Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Trends in Genetics

CellPress

Opinion

Self-Peptidome Variation Shapes Individual Immune Responses

Pierre Pontarotti,[1,2,4,*] Laurent Abi-Rached,[1,2] Jung-Hua Yeh,[3] and Julien Paganini[4]

The relationship between human genetic variation and disease has not been fully elucidated. According to the present view on infectious diseases pathogen resistance is linked to human leukocyte antigen (HLA) class I/II variants and their individual capacity to present pathogen-derived peptides. Yet, T cell education in the thymus occurs through negative and positive selection, and both processes are controlled by a combination of HLA class I/II variants and peptides from the self. Therefore, the capacity of given HLA class I/II variants to bind pathogen-derived peptides is only one part of the selective process to generate effective immune responses. We thus propose that peptidome variation contributes to shaping T cell receptor (TCR) repertoires and hence individual immune responses, and that this variation represents inherent modulator epitopes.

Highlights

TCR repertoires emerge in the thymus in each individual as T cells undergo positive and negative selection.

T cell education is controlled by the combination of HLA class I/II molecules and their peptide pools (peptidome).

HLA class I/II molecules are highly plastic in human populations but the peptidome is also a source of variation. Hence combined diversity of HLA class I/II molecules and of self-peptides shapes individual immune responses.

Self-peptide variants that affect T cell repertoires represent inherent modulator epitopes.

Introduction: Disease Resistance/Susceptibility Is Linked to Inherent Modulator Epitopes

Despite the rapidly increasing availability of genetic data, the impact of genetic variation on disease resistance/susceptibility is not fully elucidated [1]. In humans, resistance/susceptibility to pathogens was linked to **human leukocyte antigen** (HLA; see Glossary) class I and II molecules because of their capacity to present pathogen-derived peptides to induce immune responses. Therefore, a major research focus was to study the genetic diversity of HLA class I/II molecules and the individual capacity of the variants to bind pathogen-derived peptides. This research supported the view that HLA class I and II genetic diversity impacts disease resistance/susceptibility and that this diversity evolved to meet the challenges of the pathogens encountered in different geographical areas [2,3].

We propose here that this conclusion is incomplete and that **polymorphism** in the peptides derived from the self also impacts disease resistance/susceptibility, playing a role as **inherent modulator epitopes** (IMEs). T cell education in the thymus is done through positive and negative selection. First, a positive selection of T cells capable of antigen recognition and then **negative selection** to eliminate T cells displaying high affinity for the combination of HLA class I/II plus self-peptide (self-antigen tolerance) [4]. While the shaping of **T cell receptor** (TCR) repertoires is a complex process involving different mechanisms, some of them still not well described or understood, it is accepted that these two selection steps rely, in part, on peptide pools derived from the self (**peptidome**) [5]. Peptide polymorphisms can thus impact TCR repertoires and modulate the capacity of these repertoires to recognize pathogen-derived peptides. This would also explain differences in disease resistance/susceptibility for individuals that display the same sets of HLA class I/II variants.

The first part of this opinion article corresponds to a brief state-of-the-art regarding HLA class I/II and TCR structural biology, particularly in the context of thymic selection. We then discuss how

[1]Aix Marseille Université, IRD, APHM, MEPHI, IHU Méditerranée Infection, 19–21 Boulevard Jean Moulin, 13005 Marseille, France
[2]SNC5039 CNRS, 19–21 Boulevard Jean Moulin, 13005 Marseille, France
[3]Prokarium Ltd., London Bioscience Innovation Centre, 2 Royal College Street, London NW1 0NH, UK
[4]XEGEN, 15 rue Dominique Piazza, 13420 Gemenos, France

*Correspondence: Pierre.pontarotti@univ-amu.fr (P. Pontarotti).

thymic selection impacts HLA class I/II polymorphism selection. We then proceed to elaborate on how peptide pools vary between individuals and how such variation can impact immune responses. Finally, we will conclude with the proposal that some of these polymorphic peptides represent IME.

