Trends in Genetics

CellPress
REVIEWS

Review

Advances in the Evolutionary Understanding of *MHC* Polymorphism

Jacek Radwan,[1,*] Wiesław Babik,[2,8] Jim Kaufman,[3,4,7,8] Tobias L. Lenz,[5,8] and Jamie Winternitz[6,8]

Proteins encoded by the classical major histocompatibility complex (*MHC*) genes incite the vertebrate adaptive immune response by presenting peptide antigens on the cell surface. Here, we review mechanisms explaining landmark features of these genes: extreme polymorphism, excess of nonsynonymous changes in peptide-binding domains, and long gene genealogies. Recent studies provide evidence that these features may arise due to pathogens evolving ways to evade immune response guided by the locally common *MHC* alleles. However, complexities of selection on *MHC* genes are simultaneously being revealed that need to be incorporated into existing theory. These include pathogen-driven selection for antigen-binding breadth and expansion of the *MHC* gene family, associated autoimmunity trade-offs, hitchhiking of deleterious mutations linked to the *MHC*, geographic subdivision, and adaptive introgression.

Highlights

Novel MHC alleles have been demonstrated to confer better resistance to local parasites.

MHC alleles may differ by orders of magnitude in the range of antigens they bind.

Promiscuous alleles and species with more MHC genes appear to be more common in pathogen-rich populations.

The number of MHC class I alleles correlates negatively with the size of the T cell receptor repertoire, supporting the role of constraints associated with increasing ranges of bound antigens.

Deleterious mutations accumulate around MHC genes and likely affect the evolutionary dynamics of MHC haplotypes.

MHC: The Enigma Continues

The MHC is a gene-dense region in jawed vertebrate genomes enriched for immunity genes. The classical *MHC* genes, which will be the subject of this review, encode glycoproteins that bind peptides, both self and non-self, inside the cell and deliver them to the surface for inspection by T cells and natural killer (NK) cells [1,2] (Boxes 1 and 2). This antigen presentation is a crucial step in the adaptive immune response as it allows self/non-self discrimination by T cells, ultimately facilitating the recognition of infecting pathogens. The feature that distinguishes classical *MHC* genes (*MHC* genes hereafter) from other genes in the MHC region is their extreme polymorphism, with dozens to hundreds of allelic variants segregating in natural populations [3–5]. The polymorphism is most pronounced in the **peptide-binding domain** (**PBD**; see Glossary), in particular at **peptide-binding sites (PBSs)**, amino-acid residues interacting directly with antigens [6]. Consequently, molecules coded by different *MHC* alleles differ in their antigen-binding profiles [7,8], which in turn affect susceptibility to disease [9–11]. Polymorphism apparently evolves **adaptively**, as evidenced by the high relative nonsynonymous substitution rate within the PBD [12], particularly at PBSs [6,13,14], as well as by large short-term **selection coefficients** (Figure 1). High polymorphism coupled with evidence for **positive selection** has made *MHC* genes an attractive model for studying how selection can promote and maintain genetic variation in natural populations.

Evidence is accumulating, as has long been suspected based on the function of MHC proteins, that pathogens impose significant selection on *MHC* (Figure 1) and, importantly, drive *MHC* allele frequency changes in natural populations [3,15]. However, the specific selection mechanisms that shape the extraordinary diversity of *MHC* genes are still controversial (Figure 2, Key Figure). An associated question is whether these mechanisms can explain the evolutionary persistence of *MHC* allelic lineages for a much longer time than expected under neutrality, leading to **trans-species polymorphism (TSP)** [16,17], and an excess of nonsynonymous changes in *MHC* sequences. Yet another enigma is why *MHC* diversity at the individual level is limited, constraining an individual's ability to raise an effective response to parasites, even though expressing more MHC molecules or molecular variants capable of binding a broader spectrum of antigens

[1]Evolutionary Biology Group, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland
[2]Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Krakow, Poland
[3]Department of Pathology, University of Cambridge, Cambridge, UK
[4]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK
[5]Research Group for Evolutionary Immunogenomics, Max Planck Institute for Evolutionary Biology, Ploen, Germany
[6]Department of Animal Behavior, Bielefeld University, Bielefeld, Germany
[7]Current address: University of Edinburgh, School of Biological Sciences, Institute for Immunology and Infection Research, Ashworth Laboratories, Edinburgh, UK
[8]These authors are listed in alphabetical order

*Correspondence:
jradwan@amu.edu.pl (J. Radwan).

Box 1. The Major Histocompatibility Complex

The MHC was discovered as the genetic locus leading to rapid graft rejection, which is due to highly polymorphic cell surface molecules encoded by classical class I (*MHC*-I) and class II (*MHC*-II) genes [95]. For most jawed vertebrates, classical *MHC* genes are scattered throughout a large genomic region with variable levels of recombination (Figure I) along with many other genes, including some involved in classical MHC function (such as *TAP* and tapasin genes). However, classical *MHC* genes are found on several chromosomes in teleost (bony) fish, and the *MHC*-I gene family is highly expanded with loss of *MHC*-II genes in Gadiform fish like the Atlantic cod [1,2].

Classical MHC molecules bind pieces of proteins (peptides) from inside cells and allow them to be recognized on the cell surface by T lymphocytes. The peptides are bound by pockets (PBSs) in a groove (part of the PBD) exhibiting extensitve sequence variation. Normally, these peptides are derived from self (host) proteins, but upon infection (or transformation in cancer), MHC molecules present non-self (pathogen or mutated) peptides to the T cells, leading to appropriate immune responses. Class I molecules bind peptides largely from the cytoplasm and contiguous structures like the nucleus, and are recognized by CD8 cytotoxic T lymphocytes. Class II molecules bind peptides largely from intracellular vesicles which are in contact with the outside of the cell, and are recognized by CD4 helper (and regulatory) lymphocytes [96]. MHC-I molecules are also recognized by polymorphic receptors on NK cells (see Box 2).

To achieve discrimination between self and non-self, T cells are 'educated' in the thymus: T cells, which were first positively selected based on reaction with self-peptides on the MHC molecules, are then negatively selected against strong recognition with self-peptides [97].

Nonclassical genes are related to classical *MHC* genes (and often difficult to distinguish from them on the basis of sequence alone) but lack one or more of their salient features – high polymorphism, wide and high expression, and presentation of peptides to T cells. Their functions vary from immune functions of many different kinds to non-immune physiology [98,99].

