MINI REVIEW

MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years?

L. BERNATCHEZ* & C. LANDRY†

*Département de biologie, Université Laval, Ste Foy, Québec, Canada
†Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

Keywords:

evolution;
fitness;
local adaptation;
mate choice;
MHC;
pathogen resistance;
population differentiation;
selection.

Abstract

Elucidating how natural selection promotes local adaptation in interaction with migration, genetic drift and mutation is a central aim of evolutionary biology. While several conceptual and practical limitations are still restraining our ability to study these processes at the DNA level, genes of the major histocompatibility complex (MHC) offer several assets that make them unique candidates for this purpose. Yet, it is unclear what general conclusions can be drawn after 15 years of empirical research that documented MHC diversity in the wild. The general objective of this review is to complement earlier literature syntheses on this topic by focusing on MHC studies other than humans and mice. This review first revealed a strong taxonomic bias, whereby many more studies of MHC diversity in natural populations have dealt with mammals than all other vertebrate classes combined. Secondly, it confirmed that positive selection has a determinant role in shaping patterns of nucleotide diversity in MHC genes in all vertebrates studied. Yet, future tests of positive selection would greatly benefit from making better use of the increasing number of models potentially offering more statistical rigour and higher resolution in detecting the effect and form of selection. Thirdly, studies that compared patterns of MHC diversity within and among natural populations with neutral expectations have reported higher population differentiation at MHC than expected either under neutrality or simple models of balancing selection. Fourthly, several studies showed that MHC-dependent mate preference and kin recognition may provide selective factors maintaining polymorphism in wild outbred populations. However, they also showed that such reproductive mechanisms are complex and context-based. Fifthly, several studies provided evidence that MHC may significantly influence fitness, either by affecting reproductive success or progeny survival to pathogens infections. Overall, the evidence is compelling that the MHC currently represents the best system available in vertebrates to investigate how natural selection can promote local adaptation at the gene level despite the counteracting actions of migration and genetic drift. We conclude this review by proposing several directions where future research is needed.

Introduction

Elucidating how natural selection promotes local adaptation in interaction with migration, genetic drift and mutation is a central aim of evolutionary biology. Despite the fact that numerous candidate genes potentially under the influence of positive selection have been identified in model organisms (reviewed in Ford, 2002), several conceptual and practical limitations are still restraining our ability to routinely study these processes at the DNA level in most organisms. Not only it is still difficult to find appropriate genes to study in nonmodel species, but also knowledge on the function of such

Correspondence: Louis Bernatchez, Département de Biologie, Université Laval, Ste-Foy, Québec, Canada G1K 7P4.
Tel.: 418 656 3402; fax: 418 656 2043;
e-mail: louis.bernatchez@ bio.ulaval.ca

genes is often incomplete. This may limit our ability to formulate testable hypotheses regarding the possible role of selection in shaping patterns of variation at these genes and also to assess the possibility that such variation is adaptive.

As acknowledged early by Potts & Wakeland (1990), and subsequently by Hedrick (1994), genes of the major histocompatibility complex (MHC) offer several assets that make them unique candidates for studies of adaptation in natural populations. The details of the molecular processes in which they are involved have been the subject of intense research for more than two decades (Hughes & Yeager, 1998). Consequently, extensive information is available about the structure of MHC genes and their function (Klein, 1986; Stet & Egberts, 1991; Dixon *et al.*, 1996; Malaga-Trillo *et al.*, 1998; Ploegh & Watts, 1998; Van Muiswinkel *et al.*, 1999; Jeffery & Bangham, 2000; Hess & Edwards, 2002; Penn, 2002). The MHC also comprises the most polymorphic genes in the vertebrate genomes, and many studies supported the general hypothesis that allelic diversity at MHC genes is maintained by parasite-mediated balancing selection (Apanius *et al.*, 1997; Edwards & Hedrick, 1998; Hughes & Yeager, 1998; Meyer & Thomson, 2001; Dean *et al.*, 2002; Hess & Edwards, 2002; Penn *et al.*, 2002).

Yet, after 15 years of empirical research that documented MHC diversity in natural populations since the pioneer study of Watkins *et al.* (1988), it is unclear what general conclusions can be drawn regarding the role of selection in maintaining adaptive MHC diversity in species other than humans and mice. Thus far, the numerous review papers on MHC have largely focused on synthesizing current knowledge on MHC molecular architecture, functional aspects such as mate choice, and evidence supporting the hypothesis that MHC polymorphism is maintained by balancing selection in humans and mice studied in experimental or semiexperimental conditions. In this context, the general objective of this review is to complement earlier literature syntheses by focusing on MHC studies in species other than human and mice.

The structure and specific aims of this review are as follows. First, we briefly review general information about the function of the MHC genes, as well as current hypotheses and concepts concerning the role of selection on the maintenance of MHC polymorphism. Secondly, we outline the principles of selected methods that have been used to test for the role and type of selection in shaping MHC variation, both within and among populations. We then summarize the several case studies that applied such methods. Thirdly, we review the results of studies that specifically focused on assessing the relationships between MHC polymorphism and fitness. Finally, we propose several directions where future research is needed.

General background on the MHC

Function of the MHC genes

In all vertebrates studied to date, the major histocompatibility complex is a multigene family acting at the interface between the immune system and infectious diseases. These loci encode receptors on the surface of a variety of immune and nonimmune cells. The primary role of these receptors is to bind fragments of proteins in the cells and transport them to the membrane surface where the complex is recognized by T cells that can initiate the cascade of complex immune responses (Ploegh & Watts, 1998). The MHC family includes two major subfamilies; class I and class II genes. In mammals and birds, the class I and II genes are linked together in a single gene complex (Hughes & Yeager, 1998; Hess & Edwards, 2002). In contrast, recent studies have shown that class I and II genes are not localized on the same linkage groups in teleost fishes and that a large family of non-MHC linked class I genes is expressed in the amphibian *Xenopus* (Flajnik *et al.*, 1993; Bingulac-Popovic *et al.*, 1997; Hansen *et al.*, 1999). MHC class I genes are expressed on the surface of all nucleated somatic cells. They play an essential role in the immune defence against intracellular pathogens by binding endogenously derived peptides from proteins (mainly viral) in the cytoplasm. In contrast, MHC class II genes have a much more restricted expression pattern as they are primarily expressed on antigen-presenting cells of the immune system, such as B cells and macrophages. They are also primarily involved in monitoring the extracellular (rather than intracellular) environment by presenting exogenously produced (mainly bacterial) peptide antigens to helper T cells. While the general architecture of multigene MHC families appears relatively conserved within each class of vertebrates, the number of either class I or II loci has been found to vary substantially among species. For instance, up to 17 class I loci may be expressed in cichlid fishes, whereas no more than three class I loci have been documented in salmonid species (Malaga-Trillo *et al.*, 1998; Miller & Withler, 1998). The ability of both class I and II genes to face various pathogens is believed to be mainly related to sequence variation among MHC alleles in the antigen binding site or ABS (also named antigen recognition sites or ARC, and peptide binding region or PBR) (Van Eden *et al.*, 1983; Potts & Wakeland, 1990).

Hypotheses concerning the role of selection on the maintenance of polymorphism

Two main types of mechanisms may operate to maintain the unusually high level of MHC polymorphism: the disease-based and reproductive mechanisms. Disease-based models infer that genetic diversity at

MHC is maintained by balancing selection stemming from the coevolution of host with their pathogens and parasites, which can be of two basic types. The first disease-based model that took into account the function of MHC was the *overdominance hypothesis* (or the heterozygote advantage). Thus, Doherty & Zinkernagel (1975) argued that in a population exposed to an array of pathogens, heterozygous individuals are favoured as they are able to present a broader array of antigens, and therefore resist a broader array of pathogens than homozygotes (see also Hughes & Nei, 1988).

The second mechanism for MHC polymorphism is the *negative frequency dependent selection hypothesis* (or rare-allele advantage) first proposed by Clarke & Kirby (1966) (see also Bodmer, 1972 and Takahata & Nei, 1990). This model proposes that parasite antigenicity will be selected to exploit MHC-based immune response defect of the most common host genotype. This would decrease the relative fitness of the common host genotypes and provide a selective advantage to new, rare MHC alleles. The time-lag nature of these antagonistic coevolutionary responses could lead to the cycling of fitness values of different genotypes in both hosts and pathogens, and result in the maintenance of high genetic diversity.