Diversity in the Interaction between the HLA Class I/II–Peptide Complex and TCR

The central interaction in cell-mediated adaptive immunity occurs between the αβ TCR and the HLA class I/II–peptide complex (Box 1).

HLA Class I/II-Peptide Complex Diversity

HLA class I/II genes are highly polymorphic with over 15 000 allotypes encoded by the five most polymorphic genes (*HLA-A/B/C/DRB1/DQB1*) [6]. The polymorphism is concentrated around the peptide-binding cleft of the HLA class I/II proteins and thus distinct HLA class I/II allotypes typically have distinct peptide-binding preferences that are determined by anchor residues that reside within the PBD [7]. HLA class I/II variants can each present thousands of peptides [8], with an overlap between these peptides that is positively correlated with PBD similarity [9].

TCR Diversity

In contrast to HLA class I/II diversity, TCR diversity is generated during T cell ontogeny in the thymus, via random rearrangements of variable (V), diversity (D) and joining (J) gene segments from *TCR β* loci, and random rearrangement of V and J gene segments from the *TCR α* gene loci (Box 2). Yet, while this diversity is initially randomly generated, selection of αβTCR cells in the thymus involves immature thymocytes being subjected to both positive and negative selection [4] (Figure 1). The first step is the **positive selection** resulting in the survival of thymocytes that undergo intrathymic migration. This step ensures that the thymus produces T cells capable of antigen recognition. This selection is done in the thymus cortex where the self-peptide involved in positive selection seems to be generated by the thymoproteasome, a specialized type of proteasome [10,11]. The second step is the negative selection that ensures that T cell development produces functional thymocytes that are tolerant to self-antigens. This step occurs in the thymus medulla. The selection is carried out by eliminating the high binding T cells and keeping only T cells that can recognize the HLA class I/II–peptide complex with low-to-medium affinity. Because

Box 1. HLA Class I/II Proteins Present Peptide Antigens to Immune Receptors

HLA class I molecules present peptides derived from the proteins that are synthesized within the cell. The proteins are degraded by the proteasome into peptides in the cytosol and short peptides are translocated to the endoplasmic reticulum (ER) by the transporter associated with antigen presentation peptide transporter. Once in the ER, the translocated peptides are loaded into the HLA groove (peptide-binding cleft) with the help of chaperone proteins. These loaded peptides are typically 8–10 amino acids long and have specific anchor residues; longer peptides are trimmed by the aminopeptidase, ERAP1. Exocytosis of newly synthesized HLA class I molecules loaded with peptides on the cell surface initiates the interaction with the TCR of $CD8^+$ T cells. $CD8^+$ T cells are involved in cell-mediated immune responses that can directly induce cytotoxicity in virus-infected cells as well as in tumor cells.

HLA class II molecules present peptides that are longer than those presented by HLA class I molecules and that are derived from extracellular proteins degraded in the endosomal pathway. HLA class II molecules are assembled in the ER with the invariant chain Ii. Ii binds the HLA class II peptide-binding groove and directs the transport of HLA class II molecules into the endosomal pathway, where they will encounter exogenous peptides generated by proteases. Once in the endosomal pathway, Ii is degraded by proteases, leaving an invariant chain fragment (CLIP) inaccessible for proteases as it is protected by the peptide-binding groove of HLA class II molecules. This CLIP fragment will be replaced by higher affinity peptides with the help of the dedicated HLA-class-II-like chaperone DM, which locally opens the groove to release low-affinity peptides such as CLIP then locks in place the high affinity peptide replacements. The HLA class II–peptide complex will then translocate to the plasma membrane, where it can interact with $CD4^+$ T cells. These cells are helper cells that carry out multiple functions in cellular immunity, including cytokine secretion, activation of innate immune cells, and initiation of the adaptive immune responses.

