was used as the input for CodeML (included in the PAML 3.14 program suite). P is the number of parameters in the ω distribution, ω is the selection parameter and p_n is the proportion of sites falling into the ω_n site class. For models M7 and M8, p and q are the shape parameters of the β function. Positively selected sites were identified in models M2a and M8 by the Bayes empirical Bayes procedure (Yang *et al*. 2005). Sites inferred under selection at the 99% level are listed in bold, and those inferred at the 95% level are shown in italics.

No insertions or deletions (indels) were detected. Pairwise divergences ranged from 0.6 to 16.9% nucleotide substitutions. The hypothesis that dN = dS was significantly rejected by the *Z*-test for ABS (Z = 4.51, $P < 0.0001$) while it was not true for non-ABS (Z = 1.82, $P = 0.07$). The ratio of nonsynonymous to synonymous substitutions was markedly higher at the ABS (dN = 0.42 ± 0.09 vs. dS = 0.08 ± 0.02) than at non-ABS (dN = 0.06 ± 0.02 vs. dS = 0.02 ± 0.02), suggesting intensive positive selection on ABS.

Maximum likelihood-based random-site models applied to the 171 bp fragment of exon 2 of *DRB* provided additional evidence of historical positive selection acting on the *DRB* gene (Table 1). Models allowing for the inclusion of positive selection — M2a, M3 and M8 — fitted the data significantly better than simpler models — M1a, M0 and M7 ($P < 0.001$ in all three likelihood-ratio tests). BEB procedures demonstrated that selection pressure on different sites in exon 2 was highly variable and identified 15 or 20 sites under positive selection, using models M2a and M8, respectively. The sites under positive selection almost matched the ABS in the human *DRB* molecule.

Neutral and MHC marker variation within populations

No clear evidence of linkage disequilibrium between any of the 16 loci analysed (14 microsatellites, *DQA1* and *DRB*) was found. However, 13 (0.516%) of 2520 pairwise comparisons (120 pairs of loci × 21 pops) remained significant after FDR correction. If these significant comparisons corresponded to stochastic (random) events, 0.516% of the 120 pairs of loci would correspond to 0.6 cases. Nevertheless, four of the 13 significant comparisons were between *DQA1* and *DRB*, suggesting linkage between these two genes.

Table 2 Estimates of F_{IS} and the probabilities associated with rejection of the mutation-drift equilibrium hypothesis (Wilcoxon test) in BOTTLENECK for *DQA1*, *DRB*, and 14 microsatellite loci (14MS). F_{IS} estimates significantly different from zero (1000 permutations) are shown in bold. Probability values for the Wilcoxon test — $P < 0.05$ (in bold) — for *DQA1*, *DRB*, and 14 microsatellites suggest excess heterozygosity due to a demographic bottleneck or balancing selection

			F_{IS}			Mutation-drift equilibrium			
Year	Population	*N*	*DQA1*	*DRB*	MS	*DQA1*	*DRB*	14MS	Significant microsatellite loci
2001	01	27	–0.033	–0.012	0.011	0.129	**0.033**	0.548	AV3 ($P = 0.042$); AVM12 ($P = 0.011$)
	02	28	–0.013	–0.070	**0.034**	0.271	**0.000**	0.714	AT19 ($P = 0.005$)
	03	22	–0.008	**0.295**	**0.057**	**0.024**	**0.019**	**0.042**	AT9 ($P = 0.021$); AV11 ($P = 0.017$); AVM9 ($P = 0.046$)
	04	15	–0.268	–0.022	**–0.059**	0.353	0.144	**0.024**	
	05	27	–0.068	**0.250**	–0.010	**0.020**	0.265	0.749	AT2 ($P = 0.019$); AV3 ($P = 0.028$); AVM8 ($P = 0.001$)
	06	33	–0.023	0.104	–0.009	0.103	0.427	0.583	AT2 ($P = 0.016$)
	07	35	0.024	**0.231**	**0.044**	**0.048**	0.418	0.541	AT19 ($P = 0.039$); AV10 ($P = 0.041$)
2002	01	28	–0.022	**0.172**	**0.067**	**0.015**	**0.003**	0.295	
	02	26	0.133	**0.161**	–0.001	**0.002**	0.291	0.670	AV15 ($P = 0.013$)
	03	27	0.149	0.063	0.000	0.066	0.185	0.685	AV3 ($P = 0.010$); AT13 ($P = 0.020$); AT9 ($P = 0.048$);
	04	22	–0.109	0.052	0.016	0.101	0.190	0.903	
	05	23	0.119	**0.303**	0.007	0.077	**0.027**	0.173	AT19 ($P = 0.041$); AT25 ($P = 0.048$); AVM10 ($P = 0.018$); AVM8 ($P = 0.009$)
	06	27	0.085	**0.248**	–0.018	**0.001**	0.465	**0.042**	AT13 ($P = 0.022$)
	07	30	0.121	**0.277**	0.007	**0.023**	0.279	0.626	AV11 ($P = 0.036$)
2003	01	34	0.051	**0.132**	0.017	**0.006**	0.063	0.626	
	02	30	–0.065	**0.161**	–0.004	0.079	0.116	0.267	AT15 ($P = 0.030$)
	03	37	–0.077	–0.106	**0.049**	**0.010**	0.347	0.808	AV15 ($P = 0.046$)
	04	28	–0.156	0.040	0.019	**0.011**	0.247	0.626	
	05	30	–0.146	–0.011	0.006	**0.020**	**0.000**	0.217	AV11 ($P = 0.046$); AVM9 ($P = 0.035$)
	06	25	**0.212**	0.049	0.008	**0.004**	0.090	0.058	AV3 ($P = 0.040$)
	07	27	0.176	0.021	0.009	**0.001**	0.456	1.000	AT25 ($P = 0.029$); AVM9 ($P = 0.001$)

