



**HAL**  
open science

# Self-Peptidome Variation Shapes Individual Immune Responses

Pierre Pontarotti, Laurent Abi-Rached, Jung-Hua Yeh, Julien Paganini

► **To cite this version:**

Pierre Pontarotti, Laurent Abi-Rached, Jung-Hua Yeh, Julien Paganini. Self-Peptidome Variation Shapes Individual Immune Responses. *Trends in Genetics*, 2021, 37 (5), pp.414 - 420. 10.1016/j.tig.2020.10.001 . hal-03370438

**HAL Id: hal-03370438**

**<https://hal.science/hal-03370438>**

Submitted on 11 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 1 **Self-Peptidome Variation Shapes Individual Immune Responses**

2 Pierre Pontarotti<sup>1,2,4</sup>, Laurent Abi-Rached<sup>1,2</sup>, Jung-Hua Yeh<sup>3</sup>, and Julien Paganini<sup>4</sup>

3

4 1 Aix Marseille Université, IRD, APHM, MEPHI, IHU Méditerranée Infection, 19–21  
5 Boulevard Jean Moulin, 13005 Marseille, France.

6 2 SNC5039 CNRS, 19–21 Boulevard Jean Moulin, 13005 Marseille, France.

7 3 Prokarium Ltd, London Bioscience Innovation Centre, 2 Royal College Street, London  
8 NW1 0NH, UK.

9 4 XEGEN, 15 rue Dominique Piazza, 13420 Gemenos, France.

10 \*Correspondance : Pierre.pontarotti@univ-amu.fr (P. Pontarotti).

11

12 **Keywords: HLA, TCR, Disease, Inherent Modulator Epitopes, Immune responses,**  
13 **Peptide.**

14

## 15 **Abstract**

16 The relationship between human genetic variation and disease has not been fully  
17 elucidated. According to the present view on infectious diseases pathogen resistance is  
18 linked to Human Leukocyte Antigen (HLA) class I/II variants and their individual capacity  
19 to present pathogen-derived peptides. Yet, T cell education in the thymus occurs though  
20 negative and positive selection, and both processes are controlled by a combination of  
21 HLA class I/II variants and peptides from the self. Therefore, the capacity of given HLA  
22 class I/II variants to bind pathogen-derived peptides is only one part of the selective  
23 process to generate effective immune responses. We thus propose that peptidome  
24 variation contributes to shaping TCR repertoires and hence individual immune  
25 responses, and that this variation represents Inherent Modulator Epitopes.

26 **Introduction: disease resistance/susceptibility is linked to Inherent Modulator**  
27 **Epitopes (IME)**

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

38 We propose here that this conclusion is incomplete and that polymorphism in the  
39 peptides derived from the self also impacts disease resistance/susceptibility, playing a  
40 role as **Inherent Modulator Epitopes (IME)**. T cell education in the thymus is done  
41 through positive and negative selection. First, a positive selection of T cells capable of  
42 antigen recognition and then **negative selection** (see Glossary) to eliminate T cells  
43 displaying high affinity for the combination of HLA class I/II plus self-peptide (self-antigen  
44 tolerance) [4]. While the shaping of T Cell Receptor (TCR) repertoires is a complex  
45 process involving different mechanisms, some of them still not well described or  
46 understood, it is accepted that these two selection steps rely, in part, on peptide pools  
47 derived from the self (peptidome) [5]. Peptide polymorphisms can thus impact TCR  
48 repertoires and modulate the capacity of these repertoires to recognize pathogen-

49 derived peptides. This would also explain differences in disease resistance/susceptibility  
50 for individuals that display the same sets of HLA class I/II variants.

51 The first part of this opinion article corresponds to a brief state-of-the-art regarding HLA  
52 class I/II and TCR structural biology, in particular in the context of thymic selection. We  
53 then discuss how thymic selection impacts HLA class I/II polymorphism selection. We  
54 then proceed to elaborate on how peptide pools vary between individuals and how such  
55 variation can impact immune responses. Finally, we will conclude with the proposal that  
56 some of these polymorphic peptides represent IME.

