

1 The Primate Major 2 Histocompatibility Complex: An 3 Illustrative Example of Gene Family 4 Evolution

5 **Alyssa Lyn Fortier^{1,2*} and Jonathan K. Pritchard^{1,2}**

*For correspondence:

afortier@stanford.edu (ALF);

pritch@stanford.edu (JKP)

6 ¹Department of Biology, Stanford University, Stanford, CA USA; ²Department of
7 Genetics, Stanford University, Stanford, CA USA

9 **Abstract** Gene families are groups of evolutionarily-related genes. One large gene family that
10 has experienced rapid evolution is the Major Histocompatibility Complex (MHC), whose proteins
11 serve critical roles in innate and adaptive immunity. Across the ~60 million year history of the
12 primates, some MHC genes have turned over completely, some have changed function, some
13 have converged in function, and others have remained essentially unchanged. Past work has
14 typically focused on identifying MHC alleles within particular species or comparing gene content,
15 but more work is needed to understand the overall evolution of the gene family across species.
16 Thus, despite the immunologic importance of the MHC and its peculiar evolutionary history, we
17 lack a complete picture of MHC evolution in the primates. We readdress this question using
18 sequences from dozens of MHC genes and pseudogenes spanning the entire primate order,
19 building a comprehensive set of gene and allele trees with modern methods. Overall, we find that
20 the Class I gene subfamily is evolving much more quickly than the Class II gene subfamily, with
21 the exception of the Class II MHC-DRB genes. We also pay special attention to the often-ignored
22 pseudogenes, which we use to reconstruct different events in the evolution of the Class I region.
23 We find that despite the shared function of the MHC across species, different species employ
24 different genes, haplotypes, and patterns of variation to achieve a successful immune response.
25 Our trees and extensive literature review represent the most comprehensive look into MHC
26 evolution to date.

28 Introduction

29 Gene families are groups of related genes categorized by functional similarity or presumed evolu-
30 tionary relatedness. Based on clustering of their proteins' sequences, human genes fall into hun-
31 dreds to thousands of distinct families (*Gu et al. 2002; Li et al. 2001; Demuth et al. 2006; Friedman*
32 *and Hughes 2003*). Families originate from successive gene duplications, although particular gene
33 copies or entire families can also be lost (*Nei et al. 1997; Demuth et al. 2006*). For example, there
34 are hundreds of genes that are specific to human or chimpanzee and have no orthologs in the
35 other species (*Demuth et al. 2006*). This birth-and-death evolution is distinct from evolution at the
36 nucleotide or protein level (*Thornton and Desalle 2000; Hahn et al. 2005*). However, phylogenet-
37 ics can still be applied to understand the relationships within families of genes, providing insight
38 into speciation and specialization (*Thornton and Desalle 2000*).

39 One large gene family is united by a common protein structure called the MHC fold. This

40 group of genes is involved in diverse functions including lipid metabolism, iron uptake regulation,
41 and immune system function (*Hansen et al. 2007; Kupfermann et al. 1999; Kaufman 2022; Adams*
42 *and Luoma 2013*). This family includes the Class I and Class II MHC genes whose protein products
43 present peptides to T-cells (classical genes) and/or interact with other immune cell receptors like
44 killer cell immunoglobulin-like receptors (KIRs) or leukocyte immunoglobulin-like receptors (LILRs)
45 (classical and non-classical genes). The classical genes are conventionally known to be highly
46 polymorphic, have an excess of missense variants, and even share alleles across species, all in-
47 dicative of balancing selection at the allele level (*Maccari et al. 2017, 2020; Robinson et al. 2024;*
48 *Hughes and Nei 1988, 1989b; Arden and Klein 1982; Mayer et al. 1988*). In addition to variation
49 within individual genes, the region is also significantly structurally divergent across the primates
50 (*Mao et al. 2024*). Balancing selection is evident at the haplotype level as well, where haplotypes
51 with drastically different functional gene content are retained in various primate populations (*Hans*
52 *et al. 2017; de Groot et al. 2017b, 2009; Gleimer et al. 2011; Heijmans et al. 2020*). This motivates
53 the need to study the MHC holistically as a gene family. Even though species may retain different
54 sets of genes and haplotypes, related genes likely function similarly, facilitating comparisons across
55 species. Thus, by treating the genes as a related set, our understanding improves significantly com-
56 pared to considering single genes in isolation. Because gene family birth-and-death is important
57 to speciation and the MHC itself is highly relevant to organismal health, this family is an excellent
58 case for studying gene family evolutionary dynamics.

59 There are two classes of MHC genes within the greater family (Class I and Class II), and each
60 class contains two functionally distinct types of genes: classical and non-classical. Classical
61 MHC molecules perform antigen presentation to T cells—a key part of adaptive immunity—while
62 non-classical molecules have niche immune roles. The classical Class I molecules are generally
63 highly polymorphic, ubiquitously expressed, and present short, intracellularly-derived peptides to
64 T cells. Many of them also serve as ligands for other types of immune cell receptors and influence
65 innate immunity (see *Appendix 1* for an overview) (*Anderson et al. 2023; Parham and Moffett*
66 *2013; Guethlein et al. 2015; Hans et al. 2017; Wroblewski et al. 2019*). The non-classical Class I
67 molecules have limited polymorphism, restricted expression, and perform specific tasks such as
68 mediating maternal-fetal interaction and monitoring levels of MHC synthesis. In humans, the clas-
69 sical Class I genes are HLA-A, -B, and -C, and the non-classical Class I genes are HLA-E, -F, and -G
70 (*Heijmans et al. 2020*). In contrast, the classical Class II molecules are expressed only on profes-
71 sional antigen-presenting cell types and present longer, extracellularly-derived peptides to T cells
72 (*Gfeller and Bassani-Sternberg 2018; Heijmans et al. 2020; Neeffjes et al. 2011*). The non-classical
73 Class II molecules assist with loading peptides onto the classical Class II molecules before their
74 transport to the cell surface (*Dijkstra and Yamaguchi 2019; Neeffjes et al. 2011*). In humans, HLA-
75 DP, -DQ, and -DR are the classical Class II molecules, and HLA-DM and -DO are the non-classical
76 molecules (see *Appendix 3* for more detail on all of these genes).

77 However, the landscape of MHC genes differs even across closely-related species. Over evo-
78 lutionary time, the Class I gene subfamily has been extraordinarily plastic, having undergone re-
79 peated expansions, neofunctionalizations, and losses (*Hans et al. 2017; Wilming et al. 2013; Ot-*
80 *ting et al. 2020; Heijmans et al. 2020*). Convergent evolution has also occurred; in different pri-
81 mate lineages, the same gene may be inactivated, acquire a new function, or even evolve similar
82 splice variants (*Hans et al. 2017; Wilming et al. 2013; Otting et al. 2020; Heijmans et al. 2020;*
83 *Walter 2020*). As a result, it is often difficult to detect orthologous relationships in Class I even
84 within the primates (*Hughes and Nei 1989a; Piontkivska and Nei 2003; Go et al. 2003; Flügge*
85 *et al. 2002; de Groot et al. 2020*). Studies that focus only on the highly-polymorphic binding-site-
86 encoding exons are complicated by these phenomena, necessitating a more comprehensive look
87 into MHC evolution across exons and species groups.