F_{IS} estimates ranged between –0.268 and 0.494. The variability in F_{IS} (Table 2) could be accounted for by the type of marker ($F_{2,335} = 10.265$, $P < 10^{-4}$). *Post hoc* Tukey-Kramer tests revealed that F_{IS} estimates were significantly higher at the *DRB* locus than at microsatellite loci or *DQA1*, indicating a heterozygote deficit at the *DRB* locus. Overall, these results suggest the presence of null alleles in *DRB*, the frequency of which was estimated, with FREENA, to vary from 0 to 0.13 depending on year and population.

Significant deviations of genetic diversity from mutation-drift equilibrium were observed in BOTTLENECK tests, for at least at one locus, in all but one population sample (Table 2). Overall, *DQA1* gave the largest number of significant tests (13 tests in all years vs. six for *DRB* and zero to four for particular microsatellite loci; Fig. 3), and these tests were generally supported by more significant *P* values than those for other loci. The number of significant tests increased over time for *DQA1* and no such pattern was observed for *DRB* or microsatellites (Table 2, Fig. 3). We found no consistent evidence of genetic bottlenecks in our populations when using data from the whole microsatellite set (only two tests of borderline significance in 2001 and one in 2002; Table 2).

Neutral and MHC marker variation between populations

The pattern of population differentiation varied significantly during the period studied (Table 3). Overall, F_{ST} estimates

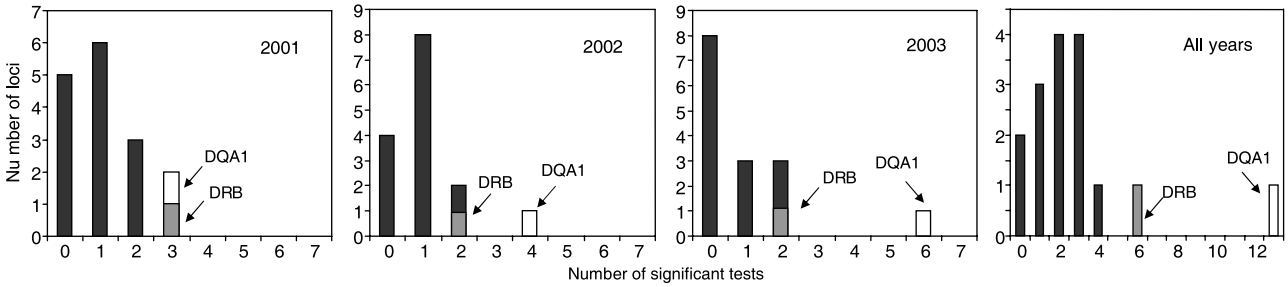

Fig. 3 Number of statistical tests at individual loci indicating departure from the mutation-drift equilibrium (excess heterozygosity, $P < 0.05$) in different years and for all years together.