Glossary

Allotype: a distinct variant at the protein level. Several allelic variants can code for the same allotype.
Human leukocyte antigen: the HLA region is a ~4-Mb genomic region on chromosome 6 (6p21.3) that contains a set of highly polymorphic genes named *HLA class I* and HLA *class II*. These genes encode cell surface proteins essential for the adaptive immune system because of their capacity to present peptide antigens to immune receptors and hence discriminate self versus non-self.
Immunopeptidome: The complete set of peptides encoded by a particular genome, or present within a particular cell type or organism that can be presented by HLA class I/II molecules.
Inherent modulator epitopes: epitopes of the self-proteome whose polymorphism can affect the negative or positive selection of T cells and can thus contribute to the shaping of immune system responses (resistance, immunotolerance).
Minor histocompatibility antigens: peptides of the self that differ between patients and donors, that are presented by HLA molecules at the cell surface, and that give rise to graft rejection in the context of bonne marrow transplantation (stem cells).
Negative selection: (i) evolution: natural selection removing deleterious alleles; (ii) immunology: deletion of T cells whose TCR recognizes with a high affinity HLA class I/II–self-peptide complex complexes of peptides–HLA class I/II molecules.
Peptidome: complete set of peptides encoded by a particular genome, or present within a particular cell type or organism.
Polymorphism: a distinct variant of the same element. In the context of peptides, polymorphism means two peptides originating from the exact same part of a protein but with a distinct sequence.
Positive selection: (i) evolution: natural selection increasing frequencies of advantageous mutations; (ii) immunology: retention of T cells whose TCR recognizes complexes of self-peptides and HLA class I/II.
T cell receptor: receptor of T cells that recognizes complexes of peptide–HLA class I/II. It is coded by V, D, and J genes: through somatic recombination,

Box 2. How TCR Variability Is Built

The generation of the TCR occurs by random rearrangement of variable (V), diversity (D), and joining (J) gene segments. Moreover, junctional diversity increases the TCR repertoire by addition of nontemplate-encoded nucleotides (N) at V(D)J junctions. V(D)J recombination is initiated early during lymphocyte development by a site-specific endonuclease: RAG1 and 2. The RAGs cleave at a conserved recombination signal that flanks each V, D, and J segments [42]. The random pairing of different TCRα and TCRβ chains increases the repertoire further. In humans, according to the number 42 V segments and 61 J segments are found for the TCR α chain locus, while 52 V, 2 D, and 13 J genes in the case TCR β chain locus [43]. Random rearrangements of genes encoding TCRα and TCRβ and random pairing can theoretically generate up to 1×10^{20} different αβ TCRs [44], although we note that the expressed diversity in any given individual is closer to 5×10^{6} αβ TCRs [45,46]. Combinations of HLA class I/II plus peptide interact with TCR hypervariable complementarity-determining regions (CDRs). The germ line V genes encode the CDR1 and CDR2 regions and the V(D)J junction encodes the CDR3 region. CDR1 and CDR2 interact with the HLA class I/II molecules, while CDR3 interacts with the peptide presented by HLA class I/II molecules [46].

these genes produce a vast number of distinct variants in each individual.

all genes are likely expressed in the thymus tissues, the whole peptidome thus plays a role at both the positive and negative selection steps.

HLA Class I/II–Peptide Complex Impacts TCR Repertoires

The role played by HLA class I/II molecules on the resistance/susceptibility to pathogens is twofold: first, through their capacity to bind pathogen-derived peptides and present them to T cells; and second, through their effect on T cell education in the thymus with the presentation of self-peptides. While the former was extensively discussed [2,3,12], the latter is not typically considered when discussing disease resistance/susceptibility in light of HLA class I/II variation.

Yet, negative thymic selection prunes T cells that are strongly self-reactive and recognize with high affinity HLA **allotypes** presenting peptides from the self. While this process does not eliminate all self-specific T cells [13], it does impact TCR repertoires. Similarly, positive thymic selection selects TCR that can recognize self-peptide–HLA class I/II complexes with enough affinity. It should be noted that there is uncertainty regarding how the strength of the affinity between HLA class I/II ligands and peptides, and between HLA class I/II–peptide complex and TCR could impact the selection.