88 In contrast to Class I, the Class II region has been largely stable across the primates, but gene
89 content still varies in other species. For example, the pig has lost the MHC-DP genes while expand-
90 ing the number of MHC-DR genes, and the cat has lost both the MHC-DQ and -DP genes, relying

91 entirely on MHC-DR (*Hammer et al. 2020; Okano et al. 2020*). The use of the different Class II
92 molecules appears to be fluid, at least over longer timescales, motivating the need to fill in the
93 gaps in knowledge in the primate tree.

94 Due to the large volume of existing MHC literature, results are scattered across hundreds of
95 papers, each presenting findings from a limited number of species or genes. Thus, we first per-
96 formed an extensive literature review to identify the genes and haplotypes known to be present
97 in different primate species. We present a detailed summary of these genes and their functions
98 in **Appendix 3**. We also performed a *BLAST* search using a custom IPD-based MHC allele database
99 against several available reference genomes to discover which genes were present on various pri-
100 mate reference haplotypes (**Figure 1—figure Supplement 2** and **Figure 2—figure Supplement 2**).
101 Our *BLAST* search and our search of NCBI RefSeq confirmed the presence of various genes in sev-
102 eral species for the first time. **Figure 1** and **Figure 2** show the landscape of MHC genes present
103 in different primate species for Class I and Class II, respectively. The inclusion of sequences from
104 dozens of new species across all genes and the often-ignored pseudogenes helps us paint a more
105 detailed picture of MHC evolution in the primates.

106 In this work, we present a large set of densely-sampled Bayesian phylogenetic trees using se-
107 quences from a comprehensive set of MHC genes across dozens of primate species. These trees
108 permit us to explore the overall evolution of the gene family and relationships between genes, as
109 well as trace particular allelic lineages over time. Across the trees, we see examples of rapid gene
110 turnover over just a few million years, evidence for long-term balancing selection retaining allelic
111 lineages, and slowly-evolving genes where orthology is retained for long time periods. In this pa-
112 per, we describe broad-scale differences between classes and discuss some specific results about
113 the relationships between genes. In a companion paper (*Fortier and Pritchard 2024*), we explore
114 the patterns of polymorphism *within* individual genes, finding evidence for deep trans-species poly-
115 morphism at multiple genes.

116 Results

117 Data

118 We collected MHC nucleotide sequences for all genes from the IPD-MHC/HLA database, a large
119 repository for MHC alleles from humans, non-human primates, and other vertebrates (*Maccari*
120 *et al. 2017, 2020; Robinson et al. 2024*). Although extensive, this database includes few or no
121 sequences from several key lineages including the gibbon, tarsier, and lemur. Thus, we supple-
122 mented our set of alleles using sequences from NCBI RefSeq (see asterisks in **Figure 1** and **Figure 2**).
123 Because the MHC genes make up an evolutionarily related family, they can all be aligned. Using
124 *MUSCLE* (*Edgar 2004*), we aligned all Class I sequences together, all Class IIA sequences together,
125 and all Class IIB sequences together. We then constructed trees for various subsets of these se-
126 quences using *BEAST2*, a Bayesian MCMC phylogenetic inference method (see Methods for more
127 detail) (*Bouckaert et al. 2014, 2019*). One major advantage of *BEAST2* over less tuneable methods
128 is that it can allow evolutionary rates to vary across sites, which is important for genes such as
129 these which experience rapid evolution in functional regions (*Wu et al. 2013*). We also considered
130 each exon separately to minimize the impact of recombination as well as to compare and contrast
131 the binding-site-encoding exons with non-binding-site-encoding exons.

132 Here, we present these densely-sampled Bayesian phylogenetic trees which include sequences
133 from 106 species and dozens of MHC genes. In this paper, we focus on the Class I, Class IIA, and
134 Class IIB multi-gene trees and discuss overall relationships between genes. Our companion paper
135 (*Fortier and Pritchard 2024*) explores individual clades/gene groups within these multi-gene trees
136 to understand allele relationships and assess support for trans-species polymorphism.