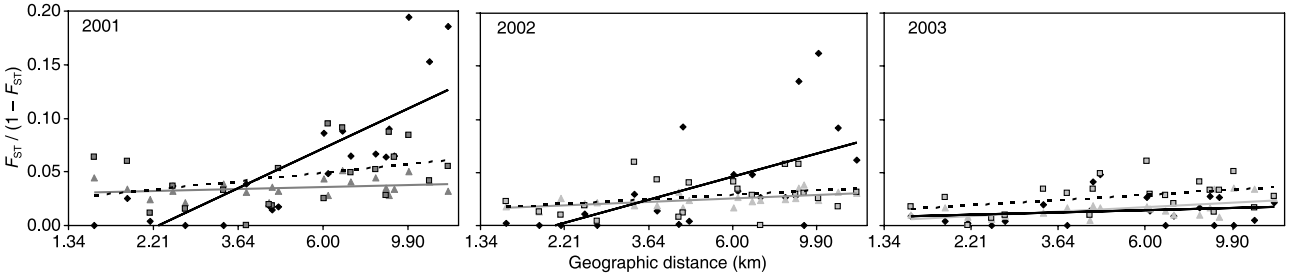

Fig. 4 Correlation between genetic $F_{ST}/(1 - F_{ST})$ and geographical distance (log scale) for three samplings of MHC and neutral markers: *DQA1* (diamond and solid black line), *DRB* (square and dashed black line), 14 microsatellites (triangle and solid grey line).

Table 3 Overall F_{ST} estimates and isolation by distance characteristics — probability (IBD *P*), correlation coefficient (R), intercept and slope — calculated for each marker type in 3 years. 95% confidence intervals (CI) for overall F_{ST} and the slope of the IBD were obtained by bootstrapping over 14 microsatellite (14MS) loci

Date	Genes	Overall F_{ST} (95% CI)	IBD *P*	R	Intercept	Slope [95% CI]
October 2001	14MS	0.034 [0.029; 0.039]	0.1039	0.274	0.004	0.0031 [–0.0024; 0.0114]
	DQA1	0.064	0.0022	0.807	–0.597	0.0768
	DRB	0.045	0.0458	0.378	–0.132	0.0212
October 2002	14MS	0.024 [0.017; 0.030]	0.0011	0.617	–0.031	0.0065 [0.0032; 0.0104]
	DQA1	0.036	0.0086	0.593	–0.346	0.0449
	DRB	0.027	0.0536	0.355	– 0.080	0.0125
October 2003	14MS	0.015 [0.011; 0.020]	0.0221	0.564	–0.054	0.0082 [0.0045; 0.0130]
	DQA1	0.009	0.2270	0.123	–0.016	0.0032
	DRB	0.023	0.0918	0.378	–0.065	0.0104

based on the 14 microsatellites showed the highest degree of neutral genetic differentiation and no significant isolation-by-distance pattern in the first sampling session (Table 3). F_{ST} estimates decreased significantly in subsequent years (one-tailed permutation test; 2003 vs. 2001: $P = 0.001$; 2002 vs. 2001: $P = 0.038$; 2003 vs. 2002: $P = 0.073$), indicating an increase in effective migration between sites with increasing density.

An overall decrease in differentiation between demes over time was also observed for both MHC genes (Table 3). One-tailed permutation tests confirmed that the decrease in F_{ST} estimates between 2001 and 2003 was significant for *DQA1* ($P = 0.018$), but not for *DRB* ($P = 0.079$). In the first sampling session, when populations were geographically isolated, both overall differentiation and isolation-by-distance were more pronounced at MHC genes than at neutral markers (Table 3, Fig. 4). F_{ST} estimates and estimates of the slope of the isolation-by-distance regression line for *DQA1* and *DRB* exceeded the upper limit of the 95% confidence intervals estimated for neutral markers (Table 3). These differences between MHC and neutral markers gradually disappeared over the next few years, with the increase in population density (Fig. 4). In the year in which population density was highest (2003), the overall F_{ST} value and slope estimate were significantly lower for *DQA1* than for neutral markers (Table 3).