57

### 58 **Diversity in the interaction between the HLA class I/II-peptide complex and the TCR**

59 The central interaction in cell-mediated adaptive immunity occurs between the  $\alpha\beta$  TCR  
60 ( $\alpha\beta$ TCR) and the HLA class I/II-peptide complex (Box 1).

61

#### 62 HLA class I/II-peptide complex *diversity*

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

## 70 *TCR diversity*

71 In contrast to HLA class I/II diversity, TCR diversity is generated during T cell ontogeny  
72 in the thymus, via random rearrangements of variable (V), diversity (D) and joining (J)  
73 gene segments from *TCR β* loci, and random re-arrangement of V and J gene segments  
74 from the *TCR α* gene loci (Box 2). Yet, while this diversity is initially randomly generated,  
75 selection of  $\alpha\beta$ TCR cells in the thymus involves immature thymocytes being subjected to  
76 both positive and negative selection [4] (Figure 1). The first step is the positive selection  
77 resulting in the survival of thymocytes that undergo intrathymic migration. This step  
78 ensures that the thymus produces T cells capable of antigen recognition. This selection  
79 is done in the thymus cortex where the self-peptide involved in positive selection seems  
80 to be generated by the thymoproteasome, a specialized type of proteasome [10, 11].  
81 The second step is the negative selection that ensures that T cell development produces  
82 functional thymocytes that are tolerant to self-antigens. This step occurs in the thymus  
83 medulla. The selection is carried out by eliminating the high binding T cells and keeping  
84 only T cells that can recognize the HLA class I/II-peptide complex with low-to-medium  
85 affinity. Because all genes are likely expressed in the thymus tissues, the whole  
86 peptidome thus plays a role at both the positive and negative selection steps.

87

88

89 **The HLA class I/II-peptide complex impacts TCR repertoires**

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

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

104 Hence with both types of thymic selection, the nature of the peptides from the self  
105 (immunopeptidome) contributes to the shaping of TCR repertoires. In the next section,  
106 we show that some of these peptides can be polymorphic and such variability thus leads  
107 to distinct TCR repertoires and adaptive immune responses.

108

109

110 **Impact of the polymorphism of the immunopeptidome on the immune function:**  
111 **the case of the Minor Histocompatibility Antigens**

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

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

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

134

### 135 **Peptidome polymorphism and Inherent Modulator Epitopes**

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

148 Polymorphic peptides could correspond to non-synonymous mutations on coding genes,  
149 but also to the presence/absence of genomic sequences that are potentially coding for  
150 proteins. Therefore, a wide variety of variation distributed along the genome can modify  
151 the immunopeptidome and hence impact TCR repertoires, and ultimately the ability of T  
152 cell to respond to diseases. Peptidome polymorphism can thus explain individual



153 variation in disease resistance/susceptibility through their impact on TCR repertoires  
154 (Figure 2). This hypothesis could be tested by analysis of immunopeptidome variation  
155 from individuals with the same HLA types but with different disease  
156 resistance/susceptibility phenotypes.

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

159  
160 **Genomic plasticity increases peptidome variability: inventory of variation and**  
161 **detail of the common mechanisms of endogenization.**

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

168 Other genomic variation, such as presence/absence of genes or indels in coding regions  
169 will also change the peptidome repertoire by adding or suppressing specific peptides.  
170 Similarly, variations affecting transcription can play a role: alternative splicing events will  
171 produce new peptides for the same coding regions. The transcription step itself could be  
172 affected by genomic variation or be under the control of epigenetic factor that can modify  
173 peptide expression in the thymus [24]. Differences in peptide expression levels could

174 similarly modify T cell maturation, inducing a competition between peptides bound by  
175 HLA class I/II. In extreme cases, modulation of gene expression in the thymus could  
176 impact the peptidome just like gene absence/presence would. Non-coding regions and  
177 their variability could also affect the peptidome as some pseudogenes can be partially  
178 translated for example. Other genetic events, like genome rearrangement involving  
179 transposons could also lead to the expression of new peptides [25, 26].

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

192

## 193 **Conclusion**

194 The genetic basis for disease resistance/susceptibility is a complex mechanism but HLA  
195 class I/II variation is commonly involved. Consistent with this, natural selection

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

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

219

## 220 **Acknowledgements**

221 We thank Jim Kaufman and Pierre Boudinot for helpful discussion

222 This work was supported by the French Government under the «Investissements  
223 d'avenir» (Investments for the Future) program managed by the Agence Nationale de la  
224 Recherche (ANR, fr:National Agency for Research), (reference: Méditerranée Infection  
225 10-IAHU-03).