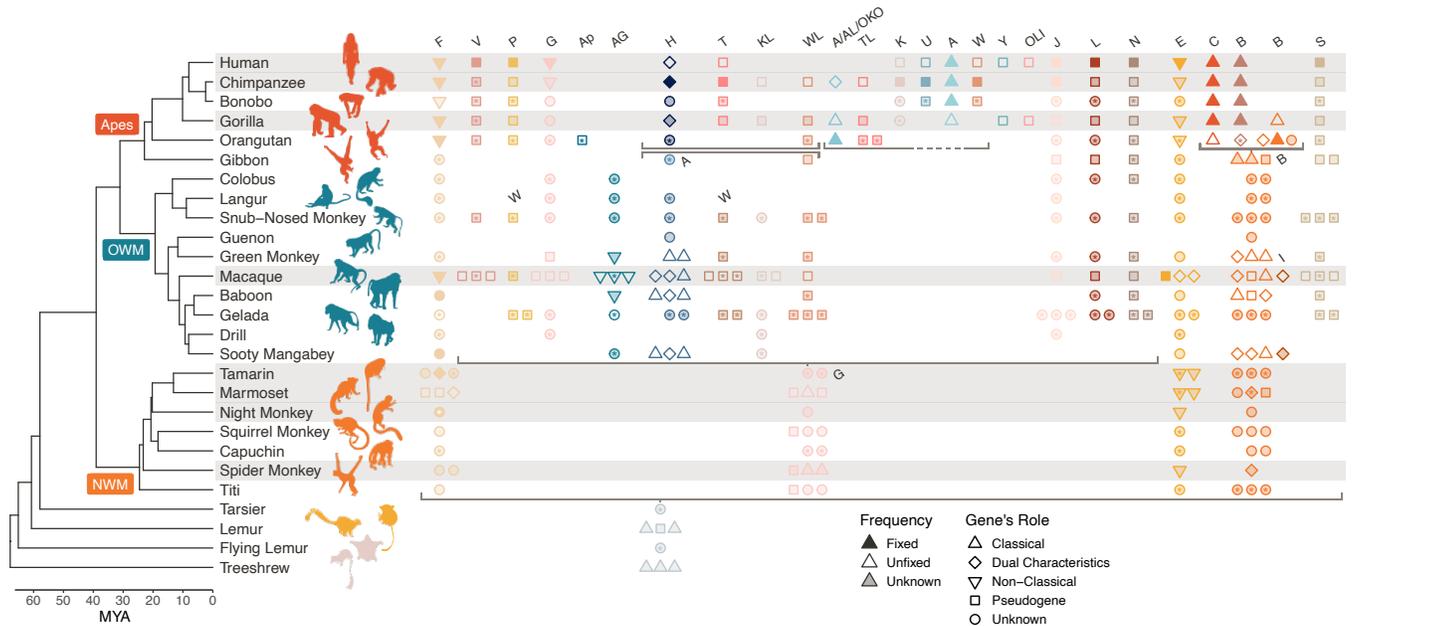


Figure 1. Class I MHC genes present in different species. The primate evolutionary tree (*Kuderna et al. 2023*) is shown on the left hand side (nonprimate icons are shown in beige). The MHC region has been well characterized in only a handful of species; the rows corresponding to these species are highlighted in gray. Species that are not highlighted have partially characterized or completely uncharacterized MHC regions. Asterisks indicate new information provided by the present study, typically discovery of a gene's presence in a species. Each column/color indicates an orthologous group of genes, labeled at the top and ordered as they are in the human genome (note that not all genes appear on every haplotype). A point indicates that a given gene is present in a given species; when a species has 3 or more paralogs of a given gene, only 3 points are shown for visualization purposes. Filled points indicate that the gene is fixed in that species, outlined points indicate that the gene is unfixed, and semi-transparent points indicate that the gene's fixedness is not known. The shape of the point indicates the gene's role, either a pseudogene, classical MHC gene, non-classical MHC gene, a gene that shares both features (dual characteristics), or unknown. The horizontal gray brackets indicate a breakdown of 1:1 orthology, where genes below the bracket are orthologous to 2 or more separate loci above the bracket. The set of two adjacent gray brackets in the top center of the figure show a block duplication. Gene labels in the middle of the plot (W , A , G , B , and I) clarify genes that are named differently in different species. OWM, Old-World Monkeys; NWM, New World Monkeys.

Figure 1—source data 1. References for *Figure 1*.

Figure 1—figure supplement 1. A version of *Figure 1* without asterisks.

Figure 1—figure supplement 2. BLAST hits for MHC Class I genes in various reference genomes.

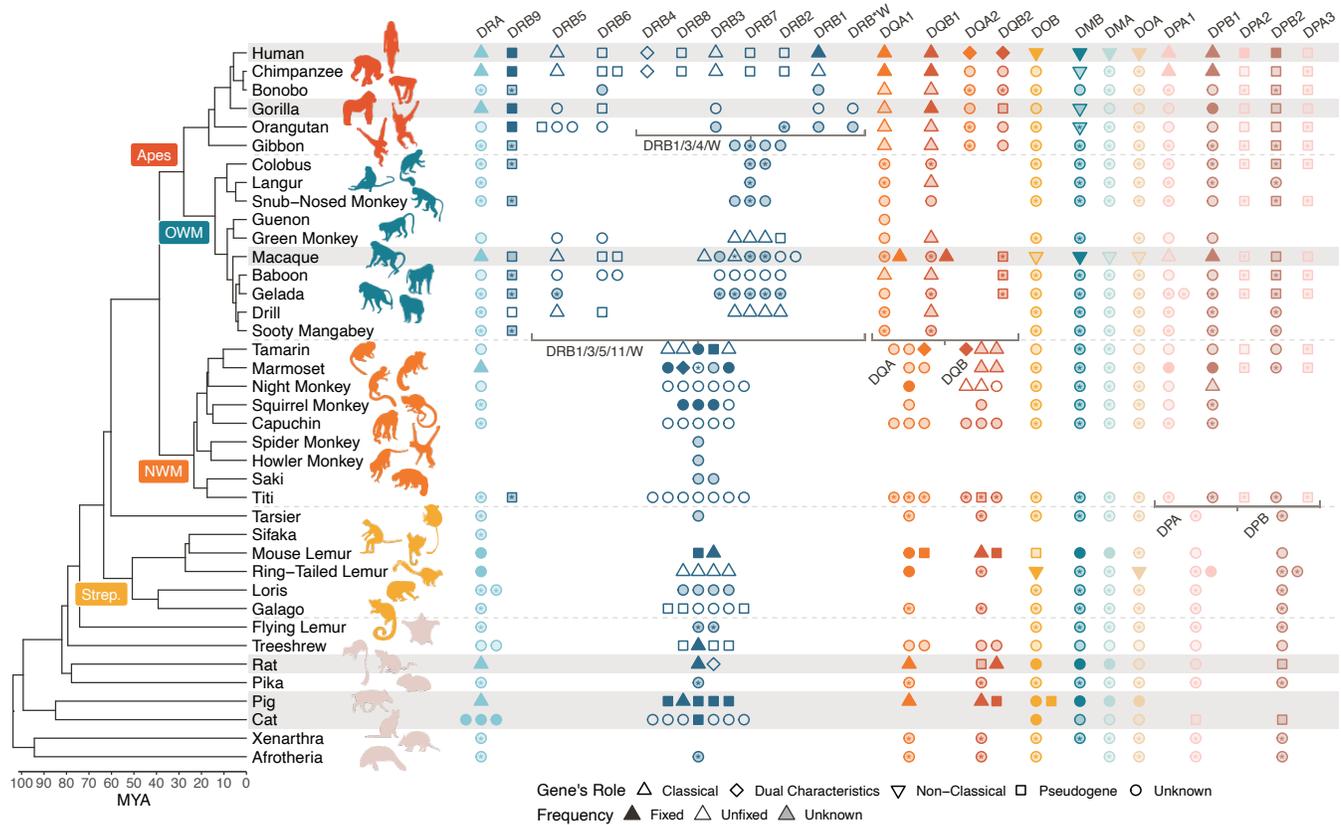


Figure 2. Class II MHC genes present in different species. The mammalian evolutionary tree is shown on the left hand side, with an emphasis on the primates (Foley et al. 2023; Kuderna et al. 2023). The rest of the figure design follows that of Figure 1, except that we did not need to limit the number of points shown per locus/species due to space constraints. OWM, Old-World Monkeys; NWM, New World Monkeys; Strep., Strepsirrhini.

Figure 2—source data 1. References for Figure 2.

Figure 2—figure supplement 1. A version of Figure 2 without asterisks.

Figure 2—figure supplement 2. BLAST hits for MHC Class II genes in various reference genomes.