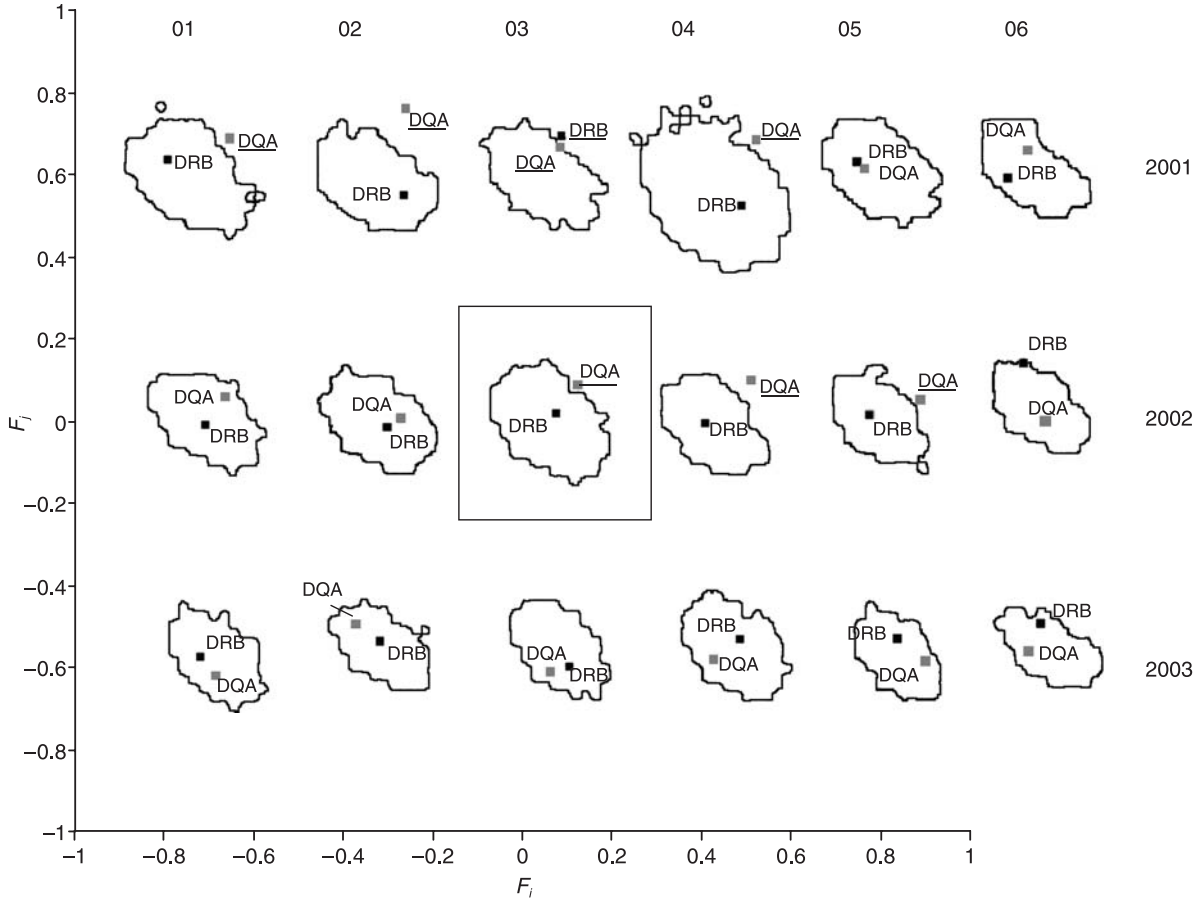

Fig. 5 Expected distribution of pairs of F_i and F_j values estimated from 14 microsatellites and two MHC loci for pairwise comparisons of population 07 with all other populations (01–06) in the years 2001, 2002, and 2003. Axes correspond to the comparison of populations 03 and 07 in 2002 (in the frame) but all the other graphs are on the same scale. Solid lines enclose a region in which 95% of the simulated data points (10 000 permutations in DETSEL) are expected to lie. Only the positions of the *DQA1* and *DRB* loci are shown. Significant outliers are underlined.

Corrected estimates of F_{ST} provided, in essence, very similar results, and did not call into question the result obtained with uncorrected estimates. First, correction for the presence of null alleles at the *DRB* locus gave estimates very similar to those obtained with standard estimators (2001: $F_{ST} = 0.045$ vs. $F_{STcorr} = 0.043$; 2002: $F_{ST} = 0.027$ vs. $F_{STcorr} = 0.026$; 2003: $F_{ST} = 0.023$ vs. $F_{STcorr} = 0.024$). Isolation-by-distance parameters were also almost identical (not shown). Second, although the absolute estimated values of F'_{ST} were different from those of F_{ST} estimators, the differences between MHC genes and microsatellites remained very similar and were of identical statistical significance (see Table S1 and Fig. S1, Supplementary material).

Model-based simulations in FDIST2 identified no MHC gene as an outlier in any of the studied years, suggesting the absence of a global effect of selection on F_{ST} at these loci (not shown). However, in seven pairwise comparisons of the 63 carried out with DETSEL the *DQA1* gene was found to lie significantly outside the 95% confidence interval. All seven pairs (four in 2001, three in 2002, and none in 2003) included population 07 (Fig. 5). *DRB* lay slightly outside the 95% confidence interval in only two pairwise comparisons in 2001 and one in 2002 (Fig. 5—only pairs with population 07 are shown).