226

## 227 **References**

228

- 229 1. Antonelli, G. and E. Roilides, *Host genetics: deciphering the variability in susceptibility to*  
230 *infections*. Clin Microbiol Infect, 2014. **20**(12): p. 1235-6.
- 231 2. Radwan, J., et al., *Advances in the Evolutionary Understanding of MHC Polymorphism*. Trends  
232 Genet, 2020. **36**(4): p. 298-311.
- 233 3. Spurgin, L.G. and D.S. Richardson, *How pathogens drive genetic diversity: MHC, mechanisms and*  
234 *misunderstandings*. Proc Biol Sci, 2010. **277**(1684): p. 979-88.
- 235 4. Inglesfield, S., et al., *Rethinking Thymic Tolerance: Lessons from Mice*. Trends Immunol, 2019.  
236 **40**(4): p. 279-291.
- 237 5. Kosmrlj, A., et al., *Effects of thymic selection of the T-cell repertoire on HLA class I-associated*  
238 *control of HIV infection*. Nature, 2010. **465**(7296): p. 350-4.
- 239 6. Robinson, J., et al., *IPD-IMGT/HLA Database*. Nucleic Acids Res, 2020. **48**(D1): p. D948-D955.
- 240 7. Rossjohn, J., et al., *T cell antigen receptor recognition of antigen-presenting molecules*. Annu Rev  
241 Immunol, 2015. **33**: p. 169-200.
- 242 8. Sarkizova, S., et al., *A large peptidome dataset improves HLA class I epitope prediction across*  
243 *most of the human population*. Nat Biotechnol, 2020. **38**(2): p. 199-209.
- 244 9. Gfeller, D. and M. Bassani-Sternberg, *Predicting Antigen Presentation-What Could We Learn*  
245 *From a Million Peptides?* Front Immunol, 2018. **9**: p. 1716.
- 246 10. Takada, K., K. Kondo, and Y. Takahama, *Generation of Peptides That Promote Positive Selection in*  
247 *the Thymus*. J Immunol, 2017. **198**(6): p. 2215-2222.
- 248 11. Takahama, Y., et al., *Thymoproteasome and peptidic self*. Immunogenetics, 2019. **71**(3): p. 217-  
249 221.
- 250 12. Pierini, F. and T.L. Lenz, *Divergent Allele Advantage at Human MHC Genes: Signatures of Past*  
251 *and Ongoing Selection*. Mol Biol Evol, 2018. **35**(9): p. 2145-2158.
- 252 13. Yu, W., et al., *Clonal Deletion Prunes but Does Not Eliminate Self-Specific alphabeta CD8(+) T*  
253 *Lymphocytes*. Immunity, 2015. **42**(5): p. 929-41.