137 **The MHC Across the Primates**

138 The MHC is a particularly dynamic example of a gene family due to intense selective pressure driven
139 by host-pathogen co-evolution (*Radwan et al. 2020; Ebert and Fields 2020*). Within the family,
140 genes have duplicated, changed function, and been lost many times in different lineages. As a
141 result, even closely-related species can have different sets of MHC genes. Thus, while the MHC has
142 been extensively studied in humans, there is a limit to how much we can learn from a single species.
143 Leveraging information from other species helps us understand the evolution of the entire family
144 and provides key context as to how it currently operates in humans (*Thornton and Desalle 2000;*
145 *Adams and Parham 2001b*). In *Figure 1* and *Figure 2*, we compare the genes present in different
146 species. In both, each column represents an orthologous gene, while the left-hand-side shows the
147 evolutionary tree for primates and our closest non-primate relatives (*Kuderna et al. 2023; Foley*
148 *et al. 2023*). Humans are part of the ape clade (red label), which is most closely related to the
149 old-world monkeys (OWM; blue label). Next, the ape/OWM clade is most closely related to the
150 new-world monkeys (NWM; orange label), and the ape/OWM/NWM clade is collectively known as
151 the *Simiiformes*. Only species with rows highlighted in gray have had their MHC regions extensively
152 studied. Gene presence in each species is indicated by points in each column, and the points also
153 indicate the function of the gene and whether it is fixed in the species. Points with an asterisk
154 indicate contributions from this work.

155 *Figure 1* shows that not all Class I genes are shared by apes and OWM, and much fewer are
156 shared between apes/OWM and NWM. Genes have also been differently expanded in different
157 lineages. While humans and most other apes have a single copy of each gene, the OWM and NWM
158 have multiple copies of nearly all genes. Additionally, many genes exhibit functional plasticity;
159 for example, MHC-G is a non-classical gene in the apes and a pseudogene in the OWM (it is not
160 1:1 orthologous to NWM MHC-G). The differences between even closely-related primate groups
161 indicate that the Class I region is evolving very rapidly.

162 In contrast, the Class II genes are more stable, as the same genes can be found in even distantly-
163 related mammals (*Figure 2*). The notable exception to this pattern is the MHC-DRB group of genes,
164 indicated by dark blue points in the middle of *Figure 2*. While some of the individual MHC-DRB
165 genes are orthologous between apes and OWM, indicated by points in the same column, others
166 are limited to the apes alone. Furthermore, no individual MHC-DRB genes (with the possible ex-
167 ception of MHC-DRB9) are shared between apes/OWM and NWM, pointing to their extremely rapid
168 evolution. While the other genes have been relatively stable, there have been expansions in cer-
169 tain lineages, such as separate duplications of the MHC-DQA and -DQB genes in apes/OWM, NWM,
170 and mouse lemur. Thus, both of the MHC Class I and Class I gene subfamilies appear to be subject
171 to birth-and-death evolution, with Class I and MHC-DRB undergoing the process more rapidly than
172 the rest of Class II.

173 **Evolution of a Gene Family**

174 We performed phylogenetic inference using *BEAST2* on our aligned MHC allele sequences collected
175 from NCBI RefSeq and the IPD-MHC database. *BEAST2* is a Bayesian method, meaning the set of
176 trees it produces represents the posterior space of trees (*Bouckaert et al. 2019*). For visualization
177 purposes, we collapsed the space of trees into a single summary tree that maximizes the sum of
178 posterior clade probabilities (*BEA 2024*). In each tree, the tips represent sequences, either named
179 with their RefSeq identifier or with standard allele nomenclature (see *Appendix 2*). The summary
180 tree for Class I is shown in *Figure 3*, while the summary trees for Class IIA and Class IIB are shown
181 in *Figure 4*.

182 *Figure 3A* shows the Class I multi-gene tree using sequences from exon 4, a non-peptide-binding-
183 region-encoding (non-PBR) exon equal in size to each of the peptide-binding-region-encoding (PBR)
184 exons 2 and 3. This exon is the least likely to be affected by convergent evolution, making its tree's
185 structure easier to interpret. This tree—which contains hundreds of tips—has been further simpli-
186 fied by collapsing clades of related tips, although two fully-expanded clades are shown in panels

187 B and C. Sequences do not always assort by locus. For example, ape MHC-J is separated from
188 OWM MHC-J, which is more closely related to ape/OWM MHC-G. Meanwhile, NWM MHC-G does
189 not group with ape/OWM MHC-G, instead falling outside of the clade containing ape/OWM MHC-A,
190 -G, -J and -K. This supports the fact that the NWM MHC-G genes are broadly orthologous to a large
191 group of genes which expanded within the ape/OWM lineage.

192 However, some clades/genes do behave in the expected fashion; that is, with their trees match-
193 ing the overall species tree. One such gene is non-classical MHC-F, shown in **Figure 3B**. Although the
194 gene has duplicated in the common marmoset (Caja-F), this subtree closely matches the species
195 tree shown in the upper right. This indicates that MHC-F is orthologous across apes, OWM, and
196 NWM. Orthology between apes and OWM is also observed for pseudogenes MHC-L, -K, -J, and
197 -V and non-classical MHC-E and -G (**Figure 3—figure Supplement 2** and **Figure 5—figure Supple-
198 ment 1**). For the other NWM genes, orthology with apes/OWM is less clear.

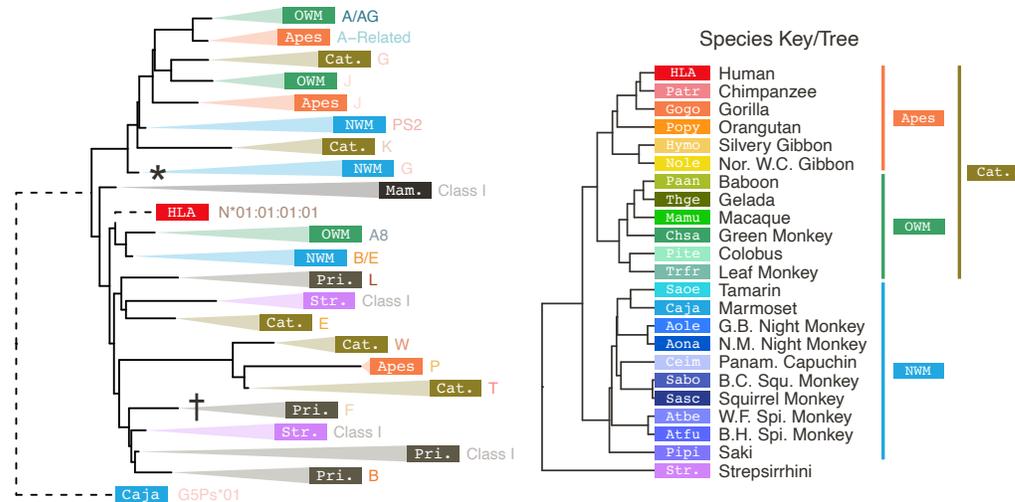
199 Other genes do not at all follow the species tree, such as NWM MHC-G. This gene group is
200 broadly orthologous to a large set of ape/OWM genes and pseudogenes, as its ancestor expanded
201 independently in both lineages. In NWM, the many functional MHC-G genes are classical, and
202 there are also a large number of MHC-G-related pseudogenes. Shown in **Figure 3C**, NWM MHC-G
203 sequences do not always group by species (colored box with abbreviation), instead forming mixed
204 clades. Thus, while some duplications appear to have occurred prior to speciation events, others
205 are species-specific. Similar expansions are seen among the MHC-A and -B genes of the OWM and
206 the MHC-B genes of the NWM (**Figure 3—figure Supplement 2**, **Figure 3—figure Supplement 3**, and
207 **Figure 3—figure Supplement 4**).