Hence, with both types of thymic selection, the nature of the peptides from the self (**immunopeptidome**) contributes to the shaping of TCR repertoires. In the next section, we

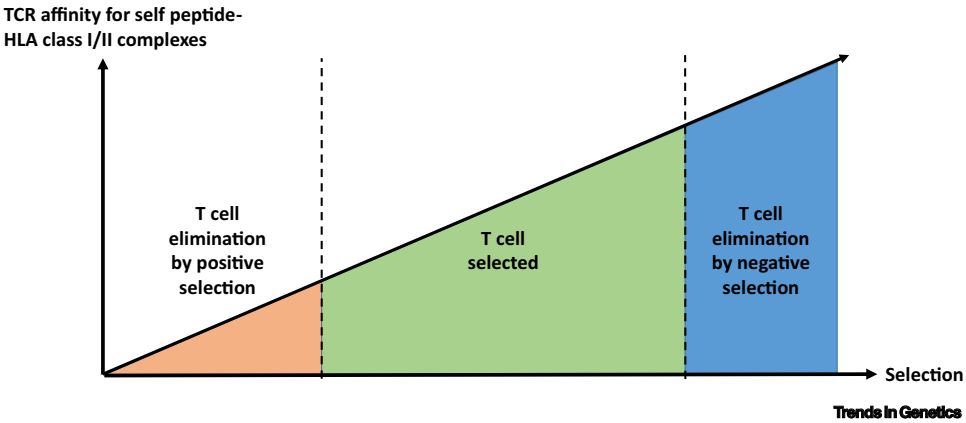

Figure 1. T Cell Selection Depends on the Affinity of Its TCR to Self Peptide-HLA Class I/II Complexes. If TCR affinity is low (orange), T cells are eliminated by positive selection. If TCR affinity is too high, T cells are eliminated by negative selection (blue). All others T cells are selected (green). Abbreviations: HLA, human leukocyte antigen; TCR, T cell receptor.

show that some of these peptides can be polymorphic and such variability thus leads to distinct TCR repertoires and adaptive immune responses.

Impact of Polymorphism of the Immunopeptidome on Immune Function: Case of Minor Histocompatibility Antigens

The importance of the polymorphism of the immunopeptidome is best illustrated by the **minor histocompatibility antigens** (MHAs) [14,15] that were first discovered when HLA class I/II identical donors and recipients gave immunological responses during bone marrow transplants. HLA class I/II molecules are a major factor associated with graft rejection as distinct HLA class I/II allotypes present distinct sets of peptides to educate T cells in the thymus and such TCR repertoire differences lead to allograft rejections [16]. Similarly, graft rejection can occur even in autologous situations (same HLA class I/II allotypes) because of the peptides that are polymorphic between these individuals and hence can lead to a differential pruning of TCRs that recognize with high affinity the combination of HLA class I/II allotypes and peptides from the self [17].

Several polymorphic peptides acting as MHA have long been described [17] but more are being discovered thanks to extensive genome sequencing and the development of high throughput approaches to predict or elute HLA class I/II peptides [18,19]. In particular, discovery of new MHAs is ensured by either forward or reverse immunology approaches. The former includes the isolation of alloreactive clones from the patient and subsequent inference of immunogenic missense nucleotide polymorphisms [18,20]. The latter is based on mass spectrometry (MS) analysis of HLA-eluted peptides, or by *in silico* predictions. The peptides detected by MS are then linked to polymorphic genomic regions [19,21].

The question that arises here is whether immunopeptidome polymorphisms have a role that goes beyond graft transplants and affect more generally resistance/susceptibility to diseases.