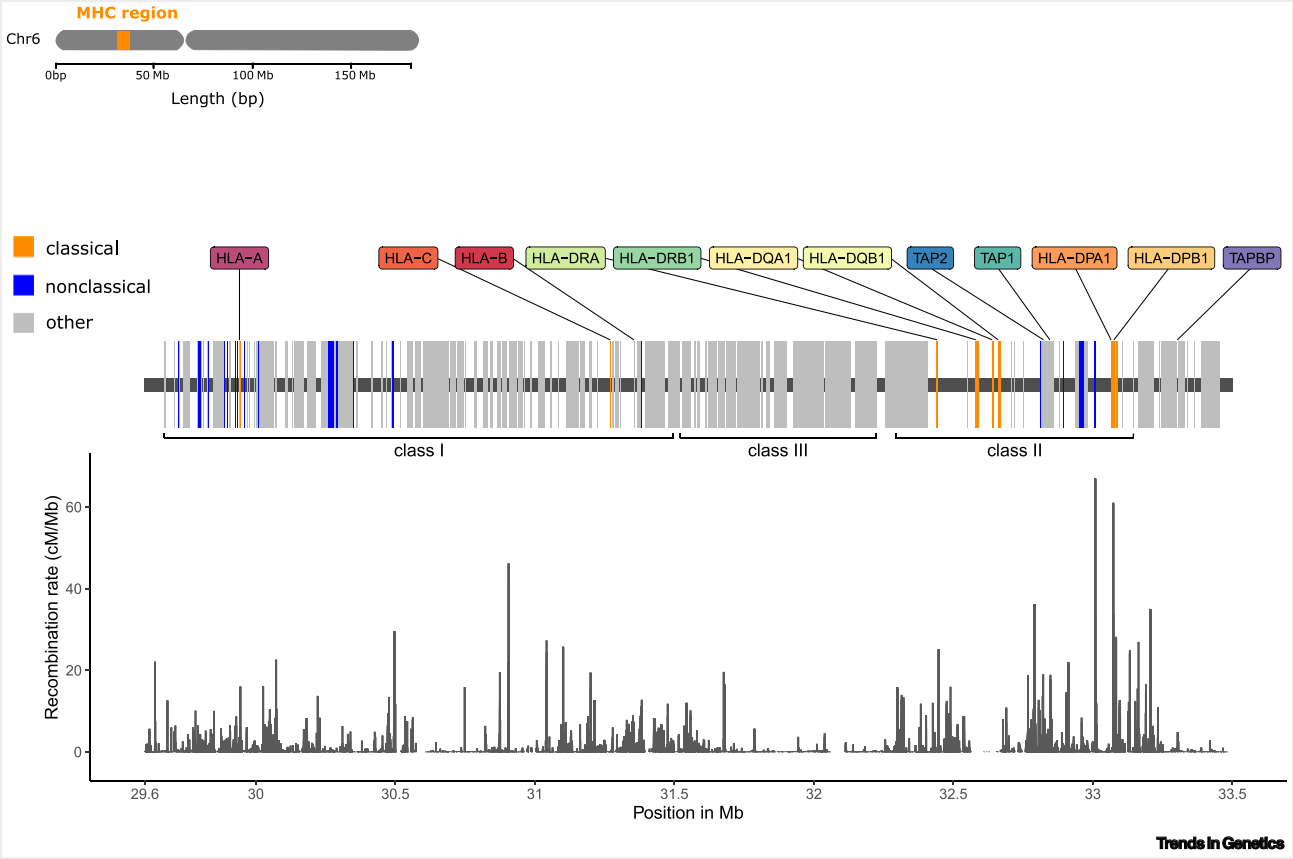

Figure I. Genomic Map of Classical Human MHC (HLA) Region with Associated Recombination Rate. (Top) The classical MHC region is approximately 3.5 Mb, comprising more than (middle) 280 genes, including those of the classical and nonclassical class I, II, and III genes. The mean recombination rate (bottom) in the class I region (0.443 cM/Mb) is lower, and in the class II region (1.712 cM/Mb) higher than the genomic average (1.2 cM/Mb); the recombination rate varies widely throughout the region (range 0.001–67 cM/Mb), which includes hotspots of extreme recombination [126,127].

Box 2. Interactions between MHC, T Cell, and Natural Killer Cell Receptors

TCRs interact with peptides bound by MHC molecule, as well as parts of PBD (formed by α1 and α2 chains in case of MHC-I, Figure I). Some classical (and nonclassical) MHC-I molecules can be ligands for NK and myeloid cells, resulting in another level of selection beyond T cells. Some NK receptors (NKRs), notably the killer inhibitor receptors (KIRs) found in humans, also recognize portions of bound peptide, potentially influencing the PBSs [100,101].

NK cells can kill cells that lose cell surface expression of MHC-I molecules due to viral infection, cancer, or even stress. However, some NKRs evolved to recognize decoy MHC-I molecules (co-opted by viruses to prevent killing) and they activate killing. Moreover, NKRs found on T cells are involved in driving cell proliferation, as are the leukocyte immunoglobulin-like receptor (LILR, synonym LIR) molecules of myeloid cells and lymphocytes. In addition, some NK cells in humans and mice bind certain MHC-I molecules to affect proliferation of invasive trophoblasts in the placenta [102,103].

NKR systems typically evolve very rapidly, with both copy number variation and high allelic polymorphism. Some species have predominantly lectin-like NKRs (like Ly49 in mice), others have predominantly immunoglobulin (Ig)-like NKRs (like KIRs in humans), and still others have both (like cattle) or neither (like marine mammals). In humans, the *LILR* genes are located next to the *KIR* genes, to which they are related [102–104]. The various receptors recognize (and thus put selective pressure on) different parts of the MHC-I molecules (Figure I): lectin-like receptors bind under the PBD, KIRs bind the top of the PBD including the C-terminal end of the bound peptide, and LILRs typically bind the MHC α3 domain and small subunit β2-microglobulin [101].

Given that both the NKR and MHC genetic systems are highly polymorphic but located on different chromosomes, there can be epistasis strongly affecting traits like resistance to infectious disease, susceptibility to autoimmunity, and aspects of reproduction. For instance, *HLA*-C expressed in human fetal trophoblasts are recognized by KIRs on maternal NK cells, with the strength of interaction between particular paternal *HLA*-C alleles and particular maternal *KIR* alleles eventually determining the blood supply to the developing embryo and pregnancy success (see Box 4) [103].

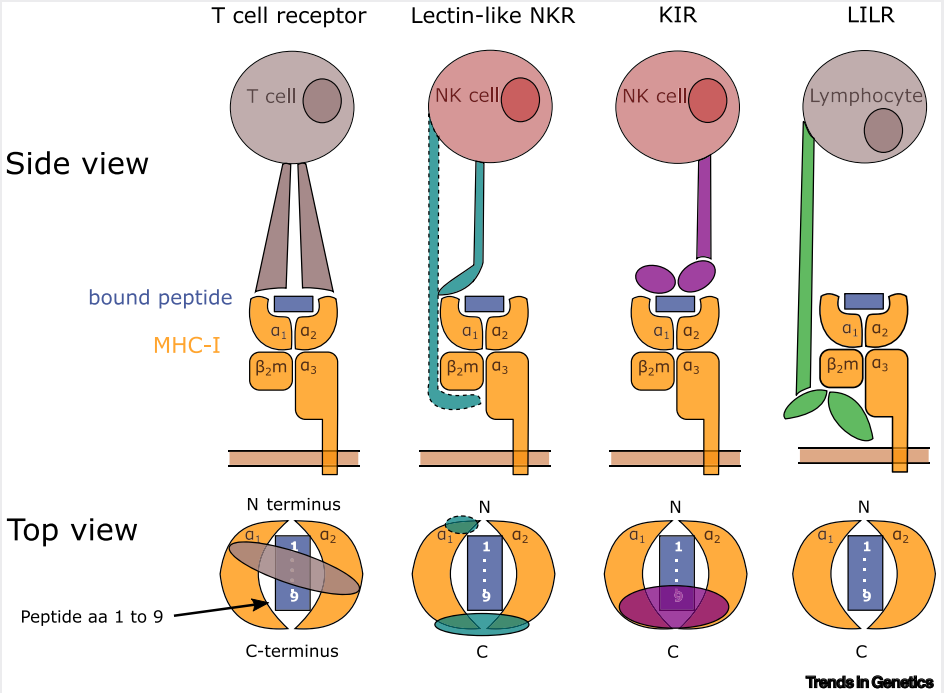

Figure I. Schematic Illustration of Interactions between the MHC-I, Bound Peptide, and T Cell and NK Cell Receptors. The MHC-I molecules and their interactions with various receptors are shown from their side and top views. The nine amino acids (aa) of the peptide (in blue) interact with the MHC-I PBD formed by the α1 and α2 chains. From left to right: the T cell receptor interacts with the MHC–peptide complex, generally binding with positions 4, 5, and 6 of the peptide. The lectin-like NK receptors generally bind in two sites: site 1 is off the N-terminal end of the α1 chain, avoiding the peptide, and site 2 (receptor with broken line) is underneath the α1–α2 domain in contact with α3 and β2m. The KIRs bind the top of the α1–α2 domain including the C-terminal end of the bound peptide. The LILRs bind the MHC α3 domain and β2-microglobulin.

Glossary

Adaptive evolution: evolutionary change driven by natural selection that increases frequencies of beneficial alleles and decreases frequencies of deleterious alleles.
Balancing selection: any form of natural selection that maintains genetic variation within populations.
Coevolution: the process in which two or more species (or genes) evolve nonindependently by exerting selection pressures on each other.
Disassortative mating: preferential mating between individuals with dissimilar trait values (e.g., MHC genotypes).
Gene genealogy: relationships between DNA sequences in a sample taken from a population; properties of gene genealogy depend on both demographic history of the population and action of selection.
Haplotype: a genomic segment that is usually inherited as a single unit, without recombination.
Hitchhiking: process in which a neutral or deleterious variant increases in frequency due to the linkage with a positively selected variant.
Introgression: transfer of gene(s) from one species into the gene pool of another species, mediated by hybridization and repeated backcrossing.
Negative selection: (i) evolution: natural selection removing deleterious alleles; (ii) immunology: deletion of T cells with TCR that bind too strongly self-peptides presented by MHC molecules.
Overdominance: fitness of heterozygote genotype exceeds that of each homozygote carrying its constituent alleles.
Peptide-binding domain (PBD): also peptide-binding region (PBR); protein domain(s) within the MHC molecule (α1 and α2 in MHC-I; α1 and β1 in MHC-II) that forms the peptide-binding cleft.
Peptide-binding sites (PBSs): amino acid residues within the PBD that directly contact the peptide bound by the MHC molecule.
Positive selection: (i) evolution: natural selection increasing frequencies of beneficial mutations; if novel protein variants are favored, the rate of nonsynonymous changes will be elevated; (2) immunology: retention of T cells with TCRs that interact properly with MHC–peptide complexes.