208 **Figure 4** shows summary trees for exon 3 (non-PBR) for the Class IIA and IIB sequence sets. In
209 the Class II genes, exon 3 does not encode the binding site, but is similar in size to binding-site-
210 encoding exon 2. In contrast to Class I (**Figure 3**), Class II sequences group entirely and unambigu-
211 ously by gene, shown by the collapsed trees in **Figure 4A**. However, the subtrees for each gene
212 exhibit varying patterns. As with Class I, non-classical genes tend to evolve in a typical fashion
213 with sequences assorting according to the species tree. This is clearly the case for non-classical
214 MHC-DMA, -DMB, -DOA, and -DOB (**Figure 4—figure Supplement 1** and **Figure 4—figure Supple-
215 ment 2**). Classical genes MHC-DRA and -DPA also follow this pattern (**Figure 4B** and **Figure 4—figure
216 Supplement 1**). However, the other classical genes' subtrees look very different from the species
217 tree.

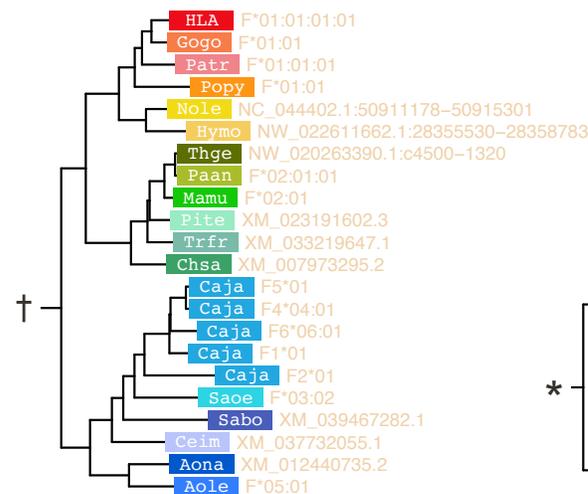
218 There are several reasons why these MHC gene trees can differ from the overall species tree. In-
219 complete lineage sorting can happen purely by chance, especially if species have recently diverged.
220 However, balancing selection can cause alleles to be longer-lived, resulting in incomplete lineage
221 sorting even among deeply-diverged species; this is called trans-species polymorphism (TSP). **Fig-
222 ure 4C** illustrates this phenomenon for MHC-DPB. Within the OWM clade (shades of green), se-
223 quences group by allelic lineage rather than by species. For example, crab-eating macaque allele
224 Mafa-DRB1*09:02:01:01 groups with green monkey allele Chsa-DPB1*09:01 (both members of the
225 DRB1*09 lineage) rather than with the other macaque alleles (Mane-, Mamu-, Math-, and Malo-),
226 despite the fact that these species are 15 million years separated from each other (**Kuderna et al.
227 2023**). We see this pattern in many Class II genes and some Class I genes (**Figure 3—figure Supple-
228 ment 3**, **Figure 3—figure Supplement 4**, **Figure 3—figure Supplement 5**, **Figure 4—figure Supple-
229 ment 8**, **Figure 4—figure Supplement 9**, and **Figure 4—figure Supplement 10**). In our companion
230 paper, we explore each of these genes further and evaluate the strength of support for TSP in each
231 gene.

232 Another way to obtain discordant trees is in the case of recent expansions of genes. Such expan-
233 sions make it difficult to assign sequences to loci, resulting in clades where sequences (ostensibly
234 from the same locus) do not group by species. An example of this is shown in **Figure 3C** for the
235 NWM Class I gene MHC-G. The Class II MHC-DRB genes have also expanded, although locus as-
236 signments are somewhat clearer. **Figure 4D** shows the Class II subtree for MHC-DRB, where ape
237 sequences (red/orange boxes) are interspersed with OWM sequences (green boxes). The MHC-

A. Sequences do not always group by gene name in the Class I multi-gene tree.



B. All MHC-F sequences group together and generally assort according to the species tree.



C. MHC-G sequences do not group together; the genes have expanded and diversified separately in the NWM.

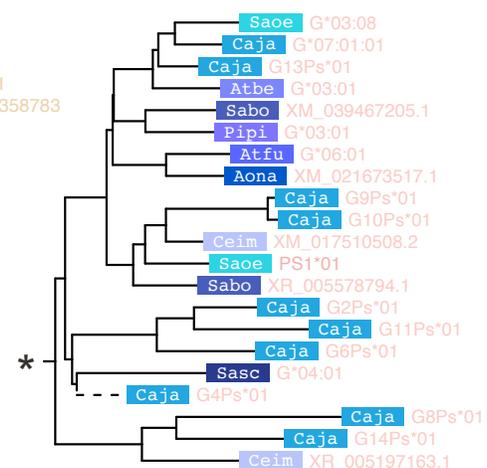


Figure 3. The Class I exon 4 multi-gene BEAST2 tree. The Class I multi-gene tree was constructed using exon 4 (non-PBR) sequences from Class I genes spanning the primates. **A)** For the purposes of visualization, each clade in the multi-gene tree is collapsed and labeled according to the main species group and gene content of the clade. The white labels on colored rectangles indicate the species group of origin, while the colored text to the right of each rectangle indicates the gene name. The abbreviations are defined in the species key to the right. **B)** The expanded MHC-F clade (corresponding to the clade in panel A marked by a †). **C)** The expanded NWM MHC-G clade (marked by a * in panel A). In panels B and C, each tip represents a sequence and is labeled with the species of origin (white label on colored rectangle) and the sequence ID or allele name (colored text to the right of each rectangle; see **Appendix 2**). The species key is on the right hand side of panel A. Dashed branches have been shrunk to 10% of their original length (to clarify detail in the rest of the tree at this scale). OWM: old-world monkeys; NWM: new-world monkeys; Cat.: Catarrhini—apes and OWM; Pri.: Primates—apes, OWM and NWM; Mam.: mammals—primates and other outgroup mammals.

Figure 3—figure supplement 1. Color and abbreviation key/tree for all species included in this study.

Figure 3—figure supplement 2. Full, non-collapsed versions of the Class I multi-gene BEAST2 trees.

Figure 3—figure supplement 3. Full, non-collapsed versions of the MHC-A-related focused BEAST2 trees.

Figure 3—figure supplement 4. Full, non-collapsed versions of the MHC-B-related focused BEAST2 trees.

Figure 3—figure supplement 5. Full, non-collapsed versions of the MHC-C-related focused BEAST2 trees.