Discussion

Evidence for historical selection acting on DQA1 *and* DRB

Major histocompatibility complex genes encode proteins with an important function in the immune response and these genes have been subject to natural selection during the evolution of species (Bernatchez & Landry 2003). Unsurprisingly, both *DQA1* and *DRB* in water voles show an excess of nonsynonymous nucleotide substitutions in the ABS and extensive trans-species evolution has been observed within the family (Bryja *et al*. 2006; Oliver & Piertney 2006; this study). However, these signs of selection do not necessarily mean that selective forces act on

contemporary populations because (i) the accumulation of a significant nonsynonymous to synonymous substitution ratio takes a long time, and (ii) once this signal has been generated, it may take a very long time to disappear in the absence of selection (Garrigan & Hedrick 2003). Nevertheless, our findings provide evidence that selection has played an important role in the history of both the MHC genes studied, and we therefore looked for signs and mechanisms of recent selection.

Contemporary selection on MHC and the interplay between immigration and selection

There is evidence to suggest that MHC genes evolve by mechanisms including spatial and temporal fluctuations in selection (Charbonnel & Pemberton 2005). For example, in salmonid fishes, greater population differentiation has often been observed at MHC loci than at neutral loci, indicating a lower effective migration rate at MHC, with particular subset of alleles locally favoured but showing limited success at invading other populations (reviewed in Bernatchez & Landry 2003). This pattern was found to be geographically limited to only a portion of the populations studied, suggesting spatial variations in the intensity of selection pressure (Landry & Bernatchez 2001; Miller *et al*. 2001; Aguilar & Garza 2006).

In line with these studies, we provide here the first evidence of spatial and temporal variations in the interplay between immigration and selection acting on MHC genes in mammals. At low densities, when fragmented populations occurred only in restricted patches, significantly higher levels of genetic differentiation between demes were detected at MHC genes than at neutral markers. This observation provides evidence of contemporary diversifying selection acting on MHC loci. A detailed examination of the data provided even stronger support for the action of local selection, as it revealed that only one of the seven populations was affected by selection at the *DQA1* gene. With increasing density, gene flow increased for both neutral and MHC genes and the spatial genetic structuring of both types of markers decreased. Nevertheless, once high levels of gene flow between demes were re-established, we again detected significant differences between the patterns of genetic structure observed at neutral and MHC genes. Allele frequencies at the *DQA1* locus were more similar than those at neutral loci, and the spatial structure of the genetic variation became significantly more homogeneous at *DQA1* than at neutral genes. Both these observations suggest that balancing selection acted on *DQA1* gene.

Locus-specific selection and necessity of multigene studies

We show here that selection can affect genetic diversity more strongly for one MHC gene than for another. Signs of local diversifying selection in fragmented water vole populations and balancing selection in continuous populations were more marked for *DQA1* than for *DRB*, suggesting that different contemporary selection pressures act on these two MHC loci. Different antigens interact with different immune receptors, and this pattern is therefore not unexpected. Unfortunately, very little is known about recent selection on more than one MHC locus in wild vertebrates and almost no data are available for other important immune system genes (Acevedo-Whitehouse & Cunningham 2006).

It is generally assumed that most MHC genes are located very close to each other on the chromosome (forming a 'complex'). However, almost no information is available concerning the physical position of MHC genes in wild vertebrates. In rodents, the precise position of class II genes on the chromosome is known only for the laboratory mouse (Blake *et al*. 2003) and rat (Hurt *et al*. 2004), in which the *DQA* gene is tightly linked to the *DRB1* and *DRB2* genes. The situation may be different in water voles, which are known to have two functional *DQA* genes (Bryja *et al*. 2006) and the only *DRB* gene is amplified (this study). The very weak linkage disequilibrium observed between *DQA1* and *DRB* and the differences in selection pressure on the two genes studied may have several different origins. First, the *DQA1* gene may be located far from *DRB* and *DQA2* on the chromosome. Second, some populations display a heterozygote deficit for *DRB*, suggesting the presence of null alleles, which may mask the linkage disequilibrium signal. Third, an increasing number of studies in nonmodel vertebrates have shown that recombination is common within the MHC locus (e.g. Richman *et al*. 2003; Schaschl *et al*. 2005). High-resolution linkage disequilibrium analysis coupled with sperm typing has shown that recombination hot spots can profoundly affect linkage disequilibrium in the MHC class II region in the human and mouse genomes (Jeffreys *et al*. 2001; Kauppi *et al*. 2004). High-resolution sperm analysis has also provided clear evidence for the presence of a hot spot within the E_β gene, which lies between the *DQA1* and *DRB1* homologues in mouse (Yauk *et al*. 2003). Selection favouring new recombinants may therefore conceal linkage between physically close genes. Although we studied only two genes of the immune system in this study, our results clearly indicate that multigene studies focusing on various components of the immune system are required to understand the importance and function of immunity in population dynamics.