Peptidome Polymorphism and IMEs

The previous section illustrates how peptidome polymorphisms impact T cell education in the thymus and can lead to histocompatibility. To understand this effect, it is important to re-emphasize how each HLA class I/II allotype presents its own set of peptides and that combinations of HLA class I/II plus self-peptide define T cell repertoires. However, due to thymic negative selection, mature T cells can display a depleted repertoire that does not fully recognize certain HLA class I/II plus pathogen-derived peptide combinations, even if in theory the HLA class I/II molecules could present such peptides. The same reasoning can be applied to the polymorphic immunopeptidome and such polymorphic peptides that affect T cell repertoires can be referred to as IMEs. These IMEs can also drive thymus positive and negative selection and select TCR that will recognize pathogen-derived peptides presented by HLA class I/II [10,11].

Polymorphic peptides could correspond to nonsynonymous mutations on coding genes, but also to the presence/absence of genomic sequences that are potentially coding for proteins. Therefore, a wide variety of variation distributed along the genome can modify the immunopeptidome and hence impact TCR repertoires, and ultimately the ability of T cell to respond to diseases. Peptidome polymorphism can thus explain individual variation in disease resistance/susceptibility through their impact on TCR repertoires (Figure 2). This hypothesis could be tested by analysis of immunopeptidome variation from individuals with the same HLA types but with different disease resistance/susceptibility phenotypes.

While HLA class I/II variation is driven by natural selection, the same could be true for peptide polymorphisms that represent IME.

Key Figure

Impact of Self Peptidome Variation on the Generation of T Cell Repertoires for the Same HLA Class I/II Variants.