Figure 3—figure supplement 6. Full, non-collapsed versions of the MHC-E-related focused BEAST2 trees.

Figure 3—figure supplement 7. Full, non-collapsed versions of the MHC-F-related focused BEAST2 trees.

Figure 3—figure supplement 8. Full, non-collapsed versions of the MHC-G-related focused BEAST2 trees.

Figure 3—source data 1. GENECONV results for the Class I focused alignments.

238 DRB genes are usually assigned to specific named loci, but in this tree only MHC-DRB5 sequences
239 group by named locus (the collapsed ape/OWM MHC-DRB5 clade is about 1/3 from the bottom of the
240 tree). The failure of the other named loci to group together indicates a lack of 1:1 orthology
241 between apes, OWM, and NWM for these genes and thus rapid evolution. This makes the MHC-
242 DRB genes unique among the Class II genes. We created a focused tree with more sequences in
243 order to explore the evolution of the MHC-DRB genes further, which is presented in a later section
244 (**Figure 4—figure Supplement 8**).

245 Gene conversion is a third way that gene trees might differ from the overall species tree. Gene
246 conversion is the unidirectional copying of a sequence onto a similar sequence (usually another
247 allele or a related locus), which results in two sequences being unusually similar even if they are
248 not related by descent. We consider this possibility in the next section.

249 **Detection of Gene Conversion**

250 Because the MHC contains many related genes in close proximity, gene conversion—the unidirec-
251 tional exchange of sequence between two similar sequences—can occur (**Chen et al. 2007**). We
252 used the program *GENECONV* (**Sawyer 1999**) to infer pairs of sequences of which one has likely
253 been converted by the other (**Figure 3—source data 1** and **Figure 4—source data 1**). We recov-
254 ered known gene conversion events, such as between human allelic lineages HLA-B*38 and HLA-
255 B*67:02, as well as novel events, such as between gorilla allelic lineages Gogo-B*01 and Gogo-B*03
256 and ape/OWM lineages MHC-DQA1*01 and MHC-DQA1*05.

257 However, most of the tracts we detected involved many different groups of species but im-
258 plicated the same pair of loci. We interpreted these as gene conversion events that must have
259 happened a long time ago in the early history of the two genes, and they are likely to blame for
260 the topological differences from exon to exon among the trees. For example, in exon 2, the Class
261 I pseudogene MHC-K groups with MHC-G, while in exon 3 it groups with MHC-F, and in exon 4 it
262 groups outside of MHC-G, -J, and -A (**Figure 3**). The uncertain early branching structure we observe
263 in our trees may be due to these ancient gene conversion events.

264 **The Importance of the Pseudogenization Process**

265 Gene birth-and-death drives the evolution of a gene family as a whole. The death can include the
266 deletion of all or part of a gene from the genome or pseudogenization by means of inactivating
267 mutations, which can leave gene remnants behind. In Class I, we find many pseudogenes that
268 have been produced in this process; while countless more have undoubtedly already been deleted
269 from primate genomes, many full-length and fragment pseudogenes still remain. Although non-
270 functional, these sequences provide insight into the granular process of birth-and-death as well as
271 improve tree inference.

272 Full Class I haplotypes including the pseudogenes are known only for human, chimpanzee, go-
273 rilla, and macaque, and even so we do not have sequences for *all* the balanced haplotypes in each
274 species (**Anzai et al. 2003; Wilming et al. 2013; Shiina et al. 2017; Karl et al. 2023**). From these
275 studies, we know that few functional Class I genes are shared by apes/OWMs and NWMs, and
276 so far no shared pseudogenes have been found (**Lugo and Cadavid 2015; Kono et al. 2014; Ca-**
277 **David et al. 1996; Maccari et al. 2017, 2020**). Therefore, the Class I genes in the two groups have
278 been generated by a largely separate series of duplications, neofunctionalizations, and losses. This
279 means that turnover has occurred on a relatively short timescale, and understanding the pseudo-
280 genes within the apes and OWM can thus shed light on the evolution of the region more granularly.
281 These ancient remnants could provide clues as to when genes or whole blocks were duplicated,
282 which regions are more prone to duplication, and how the MHC may have functioned in ancestral
283 species.

284 The Class I MHC region is further divided into three polymorphic blocks— α , κ , and β —that each
285 contain MHC genes but are separated by well-conserved non-MHC genes. The majority of the Class
286 I genes are located in the α -block, which in humans includes 12 MHC genes and pseudogenes