Comparing different genetic markers for the detection of contemporary selection: pitfalls and solutions

Microsatellites are often used as 'neutral' controls in various tests of contemporary selection. However, theoretically, there are many reasons to doubt the neutrality of a particular

microsatellite, even if we assume that the evolution of microsatellite loci as a whole is not subject to selection. For instance, microsatellites may 'hitchhike' with a linked gene evolving under selection (e.g. Kohn *et al*. 2000). In the few empirical studies that have tested the assumption of neutrality for microsatellites, selection signatures were detected in a non-negligible proportion of loci: from a few to 27% in various vertebrates (Vasemägi *et al*. 2005; Nielsen *et al*. 2006; Larsson *et al*. 2007). In our study, 17% of 17 loci were suspected to be non-neutral. Although rarely tested, selection at individual microsatellites or closely linked loci may be more common than generally recognized.

The effect of contemporary balancing selection on MHC genes within populations is often assessed with the Ewens–Watterson (EW; Ewens 1972; Watterson 1978) test of neutrality However, significant deviations from mutation-drift equilibrium due to excess heterozygosity (assessed with the EW test) may result not only from balancing selection, but also from demographic bottlenecks (Cornuet & Luikart 1996). It is therefore essential to compare neutral and presumably selected loci to separate demographic from selective causes of mutation-drift disequilibrium. However, EW tests are not suitable for microsatellites because they infer an IAM mutation model, which is not the most likely model for microsatellites (Estoup *et al*. 2001). Alternatively, the tests implemented in BOTTLENECK software (Piry *et al*. 1999) could be used, considering particular mutation models for different types of loci.

When estimating genetic differentiation between populations using different types of markers, it is necessary to control for two potential biases in F_{ST} estimation. First, a major drawback of previous studies comparing the population-genetic structure of MHC genes and microsatellites is the strong dependence of F_{ST} on the level of within-population heterozygosity. Highly variable markers tend to have lower F_{ST} values than less variable markers (Hedrick 2005). It is therefore difficult to differentiate between marker variability and actual variations in gene flow with this approach. Hedrick (2005) overcame this problem by developing a genetic differentiation measure that is standardized by dividing by the maximum possible between-population differentiation given the level of heterozygosity. This measure should now be used in place of F_{ST} in all studies aiming to compare levels of genetic differentiation at different markers.

Second, null alleles (NA) can occur at MHC loci, because primers are often designed within exons, allowing the genotyping of both genomic DNA and cDNA by the same PCR protocol. However, the most frequently analysed exon of MHC class II genes, exon 2, is the most variable part of the gene and a single set of primers within the exon is therefore highly unlikely to amplify all alleles. As NA overestimates F_{ST} values, the correction for NA (e.g. by the method proposed recently by Chapuis & Estoup 2007) should give a minimal estimate of the F_{ST} for these loci. We can then confidently state that the 'true' F_{ST} value should lie between a maximum estimated with the uncorrected F_{ST} and a minimum estimated with the corrected F_{ST}. However, the application of NA correction to genes potentially under selection still remains problematic because we cannot determine in advance whether the heterozygote deficit is linked to NA, selection, or both.

Parasitism as the potential mechanism driving the contemporary selection of MHC genes

Several factors potentially drive the evolution of MHC genes, including infectious agents and mating system (Apanius *et al*. 1997). Only the effects of pathogens are likely to vary both over time and between populations, either directly due to variations in the prevalence and community structure of parasites (e.g. in cyclic voles; Sinervo *et al*. 2000) or indirectly, because the fitness costs of infection may vary with resource availability (e.g. in birds; Holmstad *et al*. 2005). There is evidence that the composition of the vole parasite community in the study area varies (Deter *et al*. 2006, 2007; Cerqueira *et al*. 2007), and this heterogeneity of pathogen-induced selection may account for the different types of selection found to be operating on MHC genes during changes in population density.

In 2001, the higher levels of genetic differentiation observed at MHC genes than at microsatellites were mostly accounted for by a single site (population 07) where different genotypes were detected. Selection in this area, mostly during the low-abundance phase of the vole cycle may support this pattern. Local differences in the structure of parasite communities due to habitat specificity or extinction–recolonization processes may mediate this selection. In contrast, at high density (2003), some kind of balancing selection was detected at *DQA1*, suggesting that parasite-mediated selection was equivalent at the seven sites. The increase in effective migration during the phase of population growth in the vole cycle (Berthier *et al*. 2006) was probably the main factor leading to homogenization of the parasite community and its prevalence. Investigations of the repeatability of these results for the prevalence and structure of the parasite community in these vole populations, over different demographic cycles, are now required to confirm this hypothesis.