- 254 14. Mandl, J.N., et al., *T cell-positive selection uses self-ligand binding strength to optimize repertoire*  
255 *recognition of foreign antigens*. *Immunity*, 2013. **38**(2): p. 263-274.
- 256 15. Griffioen, M., C.A. van Bergen, and J.H. Falkenburg, *Autosomal Minor Histocompatibility*  
257 *Antigens: How Genetic Variants Create Diversity in Immune Targets*. *Front Immunol*, 2016. **7**: p.  
258 100.
- 259 16. Klein, J., *Natural History of the Major Histocompatibility Complex*. 1986: Wiley.
- 260 17. Wallny, H.J. and H.G. Rammensee, *Identification of classical minor histocompatibility antigen as*  
261 *cell-derived peptide*. *Nature*, 1990. **343**(6255): p. 275-8.
- 262 18. Oostvogels, R., et al., *Identification of minor histocompatibility antigens based on the 1000*  
263 *Genomes Project*. *Haematologica*, 2014. **99**(12): p. 1854-9.
- 264 19. Granados, D.P., et al., *Impact of genomic polymorphisms on the repertoire of human MHC class I-*  
265 *associated peptides*. *Nat Commun*, 2014. **5**: p. 3600.
- 266 20. Torikai, H., et al., *The human cathepsin H gene encodes two novel minor histocompatibility*  
267 *antigen epitopes restricted by HLA-A\*3101 and -A\*3303*. *Br J Haematol*, 2006. **134**(4): p. 406-16.
- 268 21. Bykova, N.A., D.B. Malko, and G.A. Efimov, *In Silico Analysis of the Minor Histocompatibility*  
269 *Antigen Landscape Based on the 1000 Genomes Project*. *Front Immunol*, 2018. **9**: p. 1819.
- 270 22. Liepe, J., et al., *A large fraction of HLA class I ligands are proteasome-generated spliced peptides*.  
271 *Science*, 2016. **354**(6310): p. 354-358.
- 272 23. Lopez de Castro, J.A., *How ERAP1 and ERAP2 Shape the Peptidomes of Disease-Associated MHC-I*  
273 *Proteins*. *Front Immunol*, 2018. **9**: p. 2463.
- 274 24. Ucar, O. and K. Rattay, *Promiscuous Gene Expression in the Thymus: A Matter of Epigenetics,*  
275 *miRNA, and More?* *Front Immunol*, 2015. **6**: p. 93.
- 276 25. Ewing, A.D. and H.H. Kazazian, Jr., *High-throughput sequencing reveals extensive variation in*  
277 *human-specific L1 content in individual human genomes*. *Genome Res*, 2010. **20**(9): p. 1262-70.
- 278 26. Wang, L. and I.K. Jordan, *Transposable element activity, genome regulation and human health*.  
279 *Curr Opin Genet Dev*, 2018. **49**: p. 25-33.
- 280 27. Mager, D.L. and N.L. Goodchild, *Homologous recombination between the LTRs of a human*  
281 *retrovirus-like element causes a 5-kb deletion in two siblings*. *Am J Hum Genet*, 1989. **45**(6): p.  
282 848-54.
- 283 28. Thomas, J., H. Perron, and C. Feschotte, *Variation in proviral content among human genomes*  
284 *mediated by LTR recombination*. *Mob DNA*, 2018. **9**: p. 36.
- 285 29. Wildschutte, J.H., et al., *Discovery of unfixed endogenous retrovirus insertions in diverse human*  
286 *populations*. *Proc Natl Acad Sci U S A*, 2016. **113**(16): p. E2326-34.
- 287 30. Young, G.R., et al., *Negative selection by an endogenous retrovirus promotes a higher-avidity*  
288 *CD4+ T cell response to retroviral infection*. *PLoS Pathog*, 2012. **8**(5): p. e1002709.
- 289 31. Wiedenheft, B., S.H. Sternberg, and J.A. Doudna, *RNA-guided genetic silencing systems in*  
290 *bacteria and archaea*. *Nature*, 2012. **482**(7385): p. 331-8.
- 291 32. Ophinni, Y., et al., *piRNA-Guided CRISPR-like Immunity in Eukaryotes*. *Trends Immunol*, 2019.  
292 **40**(11): p. 998-1010.
- 293 33. Yu, T., et al., *The piRNA Response to Retroviral Invasion of the Koala Genome*. *Cell*, 2019. **179**(3):  
294 p. 632-643 e12.
- 295 34. Dendrou, C.A., et al., *HLA variation and disease*. *Nat Rev Immunol*, 2018. **18**(5): p. 325-339.
- 296 35. Betz, R.C., et al., *Genome-wide meta-analysis in alopecia areata resolves HLA associations and*  
297 *reveals two new susceptibility loci*. *Nat Commun*, 2015. **6**: p. 5966.
- 298 36. Okada, Y., et al., *Deep whole-genome sequencing reveals recent selection signatures linked to*  
299 *evolution and disease risk of Japanese*. *Nat Commun*, 2018. **9**(1): p. 1631.
- 300 37. Tian, C., et al., *Genome-wide association and HLA region fine-mapping studies identify*  
301 *susceptibility loci for multiple common infections*. *Nat Commun*, 2017. **8**(1): p. 599.