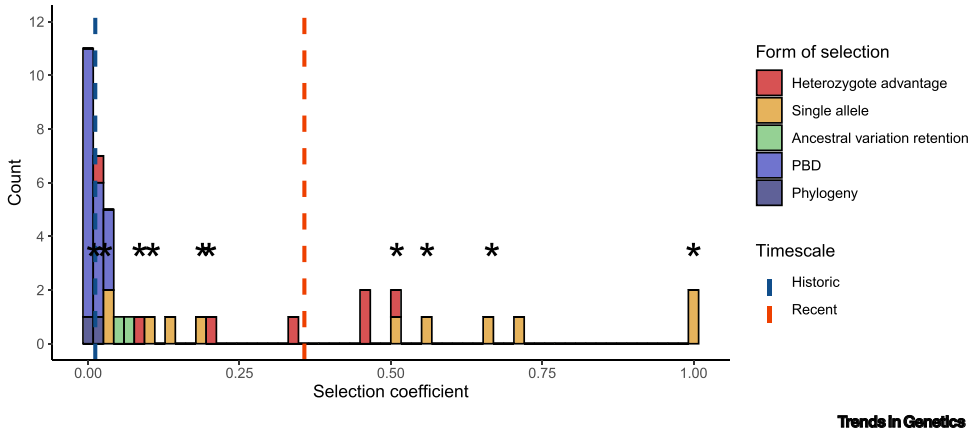

Figure 1. Distribution of Empirically Derived *MHC* Selection Coefficients from the Literature. Selection coefficients (i.e., differences in relative fitness between genotypes) were identified from 19 studies (see Table S1 in the supplemental information online). Broken vertical lines are mean values by relative timescale captured by the form of selection. Asterisks denote estimates for which selection could be ascribed to pathogens. Counts are the number of selection coefficients. Heterozygote advantage, single-allele effects, and ancestral variation retention are considered to capture more recent selection events, while PBD and phylogeny-based values are considered to capture more long-term historic selection (see Table S1 in the supplemental information online for more detail on how selection coefficients were estimated). Higher estimates in some recent selection events are likely to reflect the dynamic nature of host–parasite coevolution, resulting in bouts of strong selection acting on *MHC*.

Red Queen dynamics: a process in which organisms must constantly adapt to challenges imposed by adaptation occurring in other, coevolving organisms.
Selection coefficient: a difference in fitness between genotypes, typically measured as the difference between the fittest genotype and the genotype in question.
Supertype: a group of putatively functionally similar MHC protein variants, due to similarity of physicochemical properties of amino acids in positions key to specificity of antigen binding.
T cell receptor (TCR): receptor interacting with the MHC–peptide complex, coded by V, D, and J genes which in the process of somatic recombination produce a vast number of different variants expressed by an individual.
Trans-species polymorphism (TSP): presence of alleles in different species that are more similar to each other than some alleles within species; TSP results from retention of allelic lineages that were already distinct in the most recent common ancestor of the species; under neutral evolution TSP is transient and relatively-short lived, but it may persist for tens of millions of years under balancing selection.

[18,19] could alleviate this constraint. Recent years have brought significant progress in addressing these questions, which we review in the following sections. New studies provided clear evidence that several previously proposed evolutionary mechanisms indeed act on *MHC* in nature. In addition to addressing these long-standing enigmas, they also identified complexities of selection acting on *MHC* that have not been considered previously. These recent findings have also allowed formulation of new research questions. Here we review this recent progress and highlight outstanding and emerging questions.

How Parasites Select for *MHC* Polymorphism

The number of alleles segregating at *MHC* loci is hardly matched by any other gene, and it is thus natural that the maintenance of this polymorphism has been a focus of evolutionarily oriented MHC research. Because of the functions of MHC molecules in immune response, selection by pathogens has generally been assumed the main underlying force, and has indeed been reported in multiple studies [20,21] (Figure 1). Two mechanistic explanations have been considered ever since *MHC* gene discovery: heterozygote advantage (HA) and rare-allele advantage or negative frequency-dependent selection (NFDS) [22–24]. Another mechanism, based on fluctuating selection (FS) over time and/or space, was proposed later on and is conceptually related to HA [25] (Box 3).

HA can arise from some degree of dominance of resistance, which allows heterozygotes to respond to a wider range of pathogens or pathogen strains compared with homozygotes (Box 3), or from **overdominance**, whereby heterozygotes have intrinsically higher fitness than homozygotes [26]. HA has been extensively tested in many studies, and *MHC* heterozygosity was indeed sometimes reported to be associated with greater resistance to infection (Figure 1; reviewed in [20]). However, it may be difficult to distinguish selection favoring heterozygotes from selection favoring particular alleles, which, depending on their frequency, may be present mainly in homozygotes or heterozygotes [27,28]. HA also received some support from

Key Figure

Key Processes Shaping *MHC* Polymorphism in Populations and Within-Individual Antigen-Binding Range

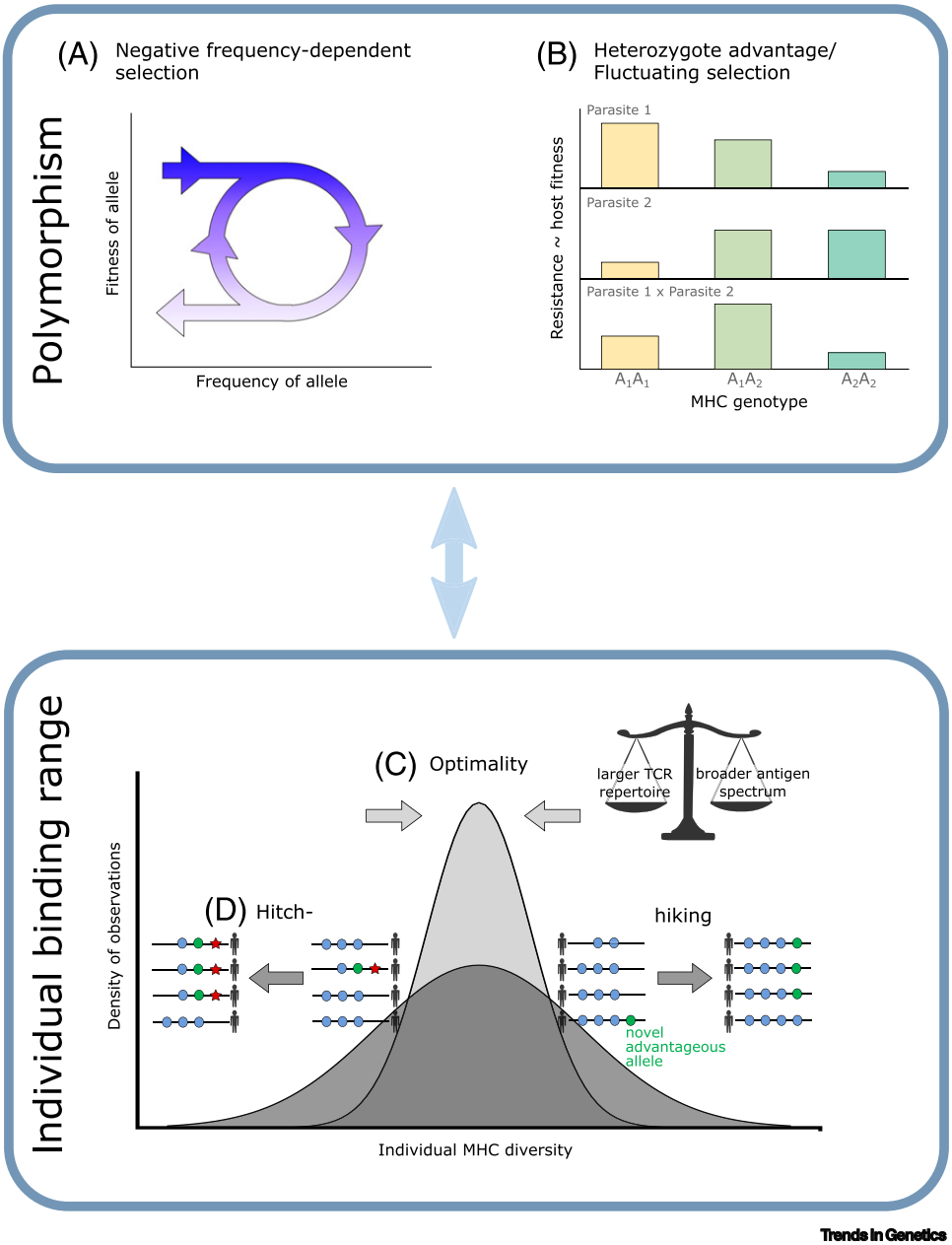

Figure 2. Upper panel: *MHC* polymorphism. (A) Fast adaptation of pathogens reduces fitness of common alleles, favoring rare *MHC* alleles, as well as functionally novel alleles (carrying nonsynonymous mutations, incoming arrow; outcoming arrows denote alleles lost due to selection or drift; adapted from [125]). (B) Exposure of host genotypes to multiple pathogens can lead to heterozygote advantage (HA) when exposure takes place in a single generation, or help maintain

(Figure legend continued at the bottom of the next page.)