The second type of mechanism relates to the genetic benefits of sexual selection. There are currently two leading explanations to account for the development of reproductive mechanisms favouring MHC polymorphism (Penn, 2002). First, disease-based fitness differences between MHC genotypes would favour reproductive mechanisms that preferentially produce high fitness genotypes in their progeny. MHC-dependent mating preferences may enhance parasite resistance in two ways. First, MHC disassortative mating preferences could increase fitness of choosy parents because a disproportionate number of offspring would be high fitness MHC heterozygotes. Secondly, mating preferences could potentially provide a mechanism to 'keep up' the molecular arms race of coevolution with parasites, a scenario coined the *moving target hypothesis* by Penn & Potts (1999).

The second model that relates the maintenance of MHC polymorphism to reproductive mechanisms is the *inbreeding avoidance hypothesis* (Potts & Wakeland, 1990). This suggests that MHC-based selective mating is similar to other genetic incompatibility systems, which function to reduce inbreeding (e.g. plants incompatibility system). Because MHC loci are often highly polymorphic, individuals that share MHC alleles are likely to be related. MHC alleles are also known to influence specific odours (reviewed in Penn, 2002), and may therefore provide a potential marker for relatedness. Thus, MHC-disassortative mating preferences may prevent kin matings and negative consequences of inbreeding, such as the expression of recessive deleterious mutations. This hypothesis therefore predicts that MHC-dependent mating preferences should be primarily favoured among species at risk of inbreeding (Jordan & Bruford, 1998).

Selected methods to test for selection and outcomes in studies of natural populations

In this section, we first briefly outline the principles of methods that have been most commonly used to test the fit of observed patterns of MHC variation to expectations under neutrality. We also present more recently developed approaches that offer particular potential in this respect. The outcomes of relevant case studies are summarized following each test presented. Methods for inferring the effect of selection on MHC variation can be divided into two basic categories: tests that allow detection of the overall effect of positive selection over evolutionary times, and tests that allow testing for the possible role of selection in shaping patterns of diversity within and among contemporary populations.

As illustrated by our review of pertinent studies (see Supplementary material on the web), there is a strong taxonomic bias in studies that applied such tests, whereby many more studies of MHC diversity in wild populations have dealt with mammals than all other vertebrates combined. This bias would be even more pronounced, had studies on mice and humans been considered as well. Thus, the apparent popularity of the MHC gene complex still has to translate into a better taxonomic coverage in order to develop a more general view of existing patterns of MHC diversity in natural populations.

Test of positive selection over evolutionary time scale

dn / ds ratio test

The most commonly applied method for detecting the long-term effect of positive selection is the *dn / ds* ratio test, first proposed by Hill & Hastie (1987), and then Hughes & Nei (1988). The rationale of this test is that if synonymous mutations are essentially neutral because they do not result in amino acid replacement, then the rate of synonymous site evolution (*ds*) will equal the mutation rate. Nonsynonymous mutations that result in change of amino acid sequences are more likely to be under the effect of selection. Thus, in situations where such mutations are deleterious the ratio of *dn / ds* is predicted to be significantly <1, and to reflect the effect of purifying selection. In contrary, if nonsynonymous mutations are maintained in populations by positive selection, then the *dn / ds* ratio is predicted to be significantly >1. A major asset of the *dn / ds* ratio test is that it does not make assumptions about population structure and equilibrium conditions (Nielsen, 2001). In the specific case of the MHC genes, differential rates of

synonymous and nonsynonymous substitution per site should be more pronounced in the codons encoding the peptide binding region (PBR) of the molecule.

The *dn / ds* ratio test was applied in 48 of the 78 studies we reviewed. All but one revealed a *dn / ds* > 1 (supplementary material on the web). The exception is a study on Xenopus Class I genes, in which the number of nonsynonymous substitution did not exceed synonymous substitution (Sammut *et al*., 2002). However, this result is to be taken with caution as the number of sequences that were compared was low, and their phylogenetic relationship and origin was unclear.

Trans-species polymorphism

The second most commonly used method to infer the long-term effect of selection on MHC is trans-species polymorphism, that is the non-neutral retention of alleles across species causing discordance between alleles and species trees (Figueroa *et al*., 1988). Coalescence theory predicts that neutral polymorphism is generally not expected to be maintained for a long time across speciation events, persisting on average 4Ne generations. However, Takahata & Nei (1990) showed that polymorphism can persist much longer under several models of balancing selection that slows down the coalescent process, leading to long genealogical relationships among alleles, even with relatively modest selection coefficients. Long living allelic lineages at MHC genes that could not be explained solely by the retention of neutral polymorphism have been reported in 40 of 42 studies that tested for trans-species polymorphism. In a typical example of trans-species polymorphism study, Garrigan & Hedrick (2001) uncovered 12 MHC class 1 exon 2 sequences in a population of Chinook salmon. Phylogenetic analysis of the 12 sequences showed that the alleles descended from two of six major allelic lineages also found among four other species of Pacific salmon. Nine of the 12 alleles belonged to an allelic lineage that began diversifying 8 million years ago, that is prior to the estimated time of Chinook salmon speciation. The most recent common ancestor of all 12 alleles was estimated to be 15 million years ago, approximately 5 million years before the radiation of the Pacific salmon species.

Despite the common use of trans-species polymorphism, one should nevertheless keep in mind that this test may be biased if mutational mechanisms such as interlocus recombination and gene conversion are operating (Martinsohn *et al*., 1999). Indeed, there is now good experimental and phylogenetic evidence for the latter which questions traditional claims of trans-species polymorphism (e.g. Hogstrand & Bohme, 1994; Bergstrom *et al*., 1998). Thus, the origin of MHC alleles could in some cases be much more recent than previously assumed, despite their apparently high level of divergence.

Coalescence-based test of alternative overdominance models

A less commonly applied, yet potentially useful method to assess the role of overdominant selection is to compare the structure and shape of allelic genealogy with a coalescent model of the evolution of balanced genetic polymorphism. The main advantage of this approach is that it potentially allows testing of two alternative hypotheses with respect to the overdominance model: (i) symmetric balancing selection (SBS), whereby all heterozygotes derive a similar selective advantage relative to homozygotes, and (ii) divergent allelic advantage (DAA), whereby heterozygotes carrying more divergent allelic sequences have a selective advantage relative to individuals carrying relatively similar alleles. Under symmetric overdominance (all heterozygous are selectively equivalent), Takahata & Nei (1990) showed that the allelic genealogy has a structure similar to a neutral gene genealogy. In contrast, a higher fitness for heterozygotes with more different alleles compared with heterozygous with more similar alleles is predicted to maintain long-terminal branches on the species allelic genealogy relative to the total coalescent time. These alternative scenarios can be assessed by quantifying the ratio of coalescent time within lineages to the length of the genealogy, and its significance of deviations from expectations can be tested using expected distribution obtained by computer simulation (Richman & Kohn, 1999; Richman, 2000).

The only application of this test was recently performed by Richman *et al*. (2001) in a study of deer mouse (*Peromyscus maniculatus*) MHC Class IIB gene. By sequencing the RT-PCR products from spleen cDNAs, the authors observed that the genealogy of allelic sequences in this species deviates significantly from theoretical expectations under a model of symmetric balancing selection, that is, alleles were more divergent than coalescent expectation. Richman *et al*. (2001) therefore suggested that this observation of high levels of pairwise allelic sequence divergence and deviation of the genealogy from theoretical expectation together provide more support for the divergent allele advantage model than for the symmetric balancing selection model for the maintenance of MHC polymorphism. This method relies on assumptions of the coalescent models in which no gene conversion is allowed and therefore conclusions should be taken with care if such a mutational process is suspected.

Tests of selection within and among contemporary populations

In this section, we consider methods that can be used to inform about the nature of selection acting on patterns of MHC allelic diversity within and among contemporary populations. As for the methods described above, these

are based on the general principle of testing the fit between observed patterns of MHC allelic variation with neutral expectations. This can be achieved by contrasting patterns of MHC variation either with theoretical predictions, or with patterns of variation empirically observed at nuclear loci assumed to behave neutrally. The use of such methods has been criticized on the basis that they can suffer from low power, and that their assumptions may frequently be violated (e.g. Ford, 2002). Our view is that with sufficient sample size and number of markers, these can potentially be very useful to assess whether selection operates within and among contemporary populations characterized by distinct pathogens, demographic and environmental conditions, or instead, if selection is episodic with long intervening periods of neutrality (Klein, 1987).