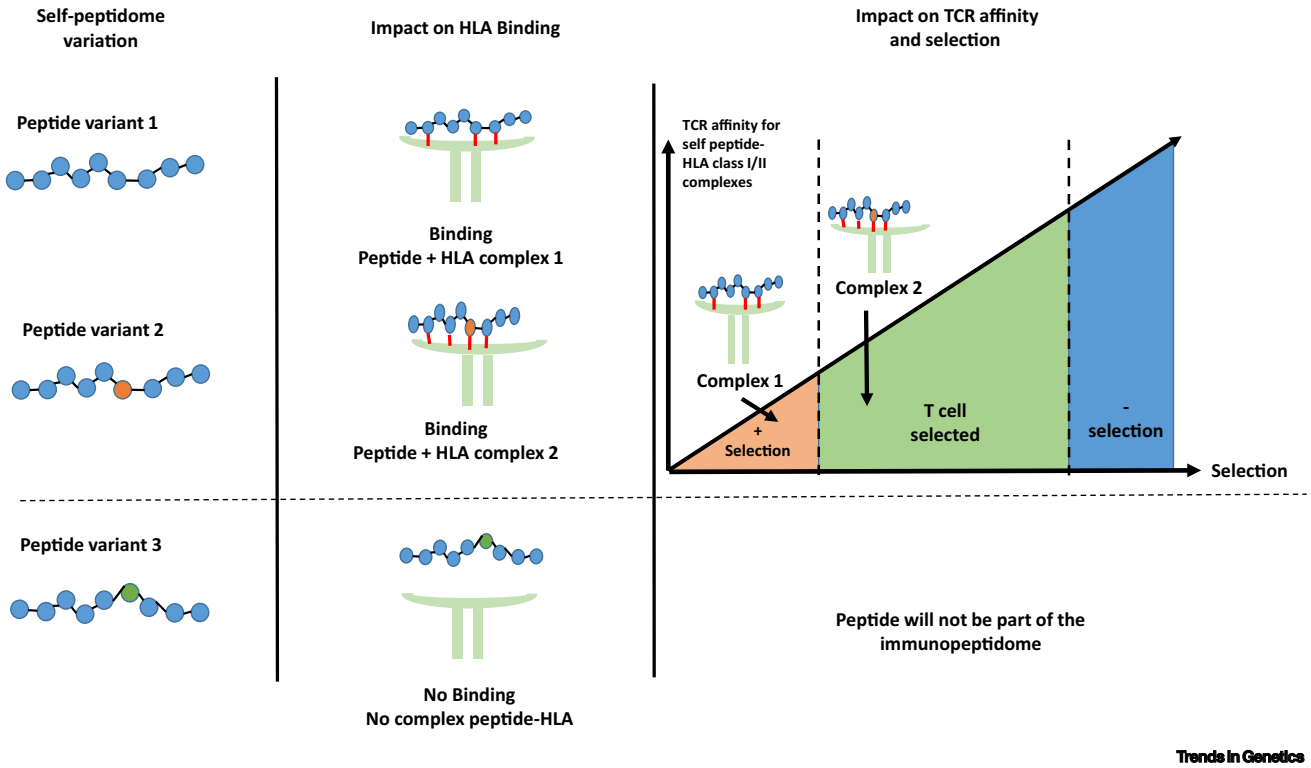

Figure 2. A self-peptidome variation can impact the binding of a peptide to the HLA class I/II variants of a given individual. This change will affect the affinity of TCRs for the HLA-peptide complex (peptide variant 1 and 2) and hence directly impact TCR repertoires. Moreover, self-peptidome variation could also generate a peptide that cannot be presented so it will not be part of the immune-peptidome, and that will also impact TCR repertoires (peptide variant 3). Abbreviations: HLA, human leukocyte antigen; TCR, T cell receptor.

Genomic Plasticity Increases Peptidome Variability: Inventory of Variation and Detail of the Common Mechanisms of Endogenization

We describe here how the peptidome contributes to shaping TCR repertoires and thus immune responses. any genomic variation modifying the peptidome can potentially impact immune responses. Such variation could occur at different levels, starting with direct and simple nonsynonymous mutations in protein-coding regions of the genome but potentially involving more indirect mechanisms like mutations impacting immunoproteome processing [22,23].

Other genomic variation, such as presence/absence of genes or indels in coding regions will also change the peptidome repertoire by adding or suppressing specific peptides. Similarly, variations affecting transcription can play a role: alternative splicing events will produce new peptides for the same coding regions. The transcription step itself could be affected by genomic variation or be under the control of epigenetic factor that can modify peptide expression in the thymus [24]. Differences in peptide expression levels could similarly modify T cell maturation, inducing a competition between peptides bound by HLA class I/II. In extreme cases, modulation of gene expression in the thymus could impact the peptidome just like gene absence/presence would. Noncoding

regions and their variability could also affect the peptidome as some pseudogenes can be partially translated for example. Other genetic events, like genome rearrangement involving transposons could also lead to the expression of new peptides [25,26].

Finally, endogenization of exogenous DNA is also potentially a phenomenon that could impact immune responses. Indeed, endogenous viruses and transposon repertoires vary between individuals [27–29], and such variation could generate IMEs. This phenomenon has the potential to enhance immune responses against exogenous retrovirus. This was, for example, reported in mice with endogenous retroviruses shaping T cell repertoires and hence allowing specific responses against a closely related exogenous retrovirus by crossreactivity [30]. Two other adaptive systems are known to use endogenization to fight pathogens: the clustered regularly interspaced short palindromic repeats (CRISPRs) in archaea and bacteria [31] and piRNA in eukaryotes [32]. In both cases foreign DNA prevents new infections [32,33]. Here, we propose that endogenization could be a widespread phenomenon that would lead to new IME and, through the resulting impact on TCR repertoires, allow immune systems to better fight pathogens.

Outstanding Questions

What is the respective contribution of IME versus HLA class I/II for TCR repertoires?

Could some MHAs also play a role as IMEs?

Could some IMEs be involved in the development of autoimmune diseases by selecting TCRs that can recognize with sufficient affinity self-peptides?

What is the part of the genome that could be under selection to preserve IMEs?

How to investigate the impact of IMEs on immune responses?