302 38. Waage, J., et al., *Genome-wide association and HLA fine-mapping studies identify risk loci and*  
303 *genetic pathways underlying allergic rhinitis*. Nat Genet, 2018. **50**(8): p. 1072-1080.  
304 39. Fuchs, K.J., et al., *Optimized Whole Genome Association Scanning for Discovery of HLA Class I-*  
305 *Restricted Minor Histocompatibility Antigens*. Front Immunol, 2020. **11**: p. 659.  
306 40. Vizcaino, J.A., et al., *The Human Immunopeptidome Project: A Roadmap to Predict and Treat*  
307 *Immune Diseases*. Mol Cell Proteomics, 2020. **19**(1): p. 31-49.  
308 41. Takahashi, T., et al., *Sex differences in immune responses that underlie COVID-19 disease*  
309 *outcomes*. Nature, 2020.  
310 42. Teng, G. and D.G. Schatz, *Regulation and Evolution of the RAG Recombinase*. Adv Immunol, 2015.  
311 **128**: p. 1-39.  
312 43. Mak, T.W., M.E. Saunders, and B.D. Jett, *Primer to the immune response*. 2013: Newnes.  
313 44. Lieber, M.R., *Site-specific recombination in the immune system*. FASEB J, 1991. **5**(14): p. 2934-44.  
314 45. Arstila, T.P., et al., *A direct estimate of the human alphabeta T cell receptor diversity*. Science,  
315 1999. **286**(5441): p. 958-61.  
316 46. Bradley, P. and P.G. Thomas, *Using T Cell Receptor Repertoires to Understand the Principles of*  
317 *Adaptive Immune Recognition*. Annu Rev Immunol, 2019. **37**: p. 547-570.

318

319

320

321

322

323

324

325

326

327

328 **Figure 1: T cell selection depends on the affinity of its TCR to self peptide-HLA**  
329 **class I/II complexes.** If TCR affinity is low (orange), T cells are eliminated by positive  
330 selection. If TCR affinity is too high, T cells are eliminated by negative selection (blue).  
331 All others T cells are selected (green).

332

333 **Figure 2: Impact of self peptidome variation on the generation of T cell repertoires**  
334 **for the same HLA class I/II variants.** A self-peptidome variation can impact the binding  
335 of a peptide to the HLA class I/II variants of a given individual. This change will affect the  
336 affinity of TCRs for the HLA-peptide complex (peptide variant #1 and #2) and hence  
337 directly impact TCR repertoires. Moreover, self-peptidome variation could also generate  
338 a peptide that cannot be presented so it will not be part of the immune-peptidome, and  
339 that will also impact TCR repertoires (peptide variant #3).

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354 **Box 1: HLA class I/II proteins present peptide antigens to immune receptors**

355 HLA class I molecules present peptides derived from the proteins that are synthesized  
356 within the cell. The proteins are degraded by the proteasome into peptides in the cytosol  
357 and short peptides are translocated to the Endoplasmic Reticulum (ER) by the  
358 transporter associated with antigen presentation (TAP) peptide transporter. Once in the  
359 ER, the translocated peptides are loaded into the HLA groove (peptide-binding cleft) with  
360 the help of chaperone proteins. These loaded peptides are typically 8-10 amino acid  
361 long and have specific anchor residues; longer peptides are trimmed by the  
362 aminopeptidase, ERAP1. Exocytosis of newly synthesized HLA class I molecules loaded  
363 with peptides on the cell surface initiates the interaction with the TCR of CD8<sup>+</sup> T cells.  
364 CD8<sup>+</sup> T cells are involved in cell-mediated immune responses that can directly induce  
365 cytotoxicity in virus-infected cells as well as in tumor cells.

366 HLA class II molecules present peptides that are longer than those presented by HLA  
367 class I molecules and that are derived from extracellular proteins degraded in the  
368 endosomal pathway. HLA class II molecules are assembled in the ER with the invariant  
369 chain Ii. Ii binds the HLA class II peptide-binding groove and directs the transport of HLA  
370 class II molecules into the endosomal pathway, where they will encounter exogenous  
371 peptides generated by proteases. Once in the endosomal pathway, Ii is degraded by  
372 proteases, leaving an invariant chain fragment (CLIP) inaccessible for proteases as it is  
373 protected by the peptide-binding groove of HLA class II molecules. This CLIP fragment  
374 will be replaced by higher affinity peptides with the help of the dedicated HLA class II-  
375 like chaperone DM, which locally opens the groove to release low-affinity peptides such  
376 as CLIP then locks in place the high affinity peptide replacements. The HLA class II-



377 peptide complex will then translocate to the plasma membrane, where it can interact  
378 with CD4<sup>+</sup> T cells. These cells are "helper cells" that carry out multiple functions in  
379 cellular immunity, including cytokine secretion, activation of innate immune cells, and  
380 initiation of the adaptive immune responses.