Box 3. Mechanisms Proposed to Maintain *MHC* Polymorphism

Parasite-Driven Mechanisms (Not Mutually Exclusive)

Heterozygote advantage (HA) Because each MHC molecular variant is able to present only a limited repertoire of antigens to T cells, being heterozygote and thus expressing two different MHC proteins should increase the probability of presenting a given antigen and thus raising an adaptive immune response [23]. For HA to be able to maintain polymorphism, resistance to a single pathogen should be overdominant, or more plausibly, dominant with different alleles conferring resistance to different pathogen species or strains, resulting in fitness over multiple infections that is overdominant (see Figure 2 in main text) [105]. However, dominance of resistance appears not to be a universal feature of MHC [30]. Furthermore, the existing theory predicts that unless fitness contributions of different alleles to resistance are similar, HA alone can maintain much fewer alleles than observed in natural populations [106].

Negative frequency-dependent selection (NFDS) arises from the fact that pathogens will tend to adapt by evading presentation by the most common MHC types [22]. Simulations of host–parasite coevolution suggest that this mechanism is capable of maintaining high levels of *MHC* polymorphism [39].

Fluctuating selection (FS) arises when there is variation in the presence of pathogens over time. The process can maintain *MHC* polymorphism under restrictions shared with HA model (dominance, similar fitness contributions of alleles), plus balanced occurrence of pathogen species in time [25]. Selection on *MHC* alleles can also vary in space, for which there is some evidence from the field (see main text) but which has been little explored theoretically.

Other Mechanisms

Mate choice for dissimilar mates can in theory maintain *MHC* polymorphism even in the absence of selection from parasites [107,108]. It should be stressed, however, that the evolution of such preferences requires pre-existing *MHC* polymorphism [109]. Nevertheless, mate choice for advantageous and compatible *MHC* genotypes can substantially affect the speed of *MHC* evolution [110].

Sheltered load would accumulate if recessive deleterious mutations linked to *MHC* were hidden from selection due to high *MHC* heterozygosity [80], analogous to a mechanism earlier postulated for plant self-incompatibility genes [111]. Similar to mate choice, this mechanism requires pre-existing balancing selection, but once at work it could potentially help maintain polymorphism in periods when selection from parasites is weak [80].

experimental infection studies with inbred mice and with frogs [26,29], but a study on outbred mice did not find support for HA in resistance to infection [30]. Even if, on average, heterozygotes are fitter than homozygotes, existing theory suggests that, similar to classical overdominance models [31], HA resulting from dominant resistance can maintain only a limited number of alleles, calling into question a major role of HA in the maintenance of *MHC* polymorphism (Box 3). This constraint is somewhat alleviated by the divergent allele advantage (DAA) version of HA [32,33]. DAA assumes that *MHC* heterozygotes carrying alleles with more divergent binding properties should present a larger overall repertoire of antigens, an assumption supported by correlational studies [34–36] as well as through computational analyses [37,38].

polymorphism via fluctuating selection (FS) if exposure to different pathogens takes place in different generations. This is particularly likely if resistance is dominant and different alleles confer resistance to different pathogens, such that fitness (a product of resistance to pathogens encountered) can be highest in heterozygotes (within or across generations for HA and FS, respectively). Lower panel: Individual binding range – a composite of *MHC* gene number and allele-binding properties. (C) By extension of the HA mechanism, *MHC* duplication and divergence should increase the spectrum of antigens an individual can present, increasing the probability of immune response. However, as the number of self-peptides presented would also increase, costs may outweigh the benefits, for example, via processes that deal with autoimmunity, such as negative selection in thymus against self-reactive T cells. The same processes are likely to shape the evolution of the binding range of particular alleles. In either case, antigen-binding range should be optimized, narrowing the distribution of within-individual binding range. (D) However, beneficial alleles [subject to continuous turnover, upper panel (A)] can occur on haplotypes carrying suboptimal number of genes, widening the distribution of *MHC* diversity via hitchhiking (dark gray area) compared with purely optimizing selection (light gray area). In addition, deleterious mutations (red stars) linked to *MHC* genes can hitchhike with beneficial alleles (green circles). The blue double-headed arrow indicates that processes shaping population and individual diversity are inter-related (see the main text for details). Abbreviation: TCR, T cell receptor.

NFDS can in theory readily maintain observed levels of *MHC* polymorphism [39]. Furthermore, the dynamic nature of host–parasite **coevolution** underlying NFDS [40] is easy to reconcile with high selection coefficients sometimes reported to act on *MHC* in the short term (Figure 1). Yet, NFDS has been harder to demonstrate than HA. Associations of infection with *MHC* alleles, reported multiple times (Figure 1; reviewed in [20]), are consistent not only with NFDS, but also with FS, HA (if resistance alleles are rare, they occur almost exclusively in heterozygotes), and even with directional selection depleting *MHC* polymorphism. Relating parasite load to snapshots of current allele frequencies may not be very informative either, because host–parasite coevolution may superimpose time lags on allele frequency changes. Therefore, a currently rare *MHC* allele that was common in the recent past can still be susceptible to pathogens and conversely, pathogens might not have adapted to a currently common, beneficial *MHC* allele if it increased in frequency only recently. That is probably why frequency dependence may be observed for some, but not for other snapshots of populations within the same system [41,42]. This shortcoming of snapshot studies has recently been overcome by using a system of allopatric guppy populations in which introduction of *MHC* alleles to which pathogens have not had a chance to adapt for many generations indeed increases resistance to a parasite [43]. Also consistent with fast local adaptation, experimental evolution in the laboratory showed that viruses passaged for a dozen generations on a mouse strain of a given MHC **haplotype** evolved improved infection success on that haplotype, but not on alternative haplotypes to which they have not been exposed [44]. These findings are consistent with the proposed mechanism of NFDS whereby rare *MHC* alleles may 'regain' advantage after pathogens are forced to adapt to more common MHC alleles [45]. Still, a cycle where a common *MHC* allele is evaded by pathogens becomes rare and then regains resistance has not yet been demonstrated in full.

By contrast, there is some evidence that at least sometimes strong selection from pathogens may in fact reduce *MHC* polymorphism, as exemplified by the chimpanzee *MHC*-I A locus (see Box 2 for functions of Class I and II loci), which shows an order of magnitude lower diversity compared with the orthologous human leukocyte antigen (*HLA*)-A locus in humans, even though we would expect the reverse based on larger effective population size in chimpanzees [46]. This is likely an effect of simian immunodeficiency virus epidemics some 3 million years ago (mya), strongly favoring *MHC* alleles functionally related to those slowing AIDS progression in humans [47]. Conversely, a recent ancient DNA study showed that the same MHC-II *DRB*1*15:01 allele that confers susceptibility to the leprosy-causing *Mycobacterium leprae* in contemporary populations was already positively associated with leprosy infection in medieval Europe [48]. Still, the allele shows only a minor reduction in frequency and remains common in contemporary Europe, suggesting that selection against it was not effective, perhaps due to pleiotropic effects and resulting fitness trade-offs, such as its protective effect against type 1 diabetes [49]. Coupled with little evidence for adaptive evolution in *M. leprae* to evade presentation by MHC-II proteins, these data provide no support for NFDS, at least over the ~24 generations covered by this study. Overall, the role of NFDS in maintaining *MHC* polymorphism, while supported by recent evidence demonstrating preconditions necessary for it to work [15,43,44], remains to be more firmly established.

In comparison to HA and NFDS, the role of FS on the maintenance of *MHC* polymorphism has received relatively less attention. A number of studies reported associations between *MHC* variation and spatial differences in parasite communities, leading to increased *MHC* variation at the meta-population level (reviewed in [20,50]). Even less evidence is available for temporal variation, but a recent study reported temporal differences in pathogen species composition, where different *MHC* alleles were associated with different pathogens [35]. The advent of ancient DNA technology might help provide a better picture of the contribution of past temporal fluctuations to *MHC* polymorphism in present-day populations [51]. Theory shows that such temporal

fluctuations can maintain polymorphism under constraints similar to those applying to HA (Box 3). By contrast, spatial variation in selection pressures on *MHC* has not yet received a comprehensive theoretical treatment, even though models of host–pathogen coevolution show that geographic subdivision can have important implications for polymorphism in hosts and parasites [52,53]. Population structure may interact with NFDS acting on *MHC* when neighboring populations are 'out of phase'. An example of this is provided by a study on bank voles, where an allele conferring resistance to a nematode species in one population increased susceptibility to the same parasite in another population [54]. Given that population structure at the *MHC* is often pronounced [55–57], its role in maintaining *MHC* diversity clearly deserves more attention.