Ewens–Watterson homozygosity test of neutrality

A commonly applied method for testing whether or not patterns of allelic diversity within-population reflect the outcomes of balancing selection is the Ewens–Watterson homozygosity test of neutrality (Ewens, 1972; Watterson, 1978). This method is based on the theoretical principle that if selection provides higher fitness to rare alleles or to heterozygous genotypes, then a more even allelic frequency distribution than under neutral expectations should be observed. Thus, the effect of balancing selection can be detected by comparing the observed allele frequency distribution with the homozygosity expected under neutrality. The neutral expectation in that case comes from a sampling distribution of neutral alleles in a finite population (Ewens, 1972). An expected homozygosity value under neutrality is then derived from a given sample size and observed number of alleles using Ewens' sampling formula. Because balancing selection favours an even distribution of allele frequencies, one expects MHC genes to have a lower homozygosity than neutral genes.

This test supported the hypothesis that MHC genes are under the effect of balancing selection in the four studies that applied it (Paterson, 1998; Landry & Bernatchez, 2001; Miller *et al*., 2001; Hambuch & Lacey, 2002). In studies implying more than one population, the effect of balancing selection appeared to vary among populations. For instance, Landry & Bernatchez (2001) recently reported a statistically significant trend ($P < 0.05$, Fisher's method for multiple independent tests) for reduced homozygosity in a study where 14 populations of wild Atlantic salmon (*Salmo salar*) were characterized at a MHC class IIB locus across a variety of habitats and geographical scales. This suggested that balancing selection is acting to maintain a higher genetic variability than expected under neutrality in this species. This pattern, however, was mainly imputable to six of the 14 populations, therefore suggesting that the strength of selection varied with the environment.

A recent extensive study of genetic variation at the MHC class II *DAB-β*1 locus among populations of the sockeye salmon (*Oncorhynchus nerka*) generated analogous results (Miller *et al*., 2001). In this study, the Ewens–Watterson and the analogous Slatkin's exact test (Slatkin, 1994) detected significant to marginally significant ($P < 0.10$) reduction in homozygosity in 13 of the 31 populations surveyed. Interestingly, deviations from neutrality were suggestive of directional selection in two populations, therefore indicating that both the intensity and the nature of selection may vary among environments.

Estimating the intensity of overdominant selection from the distribution of allele frequencies

Slatkin & Muirhead (2000) recently proposed a method that allows estimating the intensity of overdominant selection from the distribution of allele frequencies within populations. The approach is based on the assumption that, with strong balancing selection, allele frequencies are nearly at their deterministic equilibrium values and that deviation from expectation depends only on the intensity of selection scaled by the effective population size ($S = 2Ns$). The estimate of S is essentially the inverse of the variance in allelic frequencies within populations. Slatkin & Muirhead (2000) also developed a program that finds the maximum-likelihood estimate (MLE) of S and assesses its statistical significance. S most likely underestimates the true intensity of balancing selection because several factors, namely stochastic variation, may cause deviations from the stationary distribution of allele frequencies, increase the variance and hence reduce the estimate. While the estimate of S quantifies the net selective force tending to equalize allele frequencies, it does not, however, validate a particular model of balancing selection (e.g. overdominance or frequency-dependence).

Hambuch & Lacey (2002) recently applied the method of Slatkin & Muirhead (2000) to explore the effects of behavioural and demographic attributes on the intensity of balancing selection at exon 2 of the MHC class II DQ*β* locus. They compared MHC variation in a social and a solitary rodent species, the tuco-tuco (*Ctenomys* sp.), both of which occur in the same valley in south-western Argentina. Hambuch & Lacey (2002) rationale was that specific differences in demography and behaviour are expected to influence pathogen exposure and movement of alleles within and among populations. Thus, variation in these parameters could influence the intensity of balancing selection and therefore, allelic variation. The results strongly suggested that balancing selection acting at the DQ*β*B locus was enhanced in the social species compared to its solitary congener. Estimates of S were 26.8 for *C. sociabilis* compared to 16.0 for *C. haigi,* before expected differences in neutral variation because of the lower effective population size of *C. sociabilis* was taken

into account. When differences in *Ne* approximated from the ratio of $4Ne\mu$ were included, the resulting values of selection coefficient were on average 57 times greater in the social species than the solitary one. While the elucidation of the proximate causes for differences observed between both species must await further research, the study of Hambuch & Lacey (2002) shows that social interactions may influence the intensity of balancing selection acting on MHC polymorphism in natural populations.

Contrasting the patterns of MHC population structure with 'neutral' markers

There has recently been an increasing interest in using MHC to assess whether or not genetic differences *among* populations can be related to differences in local selective constraints (e.g. Boyce *et al.*, 1997; Miller & Withler, 1997; Kim *et al.*, 1999). Thus, a view is emerging that differentiation at loci potentially under selection provides much better information on the structure of adaptive variation among populations, and is therefore of more interest than neutral differentiation for defining conservation units (Miller *et al.*, 2001; Hedrick *et al.*, 2001a). Until very recently, there was no theory available to make *a priori* predictions on the expected pattern of genetic variability for genes under balancing selection in a subdivided population, and against which empirical observations could be compared. Consequently, most studies have inferred the possible role of selection in shaping patterns of MHC genetic structure among populations by using an empirical comparative approach, whereby the extent of population subdivision at MHC is contrasted by conventional statistics (e.g. F_{st}) with that observed at other markers assumed to be under the sole effect of migration and drift. Under balancing selection, differentiation between populations has been predicted to be significantly reduced compared with neutral loci for two main reasons (Schierup *et al.*, 2000). First, because alleles are kept in more equal frequencies than for neutral loci under the assumption of selective equivalence, balancing selection should increase within-population MHC diversity relative to the total diversity (Hedrick, 1999). Secondly, an incoming migrant allele that is absent in a population may be selected for, which enhances its chance of invasion compared with neutral alleles and consequently increases its effective migration rate. Thus, for a given migration rate, F_{st} estimates for loci under balancing selection should be a fraction only of the values observed at neutral loci (Schierup *et al.*, 2000). Alternatively, other evolutionary processes, such as nonselective equivalence among alleles, can promote the maintenance of distinct allelic lineages in populations (e.g. Jeffrey & Bangham, 2000). Pathogen-driven directional selection could also act differentially among individuals from distinct populations if the selective advantage of MHC alleles differs among environments that vary in the diversity and abundance of pathogens. In such a case, it is predicted that the extent of population subdivision at MHC genes should exceed that observed at neutral nuclear markers.

We found 12 studies (six in fish and six in mammals) that have contrasted the patterns of MHC diversity with that of 'neutral' markers (Supplementary material on the web). Two studies on salmonid fishes recently reported a differential pattern of population structuring between MHC and neutral loci. In a recent study on wild Atlantic salmon (*Salmo salar*), Landry & Bernatchez (2001) compared the extent of genetic differentiation between a MHC class IIB gene and microsatellite loci among populations for two types of habitat segregation: (i) among individuals from different sites within a small river drainage, and (ii) among freshwater resident and anadromous (sea-running) populations. A pronounced discrepancy in the extent of genetic differentiation at MHC and microsatellite loci was observed among individuals from different sites within river. Thus, there was no correlation between pairwise F_{st} estimates at both markers ($r = 0.17$, $P = 0.46$), and 14 of 21 pairwise F_{st} estimates were significantly higher for MHC than microsatellites. Moreover, the overall extent of population differentiation at MHC ($F_{st} = 0.047$) was significantly higher than that observed for microsatellite loci (0.028, CI 95%: 0.019–0.035). These results supported the hypothesis that directional selection partly counteracts the effect of migration at this geographical scale. This could be driven by local adaptation to specific river sites that differ in their ecological settings (Garant *et al.*, 2000), and perhaps in their pathogen community (Bakke & Harris, 1998). Comparisons among freshwater resident and sea-running populations provided more limited support for the differential effect of selection on MHC due to habitat difference. Despite a significant trend for pairwise MHC estimates of F_{st} to be higher than for microsatellites, the overall F_{st} value observed at both markers did not differ significantly (MHC: 0.157, microsatellites: 0.124, CI 95%: 0.079–0.179). There was also a strong positive correlation ($r = 0.78$, $P < 0.0001$) between pairwise F_{st} estimates between the two types of markers. Altogether, Landry & Bernatchez (2001) concluded that local adaptation may promote the maintenance of different subsets of MHC alleles among different populations in Atlantic salmon. However, this appears to prevail at small geographical scales, while genetic drift and migration may be sufficient to limit the effect of directional selection across larger habitats.