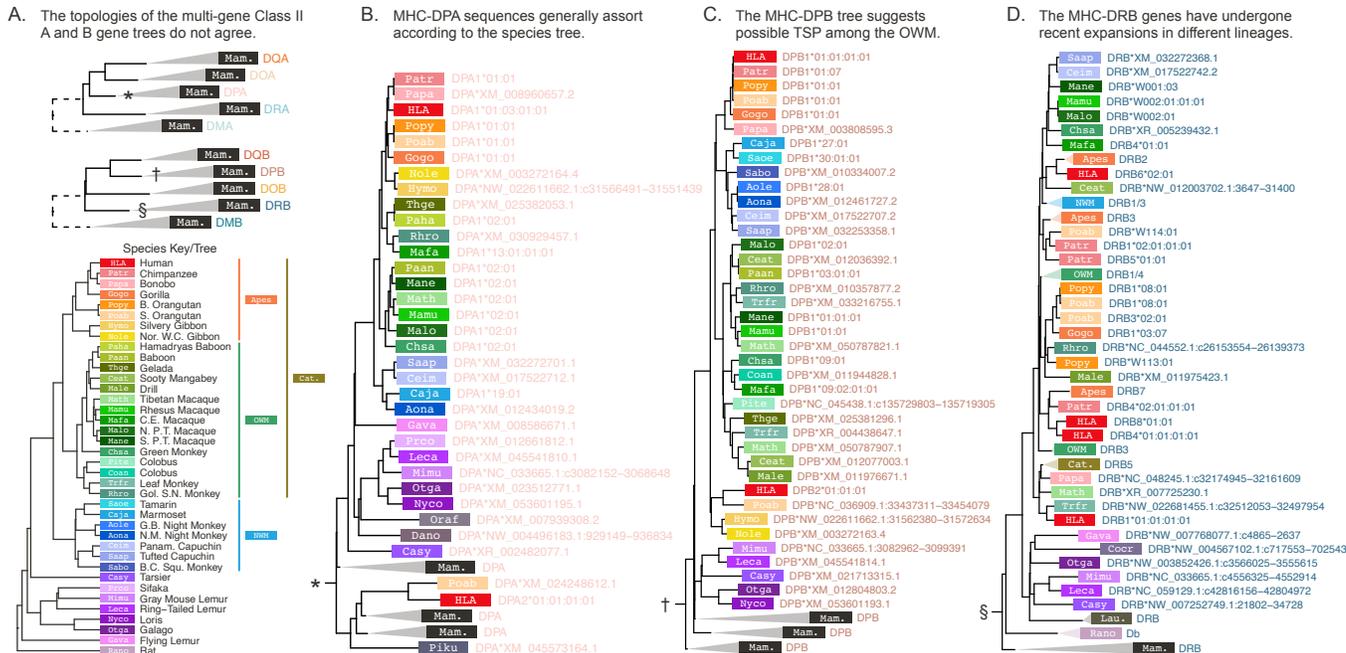


Figure 4. The Class II exon 3 multi-gene *BEAST2* trees. The trees were constructed using all Class IIA and all Class IIB exon 3 (non-PBR) sequences across all available species. The design of this figure follows *Figure 3*. **A)** The top tree shows the collapsed Class IIA gene tree, while the bottom tree shows the collapsed Class IIB gene tree. In this case, all collapsed clades are labeled with Mam. for mammals, because sequences from primates and mammal outgroups assort together by gene. **B)** The expanded MHC-DPA clade (corresponding to the clade in panel A marked by a *) **C)** The expanded MHC-DPB clade (marked by a † in panel A). **D)** The expanded MHC-DRB clade (marked by a § in panel A). OWM: old-world monkeys; NWM: new-world monkeys; Cat.: Catarrhini—apes and OWM; Mam.: mammals—primates and other outgroup mammals.

Figure 4—figure supplement 1. Full, non-collapsed versions of the Class IIA multi-gene *BEAST2* trees.

Figure 4—figure supplement 2. Full, non-collapsed versions of the Class IIB multi-gene *BEAST2* trees.

Figure 4—figure supplement 3. Full, non-collapsed versions of the MHC-DRA focused *BEAST2* trees.

Figure 4—figure supplement 4. Full, non-collapsed versions of the MHC-DQA focused *BEAST2* trees.

Figure 4—figure supplement 5. Full, non-collapsed versions of the MHC-DPA focused *BEAST2* trees.

Figure 4—figure supplement 6. Full, non-collapsed versions of the MHC-DMA focused *BEAST2* trees.

Figure 4—figure supplement 7. Full, non-collapsed versions of the MHC-DOA focused *BEAST2* trees.

Figure 4—figure supplement 8. Full, non-collapsed versions of the MHC-DRB focused *BEAST2* trees.

Figure 4—figure supplement 9. Full, non-collapsed versions of the MHC-DQB focused *BEAST2* trees.

Figure 4—figure supplement 10. Full, non-collapsed versions of the MHC-DPB focused *BEAST2* trees.

Figure 4—figure supplement 11. Full, non-collapsed versions of the MHC-DMB focused *BEAST2* trees.

Figure 4—figure supplement 12. Full, non-collapsed versions of the MHC-DOB focused *BEAST2* trees.

Figure 4—source data 1. *GENECONV* results for the Class II focused alignments.

287 (*Shiina et al. 2017*). The α -block also contains a large number of repetitive elements and gene
288 fragments belonging to other gene families, and their specific repeating pattern in humans led to
289 the conclusion that the region was formed by successive block duplications (*Shiina et al. 1999*).
290 Later, comparison of macaque and chimpanzee α -block haplotypes with the sequenced human
291 haplotype bolstered this hypothesis, although the proposed series of events is not always consistent
292 with phylogenetic data (*Kulski et al. 2005, 2004; Geraghty et al. 1992; Hughes 1995; Messer*
293 *et al. 1992; Alexandrov et al. 2023; Gleimer et al. 2011*). Improving existing theories about the
294 evolution of this block is useful for disentangling the global pattern of MHC evolution from locus-
295 and gene-specific influences. This could help us understand how selection on specific genes has
296 affected entire linked regions. We therefore created an α -block-focused tree involving sequences
297 from more species than ever before in order to strengthen and update previous hypotheses about
298 the evolution of the block, shown in **Figure 5**.

299 **Figure 5A** shows the α -block-focused tree for exon 3, with an expanded MHC-V clade. MHC-V is
300 a fragment pseudogene containing exons 1-3 which is located near MHC-F in the α -block. Previous
301 work disagrees on the age of this fragment, with some suggesting it was fixed relatively early while
302 others claiming it arose from one of the more recent block duplications (*Shiina et al. 1999; Kulski*
303 *et al. 2004, 2005*). Our tree groups ape and OWM MHC-V together and places them as an out-
304 group to all of the classical and non-classical genes, including those of the NWM. Thus, the MHC-V
305 fragment may be an ancient remnant of one of the ancestral Class I genes. We also dispute the
306 hypothesis that MHC-V and -P (a 3'-end fragment) are since-separated pieces of the same original
307 gene (*Horton et al. 2008*), as we found that both contain exon 3 and their exon 3 sequences clearly
308 do not group together.

309 In **Figure 5B**, we zoom in on a different clade within the exon 3 tree, corresponding to the
310 asterisk in panel A. Here we focus on MHC-U, a previously-uncharacterized fragment pseudogene
311 containing only exon 3. Our tree groups it with a clade of human, chimpanzee, and bonobo MHC-A,
312 suggesting it duplicated from MHC-A in the ancestor of these three species. However, it is present
313 on both chimpanzee haplotypes and nearly all human haplotypes. Since these haplotypes likely
314 diverged earlier in the ancestor of human and gorilla, we presume that MHC-U will be found in the
315 gorilla when more haplotypes are sequenced. Ours is the first work to show that MHC-U is actually
316 an MHC-A-related gene fragment.

317 The exon 4 α -block-focused tree is shown in **Figure 5C**. Here, we expand the clade for MHC-
318 K, a full-length pseudogene present in apes and OWM. In humans, only MHC-K is present, but
319 on some chimpanzee haplotypes both MHC-K and its duplicate MHC-KL are present. In gorillas,
320 haplotypes can contain either MHC-K or -KL, and in OWM there are many copies of MHC-K as
321 they are part of one of the basic block duplication units (**Figure 5—figure Supplement 2**) (*Karl*
322 *et al. 2023*). **Figure 5C** shows that MHC-K and -KL are closely related and OWM MHC-K groups
323 outside of both, indicating that the duplication (which also copied MHC-W, -A, and -T) occurred
324 after the split of apes and OWM. We did not detect MHC-K or -KL sequences in either the gibbon or
325 orangutan reference genomes during our *BLAST* search, so we cannot date this duplication event
326 more precisely. The pseudogene may have been deleted from both genomes entirely, or it may
327 be present on non-reference haplotypes. Sequencing of more haplotypes may help resolve the
328 timing of this duplication event.