Darwinian Demons Do Not Exist in the *MHC* World Either

Darwinian demons, hypothetical organisms that are able to maximize all elements of their fitness simultaneously, do not exist because of trade-offs between fitness components [58]. What trade-offs prevent MHC molecules from inciting immune response to any pathogen? Interestingly, recent evidence suggests that the probability of binding at least one immunoreactive antigen by a given MHC molecule should not constrain immune response to large viruses [7]. Differences in protection conferred by different MHC molecules can still stem from differences in binding properties (e.g., binding affinity) [59], or, in the case of fast-evolving pathogens, the ability to target antigens conserved by functional constraints [9]. This might explain why the ability to raise an immune response sometimes cannot be equated with resistance [60]. Nevertheless, the level of resistance to pathogens may in part depend on the range of different antigens bound by a given MHC molecule, which in a recent analysis was shown to be positively associated with protective effect against HIV [61]. This raises the question of why MHC variants differ in the number of different peptides they can present by nearly two orders of magnitude, as indicated by recent discoveries in chicken and humans [18,19]. Should not promiscuous binders (generalists) replace fastidious binders (specialists)? A related question was raised much earlier in the MHC literature [62,63]: why the number of *MHC* loci in the genome is not larger, providing protection against any possible pathogen that the host can encounter? Such a multilocus genotype would constitute a Darwinian demon of immunogenetics.

In the context of the MHC, it has been proposed that expansion of the *MHC* gene family is limited by the associated risk of autoimmunity, as it was recently shown for *MHC* heterozygosity in humans [64]. To avoid self-aggression, T cells with a **T cell receptor (TCR)** binding self-peptides too strongly are deleted in the process of **negative selection**. The resulting 'holes' in the TCR repertoire would increase with an increased range of MHC-presented epitopes associated with expressing more *MHC* variation across duplicated loci [63]. Consequently, the number of expressed *MHC* loci should be optimized, rather than maximized by selection [62,65]. Indirect support for this T cell depletion hypothesis was provided by studies showing that individuals with intermediate number of allelic variants across duplicated *MHC*-II genes carry the fewest parasite species [54,66]. However, the TCR depletion explanation was challenged based on the argument that expressing a wide range of MHC molecules might also enhance positive selection [67], whereby T cells, prior to negative selection, are tested for their functionality against self-peptide–MHC complexes. Technological advances now allow predictions of the T cell depletion hypothesis to be tested directly, and a recent study in a rodent [68] demonstrated that the TCR repertoire was indeed negatively associated with the number of expressed *MHC*-I (but not *MHC*-II) alleles. Whether specialist and generalist *MHC* alleles are associated with similar trade-offs [18] remains to be seen. Computational work suggests that generalists might be favored in pathogen-rich environments [69], in line with indirect evidence that pathogen richness might shift the optimum toward a higher number of *MHC* genes [70,71]. Alternatively, it has been suggested that promiscuous binders are ancestral, with particular specialists favored by specific

selection pressures from pathogens with strong fitness effects [19]. Associations of MHC haplotypes or specialist/generalist alleles with important diseases might increase the observed ranges compared with those predicted under purely optimizing selection (Figure 2). Optimizing selection might further be complicated by sex-specific selection on individual MHC diversity reported in recent studies [68,72]. Clearly, more work in this area is needed to understand evolution of antigen-binding ranges at the individual level.

The Role of Genomic Architecture

Selection on particular MHC types may sometimes be constrained by their genomic context, in particular, within the MHC genomic region, as exemplified by long-term conservation of multigene haplotypes within the MHC-I region of zebrafish [73]. Presentation of antigens on the cell surface does not only depend on MHC-binding properties, but also depend on the efficiency of the molecular apparatus assembling MHC and antigens [2], the blocking of which is basis for numerous pathogen evasion mutations [74]. In rat and in several non-mammalian vertebrates, *MHC*-I genes are located in proximity to genes coding for transporters associated with antigen presentation (TAPs) and tapasins, which can be highly polymorphic and, as evidence from chicken and rat indicates, co-evolve their specificities to work optimally with *MHC*-I alleles present on the same haplotype [75]. In most mammals, these genes are located at some distance from MHC-I region and are not polymorphic [75]. More work is needed to resolve whether differences in genomic structure represent alternative solutions to constraints imposed on *MHC* evolution, at least in part, by their linkage with genes coding for MHC peptide loading complex [2].

Selection on polymorphic *MHC* genes can also be affected by their linkage with deleterious mutations. Indeed, the MHC region of the human genome is particularly strongly associated with genetic disease risk, including various common diseases [11]. Many of them are associated with classical *MHC* genes themselves (reviewed in [76]), but multiple associations with diseases also map to genes other than classical *MHC* [77] and to intergenic polymorphisms [78,79]. This could indicate an accumulation of deleterious mutations within the MHC region, raising the question of how these are maintained despite their fitness costs. In this context, it has been proposed that recessive deleterious variants can reach high frequency in the MHC region and even fix within particular *MHC* allelic lineages because of physical linkage and excessive heterozygosity in this region. Such deleterious variants would largely be hidden from selection through the high heterozygosity that prevents their phenotypic expression, forming a sheltered load [80] (Box 3). However, although the accumulation of deleterious mutations around classical *MHC* genes has been demonstrated in humans [81], those variants are not necessarily recessive. Indeed, explicit modeling of polymorphism around targets of **balancing selection** also showed an increased frequency for 'additive' deleterious mutations. A negative correlation was observed between the frequency of these deleterious variants and their distance from classical *MHC* genes, supporting the notion that the effect is due to **hitchhiking** through linkage with targets of balancing selection [81]. Whether such accumulation of deleterious mutations in certain haplotypes can affect dynamics of *MHC* polymorphism would be worth investigating.

Furthermore, the organization of *MHC* gene copies in multilocus haplotypes leads to additional complexity for *MHC* evolution, such as epistatic interactions between loci [82,83]. A recent simulation study aimed to explore the evolutionary dynamics that govern MHC haplotype evolution in human populations, fitting their models to MHC haplotype data from 6.59 million individuals of a bone marrow donor program [84]. Their best-fitting models combined, unexpectedly, positive frequency-dependent selection, rapid fitness decline of haplotypes, and a very high haplotype recruitment rate (a composite measure of mutation, recombination, and gene flow). These simulations have been criticized for making a number of biologically implausible assumptions and for

not addressing outstanding features of *MHC* evolution, such as TSP and high nonsynonymous variation [85]. Overall, while the significance of this study thus remains unclear, the role of the genomic context in shaping evolution of polymorphic *MHC* genes and haplotypes is certainly an important field for future research.

Reconciling Landmark Features of *MHC*

Can the landmark *MHC* features – signatures of both balancing (high polymorphism and long-term maintenance of allelic linages) and positive (excess of nonsynonymous variation) selection – be explained by the set of mechanisms described earlier, or should the current paradigm be extended? Takahata and Nei [24] concluded from a simulation model that HA, and some versions of NFDS, can both cause positive selection, maintain polymorphism of *MHC* genes, and produce long **gene genealogies**. In the case of NFDS, the same mechanism of pathogens adapting to common *MHC* alleles would favor the rise of new functional MHC variants, generating the signal of positive selection, and prevent loss of alleles that become rare, as these would regain resistance due to trade-offs associated with pathogen adaptation [44] (Figure 2). 'Good genes' sexual selection might enhance selection from pathogens (Box 4), possibly accelerating the rate of nonsynonymous substitutions observed in species with higher inferred sexual selection intensity [86]. In the case of HA, although new alleles will be found initially in heterozygotes, their selective advantage will be very weak in already polymorphic populations, characterized by high heterozygosity [87]. If species inherit much polymorphism (and thus high expected heterozygosity) from ancestor species, as suggested by wide occurrence of TSP [17], contribution of HA to positive selection may be negligible, unless FS, NFDS, or nonequilibrium demography skew allele frequencies. As for FS, it is not clear how it contributes to positive selection, and we know of no theoretical treatment of this topic.

The role of selection from pathogens in maintaining TSP is more controversial. It has been argued that positive selection on the *MHC* does not necessarily reflect balancing selection, but may be a

Box 4. MHC-Based Mating Preferences, Maternal–Fetal Interactions, and *MHC* Polymorphism