In their recent study on the sockeye salmon from the Fraser River drainage in British Colombia, Miller *et al.* (2001) also reported a higher level of population differentiation at MHC than microsatellite loci, although no specific statistical criterion other than differences in point values was used to infer selection in this case. Thus, 25% of the total variation at MHC was partitioned among populations, in sharp contrast with the value of 5%

observed at microsatellites. Furthermore, pairwise F_{st} values for MHC averaged 0.19 and tended to exceed the corresponding values estimated for microsatellites neutral loci at all levels of population structure. Miller *et al.* (2001) thus concluded that a high degree of population differentiation at MHC may result from either spatial and/or temporal variability in pathogen driven directional selection, or balancing selection on genotypes with unequal fitness values.

Other studies on salmonids also suggested a general trend for a higher level of population differentiation at MHC than microsatellite loci (Supplementary material on the web). However, these results relied only on a qualitative comparison of point values, and should therefore be interpreted cautiously. In contrast, none of the studies performed on mammals showed significant (or even qualitative) differences in the extent of population differentiation between MHC and other markers (Supplementary material on the web). These results might suggest that selection has been less important in shaping patterns of allelic diversity among these mammalian populations than in salmonids. In fact, they do not rule out the possibility that selection may be very weak at the MHC loci and the alleles effectively neutral.

Approximating expectations of population subdivision under balancing selection

Schierup (1998) recently argued that current models describing expected pattern of genetic variability for genes under balancing selection in a subdivided population may not be reliable. In such a case, it may be hazardous to infer any differentiation among demes attributable to selection in departure of the amount of differentiation to be expected under migration and drift alone. To solve this problem, Muirhead (2001) proposed an analytical approach to model population structure for genes under strong balancing selection of the type seen in MHC. Muirhead (2001) thus derived (and checked against simulations) analytic solutions for the number of alleles maintained at equilibrium and the expected proportion of alleles shared between populations (and related population subdivision) for various levels of selection and migration rates. This rigorously confirmed that lower F_{st} estimates at MHC compared with 'neutral markers' are expected if balancing selection is uniformly acting at MHC. Thus, as argued by Muirhead (2001), similar or higher F_{st} estimates at MHC could imply either that balancing selection is very weak at the MHC loci and the alleles effectively neutral, or alternatively that the effective migration rate at the MHC loci is lower than for neutral loci because of differential selection between populations. In order to illustrate how her method could allow telling these alternative hypotheses apart, Muirhead (2001) re-examined the data of Miller & Withler (1997) on the Chinook salmon for which a slightly higher F_{st} value was observed at MHC compared with microsatellites. This revealed a pattern of allele-sharing vector in the Chinook MHC class I A1 locus that was much more consistent to that seen for low migration rate with selection than with the distribution predicted for neutral genes. Thus, Muirhead (2001) proposed that a higher F_{st} value at MHC could be attained if populations had different selective environments. This would locally favour particular subsets of alleles, but limit their success at invading other populations, thus resulting in higher population differentiation than expected under neutrality or a model of balancing selection that assumes selective equivalence of different alleles across different environments.

MHC polymorphism and fitness

MHC-dependent recognition: mate and kin preference

Only a handful of studies have demonstrated that MHC can directly or indirectly influence mate choice in natural populations (Table 1). Unfortunately, negative results are probably very unlikely to be reported, so establishing how frequently MHC is involved in mate choice in vertebrates is difficult. The only exception is a study by Paterson & Pemberton (1997) on the Soay sheep that found no evidence for MHC-dependent mating preferences. The earliest attempt to correlate MHC variation and phenotypic traits whose expression may be influenced by disease was performed on the ring-necked pheasant (*Phasianus colchicus*). Thus, Von Schantz *et al.* (1989) showed that females prefer to mate with long-spurred males. Moreover, data on reproductive success indicated that they may improve their chicks' survival rate by doing so, as male spur length is positively correlated with age, body size and viability. Subsequently, Von Schantz *et al.* (1996, 1997) genotyped male pheasants for both class I and class IIB genes. Multivariate analyses revealed that MHC genotype was significantly associated with variation in both male spur length and male viability, although no direct link was inferred between disease resistance and particular spur length or MHC haplotypes. Yet, this study showed for the first time in birds that MHC genes could be associated with viability and the expression of a condition-dependent selected male trait.

In a recent analogous study on adult male white-tailed deer (*Odocoileus virginianus*), Ditchkoff *et al.* (2001) tested whether antler development could provide an honest signal to females of a male's genetic quality and condition to adversaries. They thus compared antler, morphometric, hormonal, and parasitic data to characteristics of the MHC-DRB (*Odvi*) locus. They detected associations between certain allelic combinations at this gene and antler development and body mass. This suggested that antler development and body mass may be associated with pathogen resistance. Indeed, a negative relationship was found between degree of antler development and

Table 1 A summary of the main results for studies that have investigated the role of MHC polymorphism in mate preference and kin recognition in vertebrates.

Species	Observation	Proposed selection mechanism	Reference
Fish			
Atlantic salmon *Salmo salar*	Mates tend to choose patterns with dissimilar MHC class II B alleles	Enhanced heterozygosity in offspring	Landry *et al.* (2001)
Three-spined stickleback *Gasterosteus aculeatus*	Females prefer males with larger number of MHC alleles	Enhanced heterozygosity in offspring	Reusch *et al*. (2001)
Arctic charr *Salvelinus alpinus*	Fish prefer water scented with MHC identical siblings to MHC different siblings.	Discriminating kin could operate in inbreeding avoidance	Olsén *et al.* (1998, 2002)
Birds			
Ring-necked pheasant *Phasianus colchicus*	MHC associated with male ornamentation and viability	Good genes associated with MHC	Von Schantz *et al*. (1996, 1997)
Mammals			
White-tailed deer *Odocoileus virginianu*	Association between MHC and Antler development and body mass	Good genes associated with MHC	Ditchkoff *et al*. (2001)
Soay sheep *Ovis aries*	No evidence of MHC-dependent mate preference		Paterson & Pemberton (1997)

overall abundance of abomasal helminths. This study therefore provided support for the hypothesis that antler development is an honest signal of heritable male quality in terms of its MHC genotype.

Two recent studies in fish tested for direct MHC-dependent mating preferences, that is without considering possible correlation between MHC genotypes and sex-related signal traits. Although not strictly documented, these two studies therefore strongly suggested that genotypes were differentially selected through MHC-mediated odours. In the first study showing that MHC genes influence mate choice in fish, Landry *et al*. (2001) tested the null hypotheses that mate choice in wild Atlantic salmon is not dependent on the relatedness between potential partners or on the MHC similarity between mates using both the Class IIB gene and microsatellite loci. No mating preference was observed when the numbers of alleles shared between individuals was compared. However, the authors showed that mates significantly preferred to mate with fish having dissimilar functional MHC Class IIB alleles, thereby increasing the average heterozygosity of their offspring in terms of amino-acid composition at the MHC. In contrast, inbreeding avoidance was ruled out as an explanation for the observed pattern of MHC-dependent mate preference as genetic relatedness among mates as estimated with microsatellites did not differ from random expectations. In a subsequent study on wild-caught three-spined stickleback (*Gasterosteus aculeatus*), Reusch *et al*. (2001) also tested the alternative hypotheses of inbreeding avoidance and parasite resistance as a mechanism to promote sexual selection. In contrast to Atlantic salmon, sticklebacks possess up to six recently duplicated MHC class-IIB genes that could not be analysed individually. Instead, MHC-dependent mating preference was tested by relating female's male preference to either the total count of alleles in males, or mates allele sharing over all loci. Gravid female fish significantly preferred the odour of males with a large number of MHC class-IIB alleles to that of males with fewer alleles. In contrast with salmon, females did not prefer male genotypes dissimilar to their own.