Concluding Remarks

The genetic basis for disease resistance/susceptibility is a complex mechanism but HLA class I/II variation is commonly involved. Consistent with this, natural selection commonly targets HLA class I/II molecules and they display a high level of polymorphism in human populations. This point is important and necessary to understand the human genetic-based resistance and its variation between individuals and populations, yet it does not explain all the differences in HLA–TCR biology. Indeed, even if HLA polymorphism could be selected to favor binding of pathogen peptides, this selection could not be efficient if T cells do not have adequate TCR to recognize these complexes between pathogen peptides and HLA class I/II. Hence, it is likely that selection also occurs on T cell maturation during the positive and negative selection in the thymus. The whole peptidome is involved in this maturation and thus plays a role in shaping TCR repertoires. Therefore, we suggest that selection could also target the peptidome and allow selection of TCR that can recognize HLA class I/II plus pathogen-derived peptides.

Currently, disease association studies target HLA class I/II variants [34] or use genome-wide association studies to isolate genomic polymorphisms [35–38]. Similarly to what was done to isolate new MHAs [39], analyses focused on peptidome such as immunopeptidome-wide association studies [40] could help explore the impact of changes in self immunopeptidome. Such analyses take into account both HLA variation and peptidome variation, in order to identify changes that could impact self-immunopeptidome due to an increased or a decreased number of self-peptides that can be presented by the HLA class I/II variants of one individual. This type of analysis could hence lead to the discovery of IME and provide ways to test interesting hypotheses, such as the possibility that gender specific differences in the immune response against COVID-19 [41] could be explained by the Y chromosome specific immunopeptidome (see Outstanding Questions).

Acknowledgments

We thank Jim Kaufman and Pierre Boudinot for helpful discussion. This work was supported by the French Government under the Investissements d'avenir (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr:National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03).

References

1. Antonelli, G. and Roilides, E. (2014) Host genetics: deciphering the variability in susceptibility to infections. *Clin. Microbiol. Infect.* 20, 1235–1236
2. Radwan, J. *et al.* (2020) Advances in the evolutionary understanding of MHC polymorphism. *Trends Genet.* 36, 298–311
3. Spurgin, L.G. and Richardson, D.S. (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. Biol. Sci.* 277, 979–988
4. Inglesfield, S. *et al.* (2019) Rethinking thymic tolerance: lessons from mice. *Trends Immunol.* 40, 279–291