While the potential of mating preferences for MHC-dissimilar mates to maintain *MHC* polymorphism has long been discussed [105,112] (Box 3), recent meta-analyses question the paradigm of MHC-**disassortative mating** [113,114]. A systematic collection of the literature across vertebrates, including humans, corroborated the prediction that females choose more MHC diverse mates (i.e., more heterozygous or having more classical *MHC* loci), while support for preferences for MHC-dissimilar mates appears to depend on particular taxa and number of genes investigated [113], and was not found for human and non-human primates [114]. In addition to mate choice, postcopulatory selection has also been proposed as a process that can promote *MHC* diversity, if mate preferences fail to do the job. This can take shape of sperm–egg/maternal interactions [45,115,116] or maternal–fetal interactions. As for the latter, evidence also suggests that MHC similarity between mates is linked to recurrent spontaneous abortion, low birth weight, and other pregnancy-associated problems [117–120]. However, the most polymorphic *MHC*-I and *MHC*-II genes are not expressed by the fertilized embryo [121], except for slightly less polymorphic HLA-C, and there is conflicting evidence whether classical MHC molecules are expressed on the unfertilized egg and sperm at all [122]. Thus, any effect of MHC on pregnancy success may not be a manifestation of evolutionary optimization of classical MHC diversity in progeny, but rather arise from the need to stimulate maternal immunity that allows successful implantation and improves blood supply to the developing fetus [121]. Indeed, across eutherian (so-called placental) mammals, low polymorphism *MHC*-I expression has been observed on the embryo cell wall (trophoblast), and human, monkey, and mouse studies support that allogenic parental MHC-I combinations may influence likelihood of implantation, fetal growth, and successful pregnancy, but not necessarily levels of *MHC* heterozygous and homozygous offspring [118,120,123]. Thus, while MHC-dependent sexual selection has been experimentally shown in several species, precopulatory and postcopulatory mechanisms do not generally seem to favor maximizing *MHC* diversity, so that their role in maintaining *MHC* polymorphism, relative to pathogen-mediated selection, remains debated. The effects of preferences for MHC-diverse mates which emerged significant in meta-analyses are consistent with mating advantage of individuals in high phenotypic condition [110], given that heterozygotes may often show improved resistance to pathogens. The effect of such preferences remain to be explored theoretically, but intuitively they could reinforce HA (or selection on particular *MHC* alleles) [124]. Furthermore, preference for optimally dissimilar mates as an explanation for the conflicting results from meta-analyses remains to be more fully explored [110,113].

result of a series of selective sweeps, which would erase TSP [88]. In contrast to 'minority advantage' modeled by Takahata and Nei [24], which assumed entirely deterministic NFDS, host–pathogen coevolution may be more dynamic and unpredictable, and loss of *MHC* variation can occur [89], as exemplified by the chimpanzee MHC-I example [46] discussed earlier. Simulation studies have shown that while **Red Queen dynamics** strongly favor novel MHC variants, they can make gene genealogies longer or shorter compared with drift, depending on parameter combinations which are mostly unknown for natural populations [87]. The TSP phenomenon, that is, the maintenance of multiple ancient allelic lineages, thus remains a puzzling observation. Convergent evolution of *MHC* alleles due to shared parasites between species could yield patterns resembling TSP, but this explanation is not supported by recent research [90,91]. The presence of sheltered load within MHC (Box 3) could in principle contribute to TSP [80], but as discussed earlier, evidence for this hypothesis is also lacking. Long-term persistence of allelic lineages can also be facilitated by geographic subdivision and/or *MHC* **introgression** between species. In geographically structured models of host–pathogen coevolution (which do not consider *MHC* explicitly), polymorphism is maintained over longer periods than in unstructured models, probably due to asynchrony of coevolutionary interactions [52,53]. Introduction of *MHC* alleles from other species via adaptive introgression [4,55,92] could also lead to patterns consistent with TSP. Instead of being an indirect result of other dynamics, long-term coexistence of divergent allelic lineages could also be selected for directly, for example, due to their functional distinctness [33]. Lighten *et al.* [88] proposed that TSP may arise due to long-term maintenance of functional MHC groups (**supertypes**), with allelic replacement occurring mostly within these groups, but empirical support and theoretical foundations for this proposition have been questioned [93]. Still, it seems appealing that DAA could help prevent extinction of most divergent lineages [33,37], a possibility that requires an in-depth theoretical exploration. Overall, current evidence suggests that TSP can arise directly from host–pathogen coevolution under panmixia, but population subdivision and/or interspecific introgression may contribute to the long-term maintenance of *MHC* allelic lineages. The relative roles of these mechanisms deserve further investigation.

Concluding Remarks

Well-understood function and extreme polymorphism of *MHC* genes have made them a leading model for investigation of positive and balancing selection [94], but at the same time the complex nature of selective pressures acting on these genes has hampered full understanding of underlying mechanisms. Two classical mechanisms of balancing selection, HA and NFDS, have now been demonstrated to be able to operate on *MHC* in principle, and likely both contribute to *MHC* polymorphism. Existing theory questions a major role of HA in explaining the existing levels of polymorphism and nonsynonymous variation all by itself. Yet, the leading role of NFDS remains to be more firmly established given that some long-term data appear to contradict it [46,48]. The role of temporal and spatial fluctuations in pathogen composition for selection of *MHC* polymorphism also deserves more theoretical and empirical attention. While mate choice for dissimilar mates appears to lack the generality necessary to significantly impact evolution of *MHC* polymorphism, the modifying role of preferences based on partners' *MHC* diversity or complementarity deserves to be more thoroughly explored (Box 4). Selection shaping *MHC* polymorphism can also be modified by evolution of binding promiscuity of MHC molecules or *MHC* gene number, both affecting antigen-binding ranges of individuals. Conversely, selection favoring particular *MHC* alleles may result in indirect selection for haplotypes with suboptimal number of genes (or alleles with suboptimal binding range), which can explain between-individual variance in *MHC* gene number in natural populations (Figure 2). Additional selection pressures on the *MHC* that are likely to affect evolution of *MHC* polymorphism include genomic context, accumulation of deleterious mutations in MHC haplotypes, interactions with NK receptors (Box 2), or sex-specific selection (see Outstanding Questions). Future theory could also benefit from incorporating

Outstanding Questions

How does pathogen evolution rate and peptidome size modulate selection on the *MHC*? Do selective sweeps, such as those inferred for *MHC*-I in chimpanzee and characteristic of pathogens particularly effective in immune evasion (such as HIV), strongly favor MHC molecules capable of presenting highly conserved pathogen peptides?

Are trade-offs between disease protection conferred by *MHC* alleles and other fitness components, for example, increased risk of autoimmunity (see the leprosy case discussed in the main text) common and if so, how do such trade-offs interact with other processes shaping *MHC* polymorphism?

Are fastidious and promiscuous *MHC* alleles co-maintained because of such evolutionary trade-offs?

How important is the genomic architecture of the MHC region and interactions with other genes, encoded both within (e.g., *TAP*s) and outside (e.g., killer inhibitor receptors; see Box 2) this region in shaping *MHC* polymorphism?

Can accumulation of deleterious mutations (recessive or not) within the MHC region determine the evolutionary dynamics of MHC haplotypes and shape of *MHC* gene genealogies? Do deleterious alleles accumulate in MHC of species other than human? Is the accumulation affected by the landscape of recombination within the MHC region?

Can divergent allele advantage prevent the extinction of divergent allele lineages and lead to trans-species polymorphism? Does divergent allele advantage favor introgression at MHC?

How important is geographic subdivision and interspecific introgression for the long-term maintenance of *MHC* variation?

Do differences between species in the number of *MHC* genes in the genome translate into differences in the number of classical *MHC* genes expressed on the cells? If so, what selective forces shape this variation?

How does sex-specific selection on the *MHC* affect *MHC* diversity within individuals and populations?

spatial structure in models of balancing selection on *MHC* polymorphism. Introgression of *MHC* alleles from different populations, and even species, appears to be an important and perhaps widespread process with likely consequences for the level of polymorphism and genealogies of *MHC* genes. Overall, recent years have brought considerable progress in our understanding of processes shaping polymorphism of the *MHC* genes. Major mechanisms postulated by classical theory have been confirmed to operate in nature. However, recent research has also shown the need to extend the existing theory beyond classical models of balancing selection, to include additional selective pressures listed earlier. Further empirical work is needed to assess the generality and importance of these pressures.

Acknowledgments

We thank Helena Westerdahl, Manfred Milinski and two anonymous reviewers for their comments on previous version of our manuscript. W.B. was supported by the Polish National Science Centre grant 2018/31/B/NZ8/01210. J.K. was supported by a Wellcome Trust Investigator Award (110106/Z/15/Z). T.L.L. was supported by the German Research Foundation (DFG, grant LE 2593/3-1). J.W. was supported by the German Research Foundation (DFG) as part of the SFB TRR 212 (NC^3) – Project numbers 316099922 and 396780709.

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tig.2020.01.008.