Both fish studies better supported increased resistance of progeny to infectious disease through increased MHC diversity over inbreeding avoidance as a mechanism selecting for MHC-dependent male preference. Yet, recent experimental work has provided evidence that MHC can influence kin discrimination in salmonids, and therefore provide a potential marker for relatedness. Thus, Olsén *et al*. (1998) used fluvarium tests to examine whether kin recognition in juvenile Arctic char (*Salvelinus alpinus*) is influenced by polymorphism in exon 2 of the MHC class IIB gene. They showed that fish hierarchically preferred water scented by a MHC identical sibling, then water scented with an MHC different siblings, and finally, water scented with an MHC different nonsibling. More recently, Olsén *et al*. (2002) extended this study and showed that Arctic char isolated since fertilization did not show behavioural discrimination towards siblings, based on MHC genotypes. This result suggested that Arctic char may learn to discriminate between odours from individuals of different MHC types by imprinting, as previously shown for mice (Penn & Potts, 1998).

MHC-dependent survival and reproductive success

The previous sections provided evidence that the maintenance of MHC polymorphism in natural populations may be driven by selection, and that MHC is involved in

Table 2 A summary of the main results for studies that have investigated the role of MHC polymorphism in pathogen resistance and reproductive success in vertebrates.

Species	Observation	Proposed selection mechanism	Reference
Three-spined stickleback *Gasterosteus aculeatus*	Higher parasite diversity in population with higher Mhc variability Intermediate allele numbers at MHC class IIB associated with minimal parasitic load	Parasites drive MHC variability in natural populations An intermediate individual MHC diversity is optimal	Wegner *et al*. (2003)
Atlantic salmon *Salmo salar*	Survival to Aeromonas salmonicida is associated with single MHC class IIB alleles in an outbred population MHC-dependent survival is independent of family genetic background	Survival to an infection associated with specific alleles rather than to heterozygosity	Langefors *et al.* (2001)
	MHC-dependent survival is independent of family genetic background		Lohm *et al*. (2002)
Chinook salmon *Oncorhynchus tsawytscha*	MHC class IIB heterozygotes have higher survival when exposed to haematopoietic necrosis virus (IHNV)	Survival to infection is increased with heterozygosity	Arkush *et al.* (2002)
Gila topminnow *Poeciliopis o. occidentalis*	Homozygotes at a MHC class II B locus have a non statistically significant lower survival than heterozygous when infected with flukes (*Gyrodactylus turnbulli*)	Survival to a infection associated with increased heterozygosity	Hedrick *et al*. (2001b)
Soay sheep *Ovis aries*	Association between variation within the MHC, juvenile survival and resistance to intestinal nematodes	Resistance to a pathogen associated with specific alleles	Paterson *et al*. (1998)
Rhesus macaques *Macaca mulatta*	Heterozygous sire more offspring than homozygous	Heterozygotes have a higher reproductive success associated with higher disease resistance	Sauermann *et al*. (2001)

mate choice. The next step is to elucidate how an individual's fitness may be influenced by its MHC allelic composition. MHC polymorphism could first affect fitness by its direct influence on the progeny's survival to infectious diseases. As for MHC-dependent mate preference, few studies have attempted to test for an association between MHC polymorphism and disease resistance in wild vertebrates (Table 2). The most convincing case is a recent study on Atlantic salmon, in which Langefors *et al*. (2001) have illustrated the importance of genetic variation in the MHC class IIB gene on resistance to a highly pathogenic bacterium, *Aeromonas salmonicida*. In their experiment, full siblings from each of 120 families were infected with the bacterium. Fishes from high-resistance (HR, <35% mortality) and low-resistance (LR, >80% mortality) families were screened for their MHC class IIB genotypes. One allele was significantly more prevalent in HR families than in LR families such that broods carrying this allele had a 12-fold higher chance of being HR than broods without this allele. Two other alleles were also correlated with resistance to a lesser degree. The authors found no evidence that more resistant families were more heterozygous at MHC than more susceptible ones. In a subsequent study that used the progeny of uninfected siblings from Langefors *et al*.'s (2001) study, Lohm *et al*. (2002) designed full-sib broods consisting of combinations of heterozygote and homozygote genotypes with respect to resistance or susceptibility alleles identified by Langefors *et al*. (2001). By comparing full-siblings carrying different MHC genotypes, the effects on survival because of other segregating genes were minimized. This study first corroborated the differential level of resistance associated with specific alleles observed in the previous study. Moreover, it showed that a pathogenic bacterium can cause very intense selection pressure on particular MHC alleles. Indeed, Lohm *et al*. (2002) found that the relative fitness difference between individuals carrying different MHC alleles was as high as 0.5. As observed by Langefors *et al*. (2001) also, survival was not higher among heterozygous individuals. Instead, the results were more congruent with a dominant affect on disease resistance linked to one of the alleles. Thus, homozygotes at a particular allele had a higher risk of dying from infection than either heterozygotes or homozygotes for another allele, whereas no difference was observed between the latter two genotypic classes.

In another recent study on salmonids, Arkush *et al*. (2002) examined the differential MHC resistance in the Chinook salmon to three pathogens tested separately.

The results of this study differed from those on Atlantic salmon in several ways. First, the effect of MHC genotypes on resistance was less pronounced, with evidence for a genotypic effect on survival in only one of the pathogen tested (IHNV). In this case, the authors observed a higher average survival for heterozygotes (0.82) than homozygotes (0.75), which translated into a selection coefficient of 0.085. Moreover, there was apparently no evidence for an effect of specific MHC genotypes, although this observation could be confounded by variable levels of genetic relatedness among siblings from different families. Overall, their results support the heterozygote advantage hypothesis, but do not allow assessing whether this was caused by either an overdominant or a dominant effect.

In an earlier study, Paterson *et al*. (1998) investigated the association between MHC variation, juvenile survival, and infection to a nematode known to have important effect on the survivorship in a large unmanaged population of Soay sheep. Using microsatellite markers located within the MHC class II cluster, they showed that certain MHC alleles were significantly associated with low survivorship probabilities and a high incidence of parasitism and vice versa. These results supported the hypothesis that parasites are likely to play a role in the maintenance of MHC diversity in this population. As in the case of Atlantic salmon, however, the presence of individual alleles rather than heterozygosity was the critical factor determining mortality of juvenile sheep with respect to MHC variation.

Recently, the relationship between MHC variability within population and within individuals vs. the parasitic load has been addressed in natural populations of threespined sticklebacks *(Gasterosteus aculeatus L.)*. By comparing the level of variation at MHC class IIB loci and the number and diversity of parasites affecting different populations, Wegner *et al*. (2003) showed that the level of MHC diversity was correlated with parasitic diversity affecting stickleback populations, independently of the genome-wide heterozygosity as measured by microsatellites. Wegner *et al*. (2003) also showed that at the individual level, fish with an intermediate number of alleles had a lower parasitic load than both fish with high and low MHC class IIB diversity. Although apparently counterintuitive, this pattern is consistent with the model proposed by Nowak *et al*. (1992) according to which a high number of MHC alleles might not be an optimal solution for pathogen resistance. Although the effect of parasitic load on the fitness of different genotypes has still to be established, this study provides convincing evidence that MHC diversity observed in natural populations is driven by parasitic diversity.

MHC could also relate to fitness by influencing the reproductive success of individuals *per se*. In the only study that tested for an association between MHC polymorphism and reproductive success, Sauermann *et al*. (2001) evaluated the influence of MHC Class II-DQB1 genotypes on the number of offspring sired in rhesus macaques living in a seminatural environment. By correlating the number of progenies that males sired with their MHC class II DQB1 locus genotype, the authors found that males heterozygous at this locus sired significantly more offspring than homozygous males, which translated into a very high selection coefficient ($s = 0.34$). In order to assess whether the observed male heterozygote advantage was explicable in terms of genotype-dependent bias in mating preference, the authors also quantified the level of allele sharing between mothers and sires. As the pattern did not differ significantly from random expectation, the authors concluded that there was no evidence that mother–sire genotype similarity conferred a reproductive advantage or disadvantage. However, a marginally nonsignificant pattern ($P = 0.07$) was observed in one of the two groups studied. Furthermore, the authors did not test for a possible genotype-dependent bias in mating preference in terms of amino-acid allelic composition (e.g. Landry *et al*. 2001). Thus, while Sauermann *et al*. (2001) argued that the most plausible explanation for increased reproductive success in heterozygous males is increased resistance to the debilitating effects of parasite infection, a possible role for MHC-dependent mating preference conferring a reproductive advantage has not been strictly ruled out.