381

## 382 **Box 2: How TCR variability is built**

383 The generation of the TCR occurs by random rearrangement of variable (V), diversity  
384 (D), and joining (J) gene segments. Moreover, junctional diversity increases the TCR  
385 repertoire by addition of non-template encoded nucleotides (N) at V (D) J Junctions.  
386 V(D)J recombination is initiated early during lymphocyte development by a site-specific  
387 endonuclease: RAG1 and 2. The RAGs cleave at a conserved recombination signal  
388 (RSS) that flanks each V, D, and J segments [42]. The random pairing of different  
389 TCR $\alpha$ -chain and TCR $\beta$ -chains increases the repertoire further. In humans, according to  
390 the number 42 V segments and 61 J segments are found for the TCR  $\alpha$ -chain locus,  
391 while 52 V, 2 D and 13 J genes in the case TCR  $\beta$ -chain locus[43]. Random  
392 rearrangements of genes encoding TCR $\alpha$  and TCR $\beta$  and random pairing can  
393 theoretically generate up to  $1 \times 10^{20}$  different  $\alpha\beta$  TCRs [44], although we note that the  
394 expressed diversity in any given individual is closer to  $5 \times 10^6$   $\alpha\beta$  TCRs [45, 46].  
395 Combinations of HLA class I/II plus peptide interact with TCR hypervariable  
396 complementarity-determining regions (CDRs). The germ line V genes encode the CDR1  
397 and CDR2 regions and the V(D)J junction encodes the CDR3 region. CDR1 and CDR2

398 interact with the HLA class I/II molecules, while CDR3 interacts with the peptide  
399 presented by HLA class I/II molecules [46].

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416 **Glossary**

417 **Allotype:** A distinct variant at the protein level. Several allelic variants can code for the  
418 same allotype.

419 **Immunoepitidome:** The complete set of peptides encoded by a particular genome, or  
420 present within a particular cell type or organism that can be presented by HLA class I/II  
421 molecules.

422 **Inherent Modulator Epitopes (IME):** epitopes of the self-proteome whose  
423 polymorphism can affect the negative or positive selection of T Cells and can thus  
424 contribute to the shaping of immune system responses (resistance, immunotolerance).

425 **HLA:** The Human Leukocyte Antigen (HLA) region is a ~4Mb genomic region on  
426 chromosome 6 (6p21.3) that contains a set of highly polymorphic genes named *HLA*  
427 *class I* and *class II*. These genes encode cell surface proteins essential for the adaptive  
428 immune system because of their capacity to present peptide antigens to immune  
429 receptors and hence discriminate 'self' versus 'non-self'.

430 **Minor Histocompatibility Antigens (MHA):** peptides of the self that differ between  
431 patients and donors, that are presented by HLA molecules at the cell surface, and that  
432 give rise to graft rejection in the context of bone marrow transplantation (stem cells).

433 **Negative selection:** (i) evolution: natural selection removing deleterious alleles; (ii)  
434 immunology: deletion of T cells whose TCR recognizes with a high affinity HLA class I/II-  
435 self-peptide complex complexes of peptides-HLA class I/II molecules.

436 **Polymorphism:** A distinct variant of the same element. In the context of peptides,  
437 polymorphism means two peptides originating from the exact same part of a protein but  
438 with a distinct sequence.

439 **Peptidome:** The complete set of peptides encoded by a particular genome, or present  
440 within a particular cell type or organism.

441 **Positive selection:** (i) evolution: natural selection increasing frequencies of  
442 advantageous mutations; (ii) immunology: retention of T cells whose TCR recognizes  
443 complexes of self peptides-HLA class I/II.