References

1. Flajnik, M.F. (2018) A cold-blooded view of adaptive immunity. *Nat. Rev. Immun.* 18, 438–453
2. Kaufman, J. (2018) Unfinished business: evolution of the MHC and the adaptive immune system of jawed vertebrates. *Annu. Rev. Immun.* 36, 383–409
3. Biedrzycka, A. *et al.* (2018) Blood parasites shape extreme major histocompatibility complex diversity in a migratory passerine. *Mol. Ecol.* 27, 2594–2603
4. Dudek, K. *et al.* (2019) Massive introgression of MHC genes in newt hybrid zones. *Mol. Ecol.* 28, 4798–4810
5. Robinson, J. *et al.* (2017) Distinguishing functional polymorphism from random variation in the sequences of >10,000 HLA-A, -B and -C alleles. *PLoS Genet.* 13, e1006862
6. Reche, P.A. and Reinherz, E.L. (2003) Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. *J. Mol. Biol.* 331, 623–641
7. Croft, N.P. *et al.* (2019) Most viral peptides displayed by class I MHC on infected cells are immunogenic. *Proc. Natl. Acad. Sci. U. S. A.* 116, 3112–3117
8. Falk, K. *et al.* (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* 351, 290–296
9. Goulder, P.J. and Watkins, D.I. (2008) Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat. Rev. Immun.* 8, 619–630
10. Schmidt, J. *et al.* (2011) Host genetics in immune-mediated hepatitis C virus clearance. *Biomark. Med* 5, 155–169
11. Trowsdale, J. and Knight, J.C. (2013) Major histocompatibility complex genomics and human disease. *Annu. Rev. Genomics Hum. Genet.* 14, 301–323
12. Hughes, A.L. *et al.* (1990) Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. *Mol. Biol. Evol.* 7, 515–524
13. Biedrzycka, A. and Radwan, J. (2008) Population fragmentation and major histocompatibility complex variation in the spotted suslik, *Spermophilus suslicus*. *Mol. Ecol.* 17, 4801–4811
14. Schwensow, N. *et al.* (2007) Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity* 99, 265–277
15. Eizaguirre, C. *et al.* (2012) Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat. Comm.* 3, 621
16. Klein, J. (1987) Origin of major histocompatibility complex polymorphism – the transspecies hypothesis. *Hum. Immun.* 19, 155–162
17. Klein, J. *et al.* (2007) MHC, TSP, and the origin of species: from immunogenetics to evolutionary genetics. *Annu. Rev. Genet.* 41, 281–304
18. Chappell, P.E. *et al.* (2015) Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. *eLife* 4, e05345
19. Kaufman, J. (2018) Generalists and specialists: a new view of how MHC class I molecules fight infectious pathogens. *Trends Immun.* 39, 367–379
20. Spurgin, L.G. and Richardson, D.S. (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. Roy. Soc. B* 277, 979–988
21. Piertney, S.B. and Oliver, M.K. (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity* 96, 7–21
22. Bodmer, W. (1972) Evolutionary significance of the HL-A system. *Nature* 237, 139–145
23. Doherty, P.C. and Zinkernagel, R.M. (1975) Enhanced immunological surveillance in mice heterozygous at H-2 gene complex. *Nature* 256, 50–52
24. Takahata, N. and Nei, M. (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* 124, 967–978
25. Hedrick, P.W. (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution* 56, 1902–1908
26. Penn, D.J. *et al.* (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11260–11264
27. Arora, J. *et al.* (2019) HLA heterozygote advantage against HIV-1 is driven by quantitative and qualitative differences in HLA allele-specific peptide presentation. *Mol. Biol. Evol.* Published online October 25, 2019. https://doi.org/10.1093/molbev/msz249
28. Kloch, A. *et al.* (2013) MHC influences infection with parasites and winter survival in the root vole *Microtus oeconomus*. *Evol. Ecol.* 27, 635–653
29. Savage, A.E. and Zamudio, K.R. (2011) MHC genotypes associate with resistance to a frog-killing fungus. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16705–16710
30. Ilmonen, P. *et al.* (2007) Major histocompatibility complex heterozygosity reduces fitness in experimentally infected mice. *Genetics* 176, 2501–2508

31. Lewontin, R. *et al.* (1978) Heterosis as an explanation for large amounts of genic polymorphism. *Genetics* 88, 149–169
32. Stefan, T. *et al.* (2019) Divergent allele advantage provides a quantitative model for maintaining alleles with a wide range of intrinsic merits. *Genetics* 212, 553–564
33. Wakeland, E.K. *et al.* (1990) Ancestral polymorphisms of MHC class-II genes – divergent allele advantage. *Immun. Res.* 9, 115–122
34. Lenz, T.L. *et al.* (2013) Divergent allele advantage at MHC-DRB through direct and maternal genotypic effects and its consequences for allele pool composition and mating. *Proc. Roy. Soc. B* 280, 20130714
35. Osborne, M.J. *et al.* (2017) Spatio-temporal variation in parasite communities maintains diversity at the major histocompatibility complex class IIβ in the endangered Rio Grande silvery minnow. *Mol. Ecol.* 26, 471–489
36. Schwensow, N. *et al.* (2008) Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proc. Roy. Soc. B* 275, 555–564
37. Lenz, T.L. (2011) Computational prediction of MHC II-antigen binding supports divergent allele advantage and explains trans-species polymorphism. *Evolution* 65, 2380–2390
38. Pierini, F. and Lenz, T.L. (2018) Divergent allele advantage at human MHC genes: signatures of past and ongoing selection. *Mol. Biol. Evol.* 35, 2145–2158
39. Borghans, J.A.M. *et al.* (2004) MHC polymorphism under host-pathogen coevolution. *Immunogenetics* 55, 732–739
40. Woolhouse, M.E.J. *et al.* (2002) Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat. Genet.* 32, 569–577
41. Bolnick, D.I. and Stutz, W.E. (2017) Frequency dependence limits divergent evolution by favouring rare immigrants over residents. *Nature* 546, 285–288
42. Stutz, W.E. and Bolnick, D.I. (2017) Natural selection on MHC IIβ in parapatric lake and stream stickleback: balancing, divergent, both or neither? *Mol. Ecol.* 26, 4772–4786
43. Phillips, K.P. *et al.* (2018) Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1552–1557
44. Kubinak, J.L. *et al.* (2012) Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3422–3427
45. Lenz, T.L. *et al.* (2018) Cryptic haplotype-specific gamete selection yields offspring with optimal MHC immune genes. *Evolution* 72, 2478–2490
46. de Groot, N.G. *et al.* (2002) Evidence for an ancient selective sweep in the MHC class I gene repertoire of chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11748–11753
47. de Groot, N.G. *et al.* (2010) AIDS-protective HLA-B*27/B*57 and chimpanzee MHC class I molecules target analogous conserved areas of HIV-1/SIVcpz. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15175–15180
48. Krause-Kyora, B. *et al.* (2018) Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat. Comm.* 9, 1569
49. Hu, X. *et al.* (2015) Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat. Genet.* 47, 898–905
50. Eizaguirre, C. and Lenz, T.L. (2010) Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *J. Fish Biol.* 77, 2023–2047
51. Calvignac-Spencer, S. and Lenz, T.L. (2017) The One Past Health workshop: connecting ancient DNA and zoonosis research. *Bioessays* 39, 1700075
52. Thompson, J.N. (2005) *The Geographic Mosaic of Coevolution*, University of Chicago Press
53. Lion, S. and Gandon, S. (2015) Evolution of spatially structured host–parasite interactions. *J. Evol. Biol.* 28, 10–28
54. Kloch, A. *et al.* (2010) Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole *Myodes glareolus*. *Mol. Ecol.* 19, 255–265
55. Fijarczyk, A. *et al.* (2018) Balancing selection and introgression of newt immune-response genes. *Proc. Roy. Soc. B* 285, 20180819
56. Herdegen, M. *et al.* (2014) Selective pressures on MHC class II genes in the guppy (*Poecilia reticulata*) as inferred by hierarchical analysis of population structure. *J. Evol. Biol.* 27, 2347–2359
57. Sagonas, K. *et al.* (2019) Selection, drift, and introgression shape MHC polymorphism in lizards. *Heredity* 122, 468–484
58. Law, R. (1979) Optimal life histories under age-specific predation. *Am. Nat.* 114, 399–417
59. Moutaftsi, M. *et al.* (2009) Correlates of protection efficacy induced by vaccinia virus-specific $CD8^+$ T-cell epitopes in the murine intranasal challenge model. *Eur. J. Immunol.* 39, 717–722
60. Smith, A.L. *et al.* (2002) Antigenic diversity in *Eimeria maxima* and the influence of host genetics and immunization schedule on cross-protective immunity. *Infect. Immun.* 70, 2472
61. Arora, J. *et al.* (2019) HIV peptidome-wide association study reveals patient-specific epitope repertoires associated with HIV control. *Proc. Natl. Acad. Sci. U. S. A.* 116, 944–949
62. Nowak, M.A. *et al.* (1992) The optimal number of major histocompatibility complex molecules in an individual. *Proc. Natl. Acad. Sci. U. S. A.* 89, 10896–10899
63. Vidović, D. and Matzinger, P. (1988) Unresponsiveness to a foreign antigen can be caused by self-tolerance. *Nature* 336, 222–225
64. Lenz, T.L. *et al.* (2015) Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. *Nat. Genet.* 47, 1085–1090
65. Woelfing, B. *et al.* (2009) Does intra-individual major histocompatibility complex diversity keep a golden mean? *Phil. Trans. Roy. Soc. London B* 364, 117–128
66. Wegner, K.M. *et al.* (2003) Parasite selection for immunogenetic optimality. *Science* 301, 1343–1343
67. Borghans, J.A.M. *et al.* (2003) Thymic selection does not limit the individual MHC diversity. *Eur. J. Immunol.* 33, 3353–3358
68. Migalska, M. *et al.* (2019) Major histocompatibility complex class I diversity limits the repertoire of T cell receptors. *Proc. Natl. Acad. Sci. U. S. A.* 116, 5021–5026
69. Manczinger, M. *et al.* (2019) Pathogen diversity drives the evolution of generalist MHC-II alleles in human populations. *PLoS Biol.* 17, e3000131
70. Minias, P. *et al.* (2018) Evolution of copy number at the MHC varies across the avian tree of life. *Genome Biol. Evol.* 11, 17–28
71. O'Connor, E.A. *et al.* (2018) The evolution of immunity in relation to colonization and migration. *Nat. Ecol. Evol.* 2, 841–849
72. Roved, J. *et al.* (2018) Evidence for sexual conflict over major histocompatibility complex diversity in a wild songbird. *Proc. Roy. Soc. B* 285, 20180841
73. McConnell, S.C. *et al.* (2016) Alternative haplotypes of antigen processing genes in zebrafish diverged early in vertebrate evolution. *Proc. Natl. Acad. Sci. U. S. A.* 113, E5014–E5023
74. Hansen, T.H. and Bouvier, M. (2009) MHC class I antigen presentation: learning from viral evasion strategies. *Nat. Rev. Immun.* 9, 503–513
75. Kaufman, J. (2013) Antigen processing and presentation: evolution from a bird's eye view. *Mol. Immun.* 55, 159–161
76. Dendrou, C.A. *et al.* (2018) HLA variation and disease. *Nat. Rev. Immun.* 18, 325–339
77. Feder, J. *et al.* (1996) A novel MHC class I–like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.* 13, 399–408
78. Sekar, A. *et al.* (2016) Schizophrenia risk from complex variation of complement component 4. *Nature* 530, 177–183
79. Raj, P. *et al.* (2016) Regulatory polymorphisms modulate the expression of HLA class II molecules and promote autoimmunity. *Elife* 5, e12089
80. van Oosterhout, C. (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proc. Roy. Soc. B* 276, 657–665
81. Lenz, T.L. *et al.* (2016) Excess of deleterious mutations around HLA genes reveals evolutionary cost of balancing selection. *Mol. Biol. Evol.* 33, 2555–2564
82. Navarro, A. and Barton, N.H. (2002) The effects of multilocus balancing selection on neutral variability. *Genetics* 161, 849–863