Discussion and future prospects

What have we learned from tests of positive selection over evolutionary time scale?

Most studies have tested for the possible role of selection in shaping MHC diversity. In these, a strong bias was found in the type of tests applied, with the vast majority reporting results for the standard dn/ds test ratio only. A fair number of those studies ($n = 22$, Supplementary material on the web), however, did not statistically test for significant departure from neutral expectation, which limits their usefulness in rigorously inferring the possible effect of selection on MHC diversity. Studies that documented trans-species polymorphism are also relatively common. However, the use of more elaborate methods, such as the coalescence-based test of alternative overdominance models is still restricted to a very limited number of studies. Moreover, recent approaches based on codon-based likelihood models that potentially provide more accurate estimates of dn/ds ratios than do simpler approximation methods, and that also allow specific selected sites to be identified (reviewed in Yang & Bielawski, 2000) are still awaiting application in MHC studies of natural vertebrate populations (but see Cohen, 2002). Clearly, future tests of positive selection would greatly benefit from applying more statistical rigour to data analyses, as well as making a better use of the increasing number of models potentially offering higher

resolution in detecting the effect and form of selection. Despite these few caveats, however, the general view that is emerging from the studies we reviewed is that the determinant role of positive selection in shaping patterns of nucleotide diversity in MHC genes can be generalized to all vertebrates studied.

What have we learned from tests of selection within and among contemporary populations?

Studies that compared patterns of MHC diversity within and among natural populations with neutral expectations have indicated that the types of selection acting on MHC variation in natural populations differ from the traditional view of a simple model of balancing selection in which all heterozygotes are of equivalent fitness (and superior to homozygotes) over space and time. Muirhead (2001) analytically showed that observation of similar or higher F_{st} estimates at MHC compared with neutral loci is very unlikely under symmetrical balancing selection in which case, the effective migration rate should be higher than under neutrality. In contrast to this view, studies on salmonid fishes showed that the level of population differentiation at MHC may be higher than observed at neutral loci, therefore indicating a lower effective migration rate at MHC. Moreover, the differential pattern of structuring at both types of markers was geographically heterogeneous, which supports the hypothesis that the environments in which distinct populations live differ in intensity of balancing selection. Particular subsets of alleles may be locally favoured but limited in their success at invading other populations, thus resulting in higher population differentiation than for neutral markers. Pathogens could be responsible for this pattern, whereby environments with high level of pathogens abundance likely experience relatively high selection pressures for maintaining polymorphism at MHC compared with environments with low rate of infectious disease (Miller *et al*., 2001). Although this may prove to be a challenging task (Jeffrey & Bangham, 2000; but see Wegner *et al*., 2003), performing studies to associate patterns of genetic differentiation with differential gradients of infectious disease across environments may allow better inferring of how local adaptation may be responsible for maintaining MHC diversity. This will be best achieved by contrasting variation in patterns of diversity at MHC and neutral loci over time with predictions of models describing the interactions between population subdivision and balancing selection (e.g. Muirhead, 2001).

What have we learned from studies on MHC-dependent mate preference and kin recognition?

Research on mice has clearly shown that animals can discriminate and select mates based on their MHC genotypes (Tregenza & Wedell, 2000; Penn, 2002). These studies have also generally better supported the hypothesis that MHC-disassortative mating preferences function primarily to prevent kin matings (inbreeding avoidance hypothesis) than to increase the resistance of progeny to infectious diseases (Penn, 2002). However, the experimental conditions in most mice studies were conducted with few congenic, highly inbred, and strains with limited diversity in MHC haplotypes were used. Thus, it is unclear how conclusions of such studies may apply to wild, outbred populations. With the exception of Soay sheep, studies on wild fish, mammals, and birds performed thus far showed that MHC-dependent mate preference and kin recognition is not unique to mice but may provide a selective factor maintaining polymorphism in wild outbred populations as well. However, studies on salmon and stickleback were in contrast with observations on mice, as they better supported the hypothesis that MHC-disassortative mating preference is selected more for enhancing disease resistance through increased MHC diversity rather than as a mechanism for avoiding inbreeding. Moreover, the patterns observed differed substantially in both species. In salmon, mates significantly preferred to reproduce with fish with dissimilar alleles in terms of overall amino-acid composition, whereas female stickleback preferred the odour of males with a large number of alleles to that of males with fewer alleles. This suggests that the selective mechanism favouring disassortative mating may differ between both species.

Altogether, the above studies show that MHC-dependent mate preference can be empirically detected in the wild and may operate in all classes of vertebrates, even if the potential fitness benefits may be indirect. However, these studies also indicate that this reproductive mechanism is complex and context-based, possibly depending on the specific genetic architecture of the MHC (e.g. variable number of duplicated loci for a given gene class), and the selective pressures that are acting (Jordan & Bruford, 1998). For instance, it is possible that the strength of selection for inbreeding avoidance varies according to the characteristic of a given population or species relative to the *inbred–outbred* continuum. Depending on local pathogen abundance and diversity, females could also prefer males able to transmit alleles that confer more resistance to specific diseases, whereas in other cases, mating events resulting in disparate allelic combination in the progeny may be positively selected. It is also plausible that there is little opportunity for MHC-dependent female choice when other reproductive mechanisms are interacting (e.g. male–male competition for mates). This may explain why no MHC-dependent mating preference was detected in Soay sheep despite the fact that an association between variation within the MHC, juvenile survival and pathogen resistance was documented in this species (Paterson & Pemberton, 1997).

In summary, the number of studies that investigated MHC-disassortative mating preferences is still very limited, and therefore, many more are needed in order to better document the nature of this reproductive

mechanism in wild populations. Under the hypothesis that its function is condition-dependent, a comparative approach, whereby populations or closely related species that differ in inbreeding, competitive or social conditions are contrasted for their MHC-disassortative mating preferences should be particularly warranted. Such studies should also be designed so as to decipher the possible multivariate interactions between MHC and other honest signals of good genes in determining mate choice.

What have we learned about MHC-dependent fitness?

The few studies performed to date provided evidence that MHC may significantly affect fitness in organisms other than mice and humans, either by increasing reproductive success or progeny survival to pathogens infections. These also showed that the MHC-dependent selection coefficient may be very high, which provides further support for the role of selection in shaping patterns of MHC diversity in natural populations. Perhaps the most important result, however, is that experimental quantification of survival to pathogens infection provided little support for heterozygote advantage through overdominance. Instead, studies on Atlantic salmon and sheep showed that survival to infection was associated with specific alleles rather than to heterozygosity. These observations concur with the emerging view that MHC-dependent resistance in mice is generally dominant rather than overdominant (Penn, 2002). As argued by Penn (2002) also, dominance is enough to provide an advantage for heterozygotes as long as their average fitness is superior to the average value for homozygotes. However, documenting rigorously whether resistance is sufficiently dominant to provide an advantage for heterozygotes requires the quantification of differential fitness among MHC genotypes against a simultaneous infection with multiple pathogens. In the first study of multiple-strain infection in mice, Penn *et al.* (2002) indeed showed that the average fitness of heterozygotes was higher than that of homozygotes. This provided support for the hypothesis of *heterozygote advantage through dominance,* whereby the benefits of heterozygosity were caused by resistance being dominant rather than overdominant. This could also explain the correlation between parasite load and multilocus MHC allelic diversity reported for stickleback by Wegner *et al.* (2003). Clearly, there is an urgent need for more studies in natural conditions in order to assess the generality of this hypothesis. A fruitful extension to this research would be to empirically estimate the fitness consequence of MHC-dependent disassortative mating in wild populations by quantifying differential progeny survival as a function of parental MHC combinations. It would be particularly relevant to test whether all heterozygotes confer similar selective advantages, or whether there are systematic differences among heterozygotes carrying more disparate alleles. Such studies in a natural context can now be achieved thanks to recent developments in parentage assignment procedures (e.g. Marshall *et al.*, 1998; Duchesne *et al.*, 2002).