444 **T cell receptor (TCR):** receptor of T cells that recognizes complexes of peptide-HLA  
445 class I/II. It is coded by V, D, and J genes: through somatic recombination, these genes  
446 produce a vast number of distinct variants in each individual.

## Highlights

T Cell Receptor (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 Human Leukocyte Antigen (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.

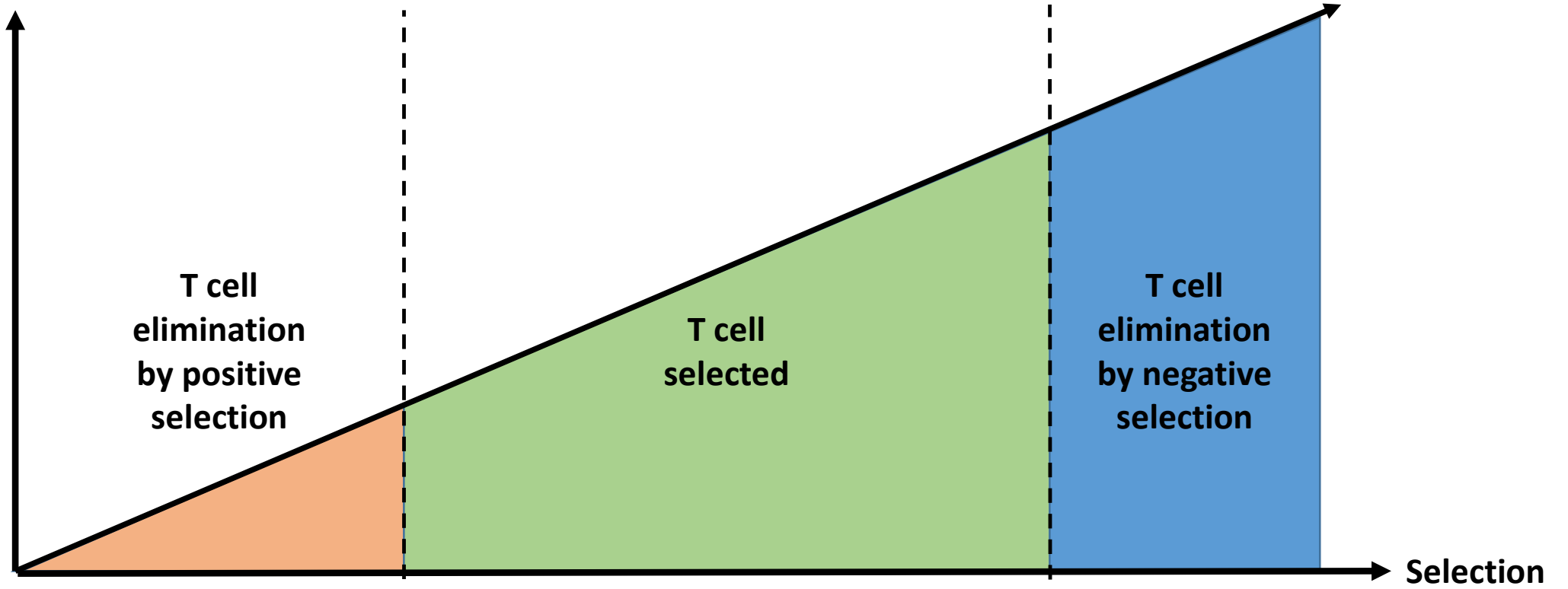
Self-peptide variants that affect T cell repertoires represent Inherent Modulator Epitopes.

## **Outstanding Questions**

- What is the respective contribution of IME vs HLA class I/II for TCR repertoires?
- Could some MHA also play a role as IME?
- Could some IME be involved in the development of autoimmune diseases by selecting TCR that can recognize with sufficient affinity self-peptides?
- What is the part of the genome that could be under selection to preserve IME?
- How to investigate the impact of IME on immune responses?