Acknowledgements

We would like to thank Y. Chaval and J. Deter for their help in the field and for providing samples. We thank F. Rousset and P. Meirmans for their help in population genetic analysis and J. Piálek, J. Gouy de Bellocq, R. Streiff and three anonymous reviewers for constructive comments on previous versions of the manuscript. This work received financial support from the Institut National de la

Recherche Agronomique, the Region of Franche-Comté, the DIREN Franche-Comté, the French Ministry of the Environment, the EU-FP6 funded EDEN Project (no. EDEN0033, European Commission, DG research, GOCE-CT-2003-010284) and GA ASCR (IAA600930608).

References

Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology & Evolution*, **21**, 433–438.

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.

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.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

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 (2001) GENETIX *4.02, Logiciel sous Windows pour la Génétique des Populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Biological Sciences*, **57**, 289–300.

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.

Berthier K, Charbonnel N, Galan M, Chaval Y, Cosson JF (2006) Migration and recovery of the genetic diversity during the increasing density phase in cyclic vole populations. *Molecular Ecology*, **15**, 2665–2676.

Berthier K, Galan M, Foltete JC, Charbonnel N, Cosson JF (2005) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Molecular Ecology*, **14**, 2861–2871.

Blake JA, Richardson JE, Bult CJ *et al*. (2003) MDG: the mouse genome database. *Nucleic Acids Research*, **31**, 193–195.

Boyce WM, Hedrick PW, MuggliCockett NE *et al*. (1997) Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. *Genetics*, **145**, 421–433.

Brown JH, Jardetzky TS, Gorga JC *et al*. (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **64**, 33–39.

Bryja J, Galan M, Charbonnel N, Cosson JF (2005) Analysis of major histocompatibility complex class II gene in water voles using capillary electrophoresis-single stranded conformation polymorphism. *Molecular Ecology Notes*, **5**, 173–176.

Bryja J, Galan M, Charbonnel N, Cosson JF (2006) Duplication, balancing selection and trans-species evolution explain the high levels of polymorphism of the DQA MHC class II gene in voles (Arvicolinae). *Immunogenetics*, **58**, 191–202.

Cerqueira D, Delattre P, De Sousa B *et al*. (2007) Numerical response of a helminth community in the course of a multi-annual abundance cycle of the water vole (*Arvicola terrestris*). *Parasitology*, **134**, 705–711.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Charbonnel N, Pemberton J (2005) A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity*, **2005**, 1–12.

Charbonnel N, Gouy de Bellocq J, Morand S (2006) Immunogenetics of macroparasites and small mammals interactions. In: *Micromammals and Macroparasites: from Evolutionary Ecology to Management* (eds Morand S, Krasnov BR, Poulin R), pp. 401–442. Springer, Tokyo.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2004.

Deter J, Berthier K, Chaval Y, Cosson JF, Morand S, Charbonnel N (2006) Influence of geographical scale on the detection of density dependence in the host-parasite system, *Arvicola terrestris* and *Taenia taeniaeformis*. *Parasitology*, **132**, 595–605.

Deter J, Chaval Y, Galan M *et al*. (2007) Linking demography and host dispersal to *Trichuris arvicolae* distribution in a cyclic vole species. *International Journal for Parasitology*, **37**, 813–824.

Di Ciccio TJ, Efron B (1996) Bootstrap confidence intervals. *Statistical Science*, **11**, 189–212.

Estoup A, Wilson IJ, Sullivan C, Cornuet JM, Moritz C (2001) Inferring population history from microsatellite and enzyme data in serially introduced cane toads, *Bufo marinus*. *Genetics*, **159**, 1671–1687.

Ewens WJ (1972) The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, **3**, 87–112.

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

Giraudoux P, Delattre P, Habert M *et al*. (1997) Population dynamics of fossorial water vole (*Arvicola terrestris scherman*): a land use and landscape perspective. *Agriculture, Ecosystems and Environment*, **66**, 47–60.

Gutierrez-Espeleta GA, Kalinowski ST, Boyce WM, Hedrick PW (2000) Genetic variation and population structure in desert bighorn sheep: implications for conservation. *Conservation Genetics*, **1**, 3–15.

Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

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

Holmstad PR, Hudson PJ, Skorping A (2005) The influence of a parasite community on the dynamics of a host population: a longitudinal study on willow ptarmigan and their parasites. *Oikos*, **111**, 377–391.

Hurt P, Walter L, Sudbrak K *et al*. (2004) The genomic sequence and comparative analysis of the rat major histocompatibility complex. *Genome Research*, **14**, 631–639.

Jeffreys AJ, Kauppi L, Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nature Genetics*, **29**, 217–222.