Despite the fact that the number of MHC studies on wild populations is still small, and that MHC-dependent mating preferences and the type of selective mechanisms maintaining MHC diversity are generally believed to be difficult to detect, progress achieved to date is encouraging. Overall, the evidence is compelling that the MHC currently represents the best system available in vertebrates to investigate how natural selection promotes local adaptation at the gene level despite the counteracting actions of migration and genetic drift. In addition to the several research directions proposed above, future MHC studies should also aim at elucidating how the effect of selection will be affected by the interspecific variation in MHC architecture (both in terms of number of loci and their chromosomal distribution), as well as variation in inter-locus interactions such as recombination and epistasis. This is particularly a crucial issue as recent simulation studies have shown that the effect of balancing selection on nucleotide diversity in single-locus systems does not extend to multilocus scenario in many circumstances (Navarro & Barton, 2002). Evolutionists have now reached a consensus that many of the key adaptive differences between organisms will not only be in the form of allelic variation at specific genes but will also be manifested as changes in gene expression (Streelman & Kocher, 2000). Genes of the MHC should not be exception to this view as extensive polymorphism at MHC regulatory sequences in humans tends to suggest (Guardiola *et al.*, 1996; Cowel *et al.*, 1998). Thus, one of the most exciting avenue of research on MHC will be to elucidate how specific environmental conditions and genotypic combination interact in determining levels of MHC gene expression, and to document the consequences of such interactions on mate choice and the resulting individual fitness.

Acknowledgments

We are grateful to Roger Butlin for inviting us to publish this review, as well as to T.B.H Reusch and one anonymous reviewer for their constructive comments. L.B.'s research is supported by the Canadian Research Chair in conservation genetics of aquatic organisms and various grants from the Natural Science and Engineering Council (NSERC, Canada). C. Landry is supported by NSERC, Frank Knox Memorial and FCAR Postgraduate Scholarships.

Supplementary material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/JEB/JEB531/JEB531sm.htm

Table S1 Studies on MHC in nonmodel vertebrates, with evidence of selection and major findings.

References

Apanius, V., Penn, D., Slev, P.R., Ruff, L.R. & Potts, W.K. 1997. The nature of selection on the major histocompatibility complex. *Crit. Rev. Immunol.* **17**: 179–224.

Arkush, K.D., Giese, A.R., Mendonca, H.L., McBride, A.M., Marty, G.D. & Hedrick, P.W. 2002. Resistance to three pathogens in the endangered winter-run chinook salmon (*Oncorhynchus tsawytscha*): effects of inbreeding and major histocompatibility complex genotypes. *Can. J. Fish Aquat. Sci.* **59**: 966–975.

Bakke, T.A. & Harris, P.D. 1998. Diseases and parasites in wild Atlantic salmon (*Salmo salar*) populations. *Can. J. Fish. Aquat. Sci.* **55**: 247–266.

Bergstrom, T.F., Josefsson, A., Erlich, H.A. & Gyllensten, U. 1998. Recent origin of HLA-DRB1 alleles and implication for human evolution. *Nat. Genet.* **18**: 237.

Bingulac-Popovic, J., Figueroa, F., Sato, A., Talbot, W.S., Johnson, S.L., Gates, M., Postlethwait, J.H. & Klein, J. 1997. Mapping of mhc class I and class II regions to different linkage groups in the zebrafish, *Danio rerio*. *Immunogenetics* **46**: 129–134.

Bodmer, W. 1972. Evolutionary significance of the HLA-system. *Nature (London)* **237**: 139–145.

Boyce, W.M., Hedrick, P.W., Muggli-Cockett, N.E., Kalinowski, S., Penedo, M.C.T. & Ramey, R.R. II. 1997. Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. *Genetics* **145**: 421–433.

Clarke, B.C. & Kirby, D.R.S. 1966. Maintenance of histocompatibility polymorphism. *Nature* **211**: 999–1000.

Cohen, S. 2002. Strong positive selection and habitat specific substitution patterns in Mhc from an estuarine fish under intense pollution stress. *Mol. Biol. Evol.* **19**: 1870–2880.

Cowell, L.G., Kepler, T.B., Janitz, M., Lauster, R. & Mitchison, N.A. 1998. The distribution of variation in regulatory gene segments, as present in MHC class II promoters. *Genome Res.* **8**: 124–134.

Dean, M., Carrington, M. & O'Brien, S.J. 2002. Balanced polymorphism selected by genetic versus infectious human disease. *Annu. Rev. Genomics. Hum. Genet.* **3**: 263–292.

Ditchkoff, S.S., Lochmiller, R.L., Masters, R.E., Hoofer, S.R. & Van Den Bussche, R.A. 2001. Major-histocompatibility-complex-associated variation in secondary sexual traits of white-tailed deer (*Odocoileus virginianus*): evidence for good-genes advertisement. *Evolution* **55**: 616–625.

Dixon, B., Nagelkerke, L.A.J., Sibbing, F.A., Egberts, E. & Stet, R.J.M. 1996. Evolution of MHC classII B chain-encoding genes in the Lake Tana barbel species flock (*Barbus intermedius* complex). *Immunogenetics* **44**: 419–431.

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

Duchesne, P., Godbout, M.-H. & Bernatchez, L. 2002. PAPA (Package for the Analysis of Parental Allocation): a computer program for simulated and real parental allocation. *Mol. Ecol. Notes.* **2**: 191–194.

Edwards, S.V. & Hedrick, P.W. 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol. Evol.* **13**: 305–311.

Ewens, W.J. 1972. The sampling theory of selectively neutral alleles. *Theo. Pop. Bio.* **3**: 87–112.

Figueroa, F., Gunther, E. & Klein, J. 1988. MHC polymorphism pre-dating speciation. *Nature* **335**: 265–267.

Flajnik, M.F., Kasahara, M., Shum, B.P., Salter-Cid, L., Taylor, E. & Du Pasquier, L. 1993. A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian Xenopus. *EMBO J.* **12**: 4385–4396.

Ford, M.J. 2002. Applications of selective neutrality tests to molecular ecology. *Mol. Ecol.* **11**: 1245–1262.

Garant, D., Dodson, J.J. & Bernatchez, L. 2000. Ecological determinants and temporal stability of whitin-river population structure in Atlantic salmon (*Salmo salar* L.). *Mol. Ecol.* **9**: 615–628.

Garrigan, D. & Hedrick, P.W. 2001. Class I MHC polymorphism and evolution in endangered California Chinook and other Pacific salmon. *Immunogenetics* **53**: 483–489.

Guardiola, J., Maffei, A., Lauster, R., Mitchison, N.A., Accolla, R.S. & Sartoris, S. 1996. Functional significance of polymorphism among MHC class II gene promoters. *Tissue Antigens* **48**: 615–625.

Hambuch, T.M. & Lacey, E.A. 2002. Enhanced selection for MHC diversity in social tuco-tucos. *Evolution* **56**: 841–845.

Hansen, J.D., Strassburger, P., Thorgaard, G.H., Young, W.P. & Du Pasquier, L. 1999. Expression, linkage, and polymorphism of MHC-related genes in rainbow trout, *Oncorhynchus mykiss*. *J. Immunol.* **163**: 774–786.

Hedrick, P.W. 1994. Evolutionary genetics of the major histocompatibility complex. *Am. Nat.* **143**: 945–964.

Hedrick, P.W. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**: 313–318.

Hedrick, P.W., Kim, T.J. & Parker, K.M. 2001b. Parasite resistance and genetic variation in the endangered Gila topminnow. *Anim. Cons.* **4**: 103–109.

Hedrick, P.W., Parker, K.M. & Lee, R.N. 2001a. Using microsatellite and MHC variation to identify species, ESUs, and MUs in the endangered Sonoran topminnow. *Mol. Ecol.* **10**: 1399–1412.

Hess, C.M. & Edwards, S.V. 2002. The evolution of the major histocompatibility complex in birds. *Bioscience* **52**: 423–431.

Hill, R.E. & Hastie, N.D. 1987. Accelerated evolution in the reactive centre regions of serine protease inhibitors. *Nature* **326**: 96–99.

Hogstrand, K. & Bohme, J. 1994. A determination of the frequency of gene conversion in unmanipulated mouse sperm. *Proc. Natl. Acad. Sci. USA* **91**: 9921–9925.

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

Hughes, A.L. & Yeager, M. 1998. Natural selection and the evolutionary history of major histocompatibility complex loci. *Front. Biosci.* **3**: d509–516.

Jeffery, K.J. & Bangham, C.R. 2000. Do infectious diseases drive MHC diversity? *Microbes. Infect.* **2**: 1335–1341.

Jordan, W.C. & Bruford, M.W. 1998. New perspectives on mate choice and the MHC. *Heredity* **81**: 239–245.