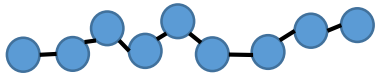
Figure1

TCR affinity for self peptide-  
HLA class I/II complexes



### Self-peptidome variation

Peptide variant 1



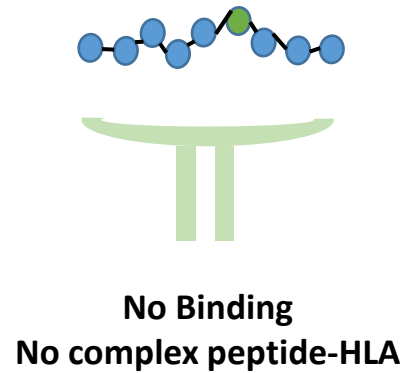
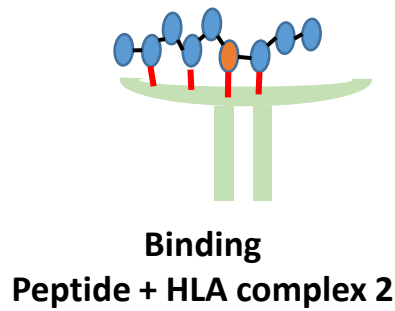
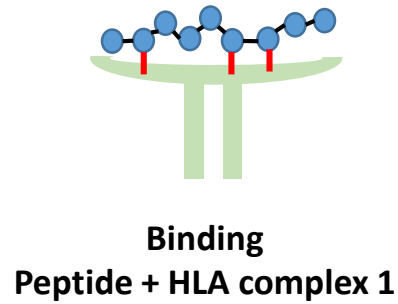
Peptide variant 2



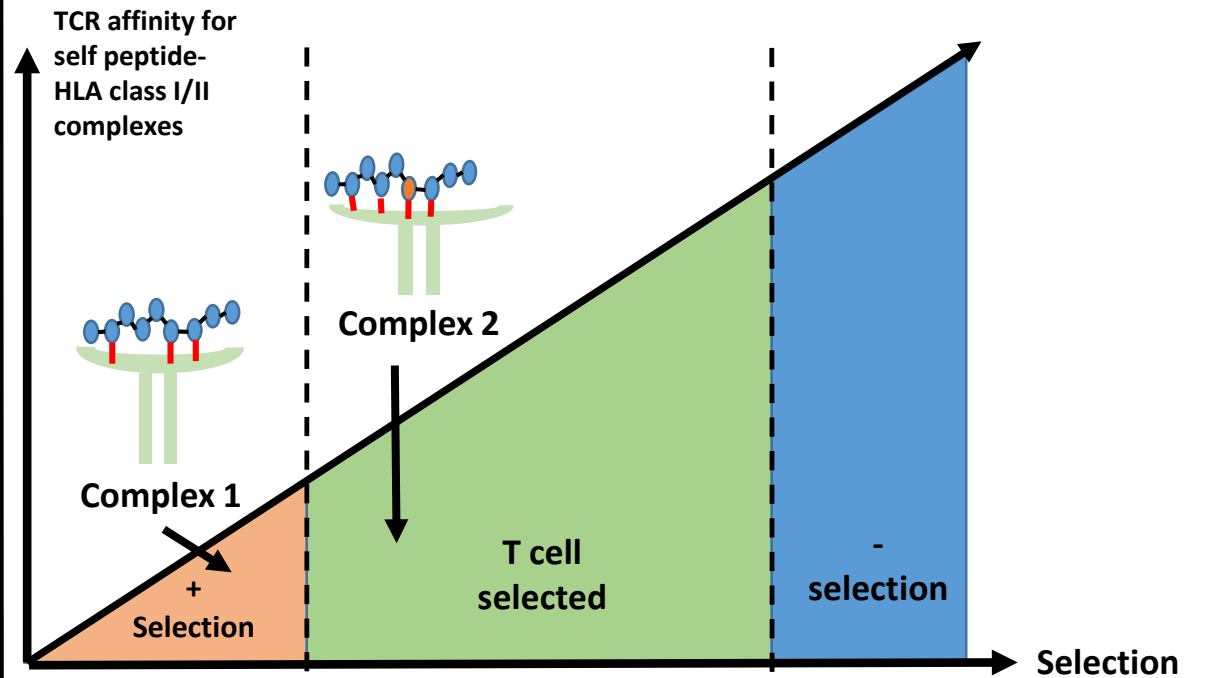
Peptide variant 3



### Impact on HLA Binding



### Impact on TCR affinity and selection



Peptide will not be part of the immunopeptidome