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: *Mammalian Protein Metabolism* (ed. Munro HN), pp. 21–132. Academic Press, New York.

Kauppi L, Jeffreys AJ, Keeney S (2004) Where the crossovers are: recombination distributions in mammals. *Nature Reviews Genetics*, **5**, 413–424.

Klein J, Bontrop RE, Dawkins RL *et al*. (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*, **31**, 217–219.

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.

Kumar S, Tamura K, Nei M (2004) MEGA 3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

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.

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.

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

Muirhead CA (2001) Consequences of population structure on genes under balancing selection. *Evolution*, **55**, 1532–1541.

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

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New-York.

Nielsen EE, Hansen MM, Meldrup D (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology*, **15**, 3219–3229.

O'Brien SJ, Troyer JL, Roelke M, Marker L, Pecon-Slattery J (2006) Plagues and adaptation: lessons from the Felidae models for SARS and AIDS. *Biological Conservation*, **131**, 255–267.

Oliver MK, Piertney SB (2006) Isolation and characterization of a MHC class II *DRB* locus in the European water vole (*Arvicola terrestris*). *Immunogenetics*, **58**, 390–395.

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

Piry S, Lukart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetic software for exact tests and eucumenism. *Journal of Heredity*, **86**, 248–249.

Redpath SM, Mougeot F, Leckie FM, Elston DA, Hudson PJ (2006) Testing the role of parasites in driving the cyclic population dynamics of a gamebird. *Ecology Letters*, **9**, 410–418.

Richman AD, Herrera LG, Nash D, Schierup MH (2003) Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in *Peromyscus maniculatus*. *Genetical Research*, **82**, 89–99.

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

Saccheri I, Hanski I (2006) Natural selection and population dynamics. *Trends in Ecology & Evolution*, **21**, 341–347.

Saucy F (1994) Density dependence in time series of the fossorial form of the water vole, *Arvicola terrestris*. *Oikos*, **71**, 381–392.

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.

Schaschl H, Suchentruk F, Hammer S, Goodman SJ (2005) Recombination and the origin of sequence diversity in the *DRB* MHC class II locus in chamois (*Rupicapra* spp.). *Immunogenetics*, **57**, 108–115.

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

Seddon JM, Ellegren H (2004) A temporal analysis shows major histocompatibility complex loci in the Scandinavian wolf population are consistent with neutral evolution. *Proceedings of the Royal Society B: Biological Sciences*, **266**, 1283–1290.

Sinervo B, Svensson E, Comendant T (2000) Density cycles and an offspring quantity and quality game driven by natural selection. *Nature*, **406**, 985–988.

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.

Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811–1823.

Vitalis R, Dawson K, Boursot P, Belkhir K (2002) DETSEL 1.0: a computer program to detect markers responding to selection. *Journal of Heredity*, **94**, 429–431.

Watterson GA (1978) The homozygosity test of neutrality. *Genetics*, **88**, 405–417.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Worley K, Carey J, Veitch A, Coltman DW (2006) Detecting the signature of selection on immune genes in highly structured populations of wild sheep (*Ovis dalli*). *Molecular Ecology*, **15**, 623–637.

Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in Biosciences*, **13**, 555–556.

Yang ZH, Swanson WJ, Vacquier VD (2000) Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Molecular Biology and Evolution*, **17**, 1446–1455.

Yang ZH, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, **22**, 1107–1118.

Yauk CL, Bois PRJ, Jeffreys AJ (2003) High-resolution sperm typing of meiotic recombination in the mouse MHC E_β gene. *EMBO Journal*, **22**, 1389–1397.

This paper originated in the CBGP laboratory during J.B.'s postdoctoral work on MHC genes in fluctuating populations of voles. J.B. now heads a genetic laboratory at IVB, where he uses molecular markers in ecological and evolutionary studies. N.C. is a population biologist whose main research interests are population genetics and host–parasite interactions. K.B., during her PhD, used demographic and genetic approaches to estimate the role of dispersal in fluctuating populations; she now studies the MHC of endangered birds. M.G. is a scientific officer specializing in molecular biology and working on various mammals. J.-F.C. is mainly interested in the population genetics of fluctuating populations in relation to landscape ecology.

Supplementary material

The following supplementary material is available for this article:

Fig. S1 Correlation between genetic $F'_{ST}/(1-F'_{ST})$ ($F'_{ST} = F_{ST}$ corrected for the Hedrick effect) and geographical distances (log scale) for three samplings of MHC and neutral markers.

Table S1 Overall F'_{ST} estimates corrected for the Hedrick effect (Hedrick 2005) and isolation by distance characteristics calulated for each marker type in three years.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2007.03584.x
(This link will take you to the article abstract).

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