Kim, T.J., Parker, K.M. & Hedrick, P.W. 1999. Major histocompatibility complex differentiation in Sacramento River chinook salmon. *Genetics* **151**: 1115–1122.

Klein, J. 1986. *The Natural History of the Major Histocompatibility Complex*. John Wiley & Sons, NY, USA.

Klein, J. 1987. Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Hum. Immunol.* **19**: 155–162.

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*). *Mol. Ecol.* **10**: 2525–2539.

Landry, C., Garant, D., Duchesne, P. & Bernatchez, L. 2001. 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proc. R. Soc. Lond. B* **268**: 1279–1285.

Langefors, A., Lohm, J., Grahn, M., Andersen, O. & von Schantz, T. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proc. R. Soc. Lond. B* **268**: 479–485.

Lohm, J., Grahn, M., Langefors, A., Adersen, O., Storset, A. & von Schantz, T. 2002. Experimental evidence for major histocompatibiltiy complex-allele-specific resistance to a bacterial infection. *Proc. R. Soc. Lond. B.* **2114**: 2029–2034.

Malaga-Trillo, E., Zaleska-Rutczynska, Z., McAndrew, B., Vincek, V., Figueroa, F., Sultmann, H. & Klein, J. 1998. Linkage relationships and haplotype polymorphism among cichlid Mhc class II B loci. *Genetics* **149**: 1527–1537.

Marshall, T.C., Slate, J., Kruuk, L.E.B. & Pemberton, J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7**: 639–655.

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

Meyer, D. & Thomson, G. 2001. How selection shapes variation on the human major histocompatibility complex: a review. *Ann. Hum. Gen.* **65**: 1–26.

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

Miller, K. & Withler, R. 1997. MHC diversity in Pacific salmon: population structure and trans-species allelism. *Hereditas* **127**: 163–164.

Miller, K.M. & Withler, R.E. 1998. The salmonids class I MHC: limited diversity in a primitive teleost. *Immuno. Rev.* **166**: 279–293.

Muirhead, C.A. 2001. Consequences of population structure on genes under balancing selection. *Evolution* **55**: 1532–1541.

Navarro, A. & Barton, N.H. 2002. The effects of multilocus balancing selection on neutral variability. *Genetics* **161**: 849–863.

Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. *Heredity* **86**: 641–647.

Nowak, M.A., Tarczy-Hornoch, K. & Austyn, J.M. 1992. The optimal number of major histocompatibility complex molecules in an individual. *Proc. Natl. Acad. Sci. USA* **89**: 10 896–10 899.

Olsén, K.H., Grahn, M. & Lohm, J. 2002. Influence of MHC on sibling discrimination in Arctic char, *Salvelinus alpinus* (L.). *J. Chem. Ecol.* **28**: 783–795.

Olsén, K.H., Grahn, M., Lohm, J. & Langefors, A. 1998. MHC and kin discrimination in juvenile Arctic charr, *Salvelinus alpinus* (L.). *Anim. Behav.* **56**: 319–327.

Paterson, S. 1998. Evidence for balancing selection at the major histocompatibility complex in a free-living ruminant. *J. Hered.* **89**: 289–294.

Paterson, S. & Pemberton, J.M. 1997. No evidence for major histocompatibility complex-dependent mating patterns in a free-living ruminant population. *Proc. R. Soc. Lond. B* **264**: 1813–1819.

Paterson, S., Wilson, K. & Pemberton, J.M. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries L.*). *Proc. Natl. Acad. Sci. USA* **95**: 3714–3719.

Penn, D.J. 2002. The scent of genetic compatibility: sexual selection and the major histocompatibility complex. *Ethology* **108**: 1–21.

Penn, D.J., Damjanovich, K. & Potts, W.K. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc. Natl. Acad. Sci. USA* **99**: 11 260–11 264.

Penn, D. & Potts, W.K. 1998. Chemical signals and parasite-mediated sexual selection. *Trends Ecol. Evol.* **13**: 391–396.

Penn, D.J. & Potts, W.K. 1999. The evolution of mating preference and major histocompatibility complex genes. *Am. Nat.* **153**: 145–164.

Ploegh, H. & Watts, C. 1998. Antigen recognition. *Curr. Opin. Immunol.* **10**: 57–58.

Potts, W.K. & Wakeland, E.K. 1990. Evolution of diversity at the major histocompatibility complex. *Trends Ecol. Evol.* **5**: 181–186.

Reusch, T.B., Haberli, M.A., Aeschlimann, P.B. & Milinski, M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature* **414**: 300–302.

Richman, A. 2000. Evolution of balanced genetic polymorphism. *Mol. Ecol.* **9**: 1953–1963.

Richman, A.D., Herrera, L.G. & Nash, D. 2001. MHC class II beta sequence diversity in the deer mouse (*Peromyscus maniculatus*): implications for models of balancing selection. *Mol. Ecol.* **10**: 2765–2773.

Richman, A.D. & Kohn, J.R. 1999. Self-incompatibility alleles from Physalis: implications for historical inference from balanced genetic polymorphisms. *Proc. Natl. Acad. Sci. USA* **96**: 168–172.

Sammut, B., Marcuz, A. & Pasquier, L.D. 2002. The fate of duplicated major histocompatibility complex class Ia genes in a dodecaploid amphibian, *Xenopus ruwenzoriensis*. *Eur. J. Immunol.* **32**: 1593–1604.

Sauermann, U., Nurnberg, P., Bercovitch, F.B., Berard, J.D., Trefilov, A., Widdig, A., Kessler, M., Schmidtke, J. & Krawczak, M. 2001. Increased reproductive success of MHC class II heterozygous males among free-ranging rhesus macaques. *Hum. Gen.* **108**: 249–254.

Schierup, M.H. 1998. The number of self-incompatibility alleles in a finite, subdivided population. *Genetics* **149**: 1153–1162.

Schierup, M.H., Vekemans, X. & Charlesworth, D. 2000. The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genet. Res.* **76**: 51–62.

Slatkin, M. 1994. An exact test for neutrality based on the Ewens sampling distribution. *Genet. Res.* **64**: 71–74.

Slatkin, M. & Muirhead, C.A. 2000. A method for estimating the intensity of overdominant selection from the distribution of allele frequencies. *Genetics* **156**: 2119–2126.

Stet, R.J.M. & Egberts, E. 1991. The histocompatibility system in teleostean fishes: from multiple histocompatibility loci to

a major histocompatibility complex. *Fish Shell. Immunol.* **1**: 1–16.

Streelman, J.T. & Kocher, T.D. 2000. From phenotype to genotype. *Evol. Dev.* **2**: 166–173.

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

Tregenza, T. & Wedell, N. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. *Mol. Ecol.* **9**: 1013–1027.

Van Eden, W., Devries, R.R.P. & Van Rood, J.J. 1983. The genetic approach to infectious disease with special emphasis on the MHC. *Dis. Mark.* **1**: 221–242.

Van Muiswinkel, W.B., Wiegertjes, G.F. & Stet, R.J.M. 1999. The influence of environmental and genetic factors on the disease resistance of fish. *Aquaculture* **172**: 103–110.

Von Schantz, T., Goransson, G., Andersson, G., Froberg, I., Grahn, M., Helgee, A. & Wittzell, H. 1989. Female choice selects for a viability-based male trait in pheasant. *Nature* **337**: 166–169.

Von Schantz, T., Wittzell, H., Goransson, G. & Grahn, M. 1997. Mate choice, male condition-dependent ornamentation and MHC in the pheasant. *Hereditas* **127**: 133–140.

Von Schantz, T., Wittzell, H., Goransson, G., Grahn, M. & Persson, K. 1996. MHC genotype and male ornamentation: genetic evidence for the Hamilton–Zuk model. *Proc. R. Soc. Lond. B* **263**: 265–271.

Watkins, D.I., Hodi, F.S. & Letvin, N.L. 1988. A primate species with limited major histocompatibility complex class I polymorphism. *Proc. Natl. Acad. Sci. USA* **85**: 7714–7718.

Watterson, G.A. 1978. The homozygosity test of neutrality. *Genetics* **88**: 405–417.

Wegner, K.M., Reusch, T.B.H. & Kalbe, M. 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J. Evol. Biol.* **16**: 224–232.

Yang, Z. & Bielawski, B. 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**: 496–503.

Received 22 October 2002; revised 18 December 2002; accepted 15 January 2003