5. Kosmrlj, A. *et al.* (2010) Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465, 350–354
6. Robinson, J. *et al.* (2020) IPD-IMGT/HLA database. *Nucleic Acids Res.* 48, D948–D955
7. Rossjohn, J. *et al.* (2015) T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* 33, 169–200
8. Sarkizova, S. *et al.* (2020) A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat. Biotechnol.* 38, 199–209
9. Gfeller, D. and Bassani-Sternberg, M. (2018) Predicting antigen presentation—what could we learn from a million peptides? *Front. Immunol.* 9, 1716
10. Takada, K. *et al.* (2017) Generation of peptides that promote positive selection in the thymus. *J. Immunol.* 198, 2215–2222
11. Takahama, Y. *et al.* (2019) Thymoproteasome and peptidic self. *Immunogenetics* 71, 217–221
12. 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
13. Yu, W. *et al.* (2015) Clonal Deletion Prunes but Does Not Eliminate Self-Specific alphabeta CD8(+) T Lymphocytes. *Immunity* 42, 929–941
14. Mandl, J.N. *et al.* (2013) T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* 38, 263–274
15. Griffioen, M. *et al.* (2016) Autosomal minor histocompatibility antigens: how genetic variants create diversity in immune targets. *Front. Immunol.* 7, 100
16. Klein, J. (1986) *Natural History of the Major Histocompatibility Complex*, Wiley
17. Wallny, H.J. and Rammensee, H.G. (1990) Identification of classical minor histocompatibility antigen as cell-derived peptide. *Nature* 343, 275–278
18. Oostvogels, R. *et al.* (2014) Identification of minor histocompatibility antigens based on the 1000 Genomes Project. *Haematologica* 99, 1854–1859
19. Granados, D.P. *et al.* (2014) Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides. *Nat. Commun.* 5, 3600
20. Torikai, H. *et al.* (2006) The human cathepsin H gene encodes two novel minor histocompatibility antigen epitopes restricted by HLA-A*3101 and -A*3303. *Br. J. Haematol.* 134, 406–416
21. Bykova, N.A. *et al.* (2018) *In silico* analysis of the minor histocompatibility antigen landscape based on the 1000 Genomes Project. *Front. Immunol.* 9, 1819
22. Liepe, J. *et al.* (2016) A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* 354, 354–358
23. Lopez de Castro, J.A. (2018) How ERAP1 and ERAP2 shape the peptidomes of disease-associated MHC-I proteins. *Front. Immunol.* 9, 2463
24. Ucar, O. and Rattay, K. (2015) Promiscuous gene expression in the thymus: a matter of epigenetics, miRNA, and more? *Front. Immunol.* 6, 93
25. Ewing, A.D. and Kazazian Jr., H.H. (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res.* 20, 1262–1270
26. Wang, L. and Jordan, I.K. (2018) Transposable element activity, genome regulation and human health. *Curr. Opin. Genet. Dev.* 49, 25–33
27. Mager, D.L. and Goodchild, N.L. (1989) Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. *Am. J. Hum. Genet.* 45, 848–854
28. Thomas, J. *et al.* (2018) Variation in proviral content among human genomes mediated by LTR recombination. *Mob. DNA* 9, 36
29. Wildschutte, J.H. *et al.* (2016) Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc. Natl. Acad. Sci. U. S. A.* 113, E2326-2334
30. Young, G.R. *et al.* (2012) Negative selection by an endogenous retrovirus promotes a higher-avidity CD4+ T cell response to retroviral infection. *PLoS Pathog.* 8, e1002709
31. Wiedenheft, B. *et al.* (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338
32. Ophinni, Y. *et al.* (2019) piRNA-guided CRISPR-like immunity in eukaryotes. *Trends Immunol.* 40, 998–1010
33. Yu, T. *et al.* (2019) The piRNA response to retroviral invasion of the koala genome. *Cell* 179, 632–643 e12
34. Dendrou, C.A. *et al.* (2018) HLA variation and disease. *Nat. Rev. Immunol.* 18, 325–339
35. Betz, R.C. *et al.* (2015) Genome-wide meta-analysis in alopecia areata resolves HLA associations and reveals two new susceptibility loci. *Nat. Commun.* 6, 5966
36. Okada, Y. *et al.* (2018) Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat. Commun.* 9, 1631
37. Tian, C. *et al.* (2017) Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat. Commun.* 8, 599
38. Waage, J. *et al.* (2018) Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis. *Nat. Genet.* 50, 1072–1080
39. Fuchs, K.J. *et al.* (2020) Optimized whole genome association scanning for discovery of HLA class I-restricted minor histocompatibility antigens. *Front. Immunol.* 11, 659
40. Vizcaino, J.A. *et al.* (2020) The Human Immunopeptidome Project: a roadmap to predict and treat immune diseases. *Mol. Cell. Proteomics* 19, 31–49
41. Takahashi, T. *et al.* (2020) Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature* Published online August 26, 2020. https://doi.org/10.1038/s41586-020-2700-3
42. Teng, G. and Schatz, D.G. (2015) Regulation and evolution of the RAG recombinase. *Adv. Immunol.* 128, 1–39
43. Mak, T.W. *et al.* (2013) *Primer to the Immune Response*, Newnes
44. Lieber, M.R. (1991) Site-specific recombination in the immune system. *FASEB J.* 5, 2934–2944
45. Arstila, T.P. *et al.* (1999) A direct estimate of the human alphabeta T cell receptor diversity. *Science* 286, 958–961
46. Bradley, P. and Thomas, P.G. (2019) Using T cell receptor repertoires to understand the principles of adaptive immune recognition. *Annu. Rev. Immunol.* 37, 547–570