83. Penman, B.S. *et al.* (2013) Pathogen selection drives nonoverlapping associations between HLA loci. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19645–19650
84. Lobkovsky, A.E. *et al.* (2019) Multiplicative fitness, rapid haplotype discovery, and fitness decay explain evolution of human MHC. *Proc. Natl. Acad. Sci. U. S. A.* 116, 14098–14104
85. Hedrick, P.W. and Klitz, W. (2019) Evolution of the human MHC: new haplotype frequency analysis is not informative. *Proc. Natl. Acad. Sci. U. S. A.* 116, 23386–23387
86. Winternitz, J.C. *et al.* (2013) Sexual selection explains more functional variation in the mammalian major histocompatibility complex than parasitism. *Proc. Roy. Soc. B* 280, 20131605
87. Ejsmond, M.J. and Radwan, J. (2015) Red Queen processes drive positive selection on major histocompatibility complex (MHC) genes. *PLoS Comput. Biol.* 11, e1004627
88. Lighten, J. *et al.* (2017) Evolutionary genetics of immunological supertypes reveals two faces of the Red Queen. *Nat. Comm.* 8, 1294
89. Ejsmond, M.J. and Radwan, J. (2011) MHC diversity in bottlenecked populations: a simulation model. *Cons. Gen.* 12, 129–137
90. Lenz, T.L. *et al.* (2013) Evaluating patterns of convergent evolution and trans-species polymorphism at MHC immunogenes in two sympatric stickleback species. *Evolution* 67, 2400–2412
91. Kaesler, E. *et al.* (2017) Shared evolutionary origin of major histocompatibility complex polymorphism in sympatric lemurs. *Mol. Ecol.* 26, 5629–5645
92. Abi-Rached, L. *et al.* (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science* 334, 89–94
93. Ejsmond, M.J. *et al.* (2018) The role of MHC supertypes in promoting trans-species polymorphism remains an open question. *Nat. Comm.* 9, 4362
94. Garrigan, D. and Hedrick, P.W. (2003) Perspective: Detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* 57, 1707–1722
95. Klein, J. (1986) *The Natural History of the Major Histocompatibility Complex*, Wiley & Sons
96. Blum, J.S. *et al.* (2013) Pathways of antigen processing. *Annu. Rev. Immun.* 31, 443–473
97. Starr, T.K. *et al.* (2003) Positive and negative selection of T cells. *Annu. Rev. Immun.* 21, 139–176
98. Adams, E.J. and Luoma, A.M. (2013) The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I–like molecules. *Annu. Rev. Immun.* 31, 529–561
99. D'Souza, M.P. *et al.* (2019) Casting a wider net: immunosurveillance by nonclassical MHC molecules. *PLoS Pathog.* 15, e1007567
100. Rossjohn, J. *et al.* (2015) T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immun.* 33, 169–200
101. Mariuzza, R. and Li, Y. (2014) Structural basis for recognition of cellular and viral ligands by NK cell receptors. *Front. Immunol.* 5, 123
102. Hirayasu, K. and Arase, H. (2015) Functional and genetic diversity of leukocyte immunoglobulin-like receptor and implication for disease associations. *J. Human Genet.* 60, 703–708
103. Parham, P. and Moffett, A. (2013) Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat. Rev. Immun.* 13, 133–144
104. Hammond, J.A. *et al.* (2009) Evolution and survival of marine carnivores did not require a diversity of killer cell Ig-like receptors or Ly49 NK cell receptors. *J. Immunol.* 182, 3618–3627
105. Apanius, V. *et al.* (1997) The nature of selection on the major histocompatibility complex. *Crit. Rev. Immun.* 17, 179–224
106. De Boer, R.J. *et al.* (2004) Heterozygote advantage fails to explain the high degree of polymorphism of the MHC. *Immunogenetics* 55, 725–731
107. Ejsmond, M.J. *et al.* (2014) Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proc. Roy. Soc. B* 281, 20141662
108. Hedrick, P.W. (1992) Female choice and variation in the major histocompatibility complex. *Genetics* 132, 575–581
109. Potts, W.K. and Wakeland, E.K. (1993) Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends Genet.* 9, 408–412
110. Milinski, M. (2006) The major histocompatibility complex, sexual selection, and mate choice. *Annu. Rev. Ecol. Evol. Syst.* 37, 159–186
111. Uyenoyama, M.K. (2003) Genealogy-dependent variation in viability among self-incompatibility genotypes. *Theor. Pop. Biol.* 63, 281–293
112. Penn, D.J. (2002) The scent of genetic compatibility: sexual selection and the major histocompatibility complex. *Ethology* 108, 1–21
113. Kamiya, T. *et al.* (2014) A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Mol. Ecol.* 23, 5151–5163
114. Winternitz, J. *et al.* (2017) Patterns of MHC-dependent mate selection in humans and nonhuman primates: a meta-analysis. *Mol. Ecol.* 26, 668–688
115. Wedekind, C. *et al.* (1996) Non-random fertilization in mice correlates with the MHC and something else. *Heredity* 77, 400–409
116. Yeates, S.E. *et al.* (2009) Atlantic salmon eggs favour sperm in competition that have similar major histocompatibility alleles. *Proc. Roy. Soc. B* 276, 559–566
117. Craenmehr, M.H.C. *et al.* (2019) Reciprocal HLA-DR allogenicity between mother and child affects pregnancy outcome parameters. *J. Repr. Immun.* 133, 15–17
118. Madeja, Z. *et al.* (2011) Paternal MHC expression on mouse trophoblast affects uterine vascularization and fetal growth. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4012–4017
119. Meuleman, T. *et al.* (2015) HLA associations and HLA sharing in recurrent miscarriage: a systematic review and meta-analysis. *Hum. Immun.* 76, 362–373
120. Ober, C. *et al.* (1998) Human leukocyte antigen matching and fetal loss: results of a 10 year prospective study. *Hum. Reprod.* 13, 33–38
121. Moffett, A. and Loke, C. (2006) Immunology of placentation in eutherian mammals. *Nat. Rev. Immun.* 6, 584–594
122. Fernandez, N. *et al.* (1999) A critical review of the role of the major histocompatibility complex in fertilization, preimplantation development and feto-maternal interactions. *Hum. Reprod. Update* 5, 234–248
123. Sterck, E. *et al.* (2017) No postcopulatory selection against MHC-homozygous offspring: evidence from a pedigreed captive rhesus macaque colony. *Mol. Ecol.* 26, 3785–3793
124. Eizaguirre, C. *et al.* (2009) MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Mol. Ecol.* 18, 3316–3329
125. Lenz, T.L. (2018) Adaptive value of novel MHC immune gene variants. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1414–1416
126. Miretti, M.M. *et al.* (2005) A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am. J. Hum. Genet.* 76, 634–646
127. De Bakker, P.I. *et al.* (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* 38, 1166