329 Both the exon 3 and exon 4 trees show a clear separation between the clade of MHC-W, -WL,
330 -P, -T, -TL, and -OLI fragment pseudogenes and the rest of the genes (Figures 5A and C). On the
331 chromosome, members of these two subfamilies are interleaved throughout the α -block, suggest-
332 ing that both groups are old and a series of block duplications occurred. Previous work on human
333 sequences has shown that HLA-P, -W, -T, and -OLI are related (*Alexandrov et al. 2023; Hughes*
334 *1995; Kulski et al. 2005*). However, humans do not have orthologs of every single primate gene,
335 so utilizing other primate sequences is critical to understanding this subfamily's evolution.

336 The MHC-W/WL/P/T/TL/OLI clade, marked with a † in **Figure 5C**, is expanded in panel D. We
337 expected OWM MHC-W sequences to form a monophyletic clade either outside of all of the ape

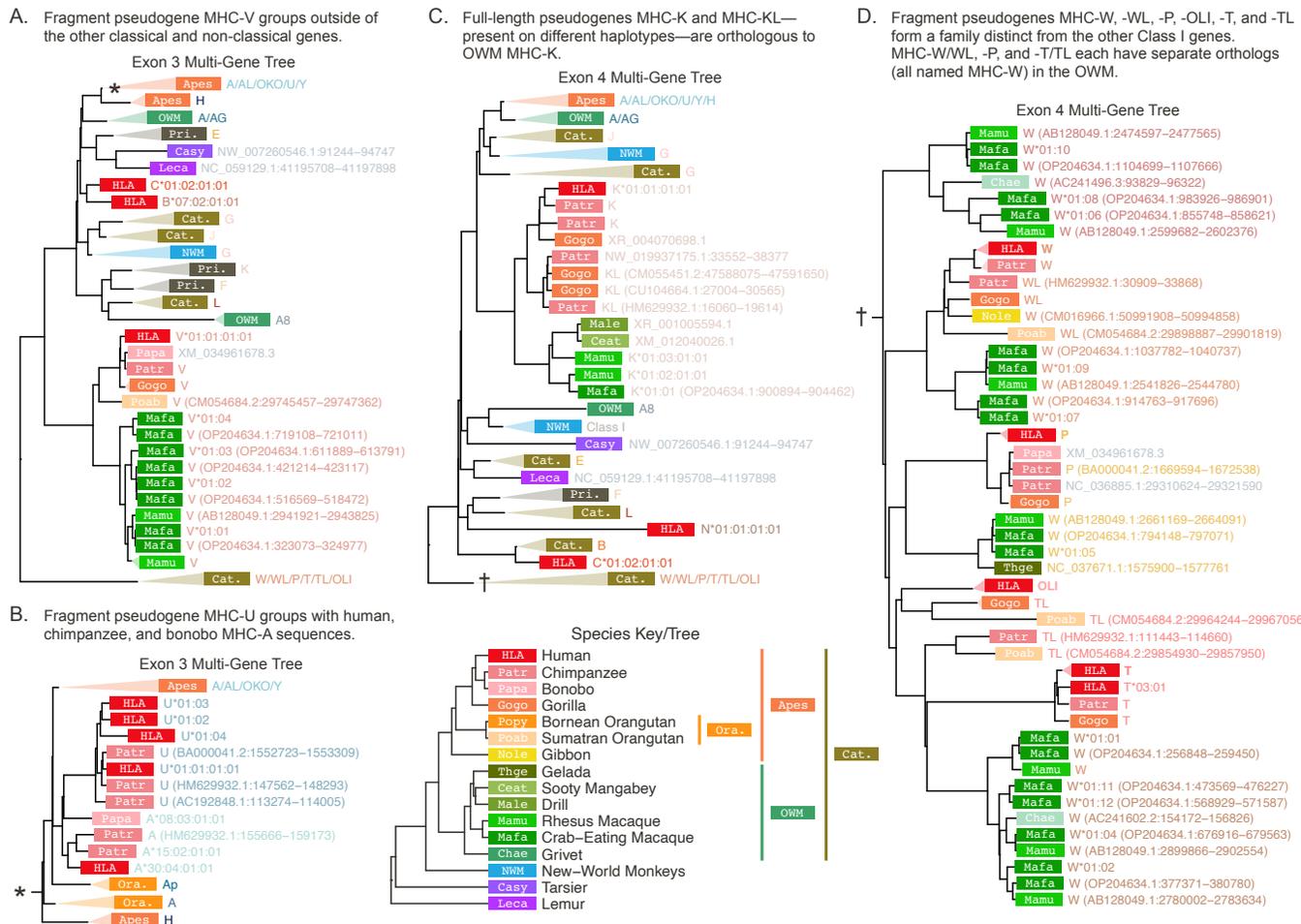


Figure 5. Class I α -block-focused multi-gene BEAST2 trees. The α -block-focused trees use the common backbone sequences as well as additional sequences from our custom BLAST search of available reference genomes. For the purposes of visualization, some clades are collapsed and labeled with the species group and gene content of the clade (colored text to the right of each rectangle). The white labels on colored rectangles indicate the species group of origin, while the colored text to the right of each rectangle indicates the gene or sequence name (see Appendix 2). The species abbreviations are defined in the species key at the bottom. **A**) Exon 3 α -block-focused BEAST2 tree with expanded MHC-V clade. **B**) The expanded MHC-A/AL/OKO/U/Y clade from the exon 3 tree (corresponding to the clade in panel A marked by a *), focusing on MHC-U. **C**) Exon 4 α -block-focused BEAST2 tree with expanded MHC-K/KL clade. **D**) The expanded MHC-W/WL/P/T/TL/OLI clade from the exon 4 tree (marked by a † in panel C). OWM: old-world monkeys; NWM: new-world monkeys; Cat.: Catarrhini—apes and OWM; Pri.: Primates—apes, OWM and NWM.

Figure 5—figure supplement 1. Full, non-collapsed versions of the Class I α -block pseudogene-focused BEAST2 trees.

Figure 5—figure supplement 2. Full, non-collapsed versions of the all-pseudogene-focused BEAST2 trees.

338 genes or with a single ape MHC gene, demonstrating orthology. Surprisingly, OWM MHC-W se-
339 quences instead formed four distinct clades, with one grouping with ape MHC-W/WL, one with ape
340 MHC-P, one with ape MHC-T/TL, and one outside of all. Furthermore, based on the alleles present,
341 each of these OWM MHC-W clades corresponds to a type of basic repeat block (as revealed by the
342 published macaque MHC haplotype) (Karl *et al.* 2023; Kulski *et al.* 2004). The correspondence
343 between the OWM clades and the genes' physical location lends further support to the hypothesis
344 that macaque haplotypes were generated by tandem duplications. Furthermore, the fact that the
345 OWM MHC-W clades each group with a different ape pseudogene suggests that there are three
346 ape/OWM orthologous groups. Thus, there must have been three distinct MHC-W-like genes in
347 the ape/OWM ancestor.

348 We also learned more about HLA-OLI, a recently-discovered MHC pseudogene found on the
349 same insertion segment that carries HLA-Y in a small fraction of the human population. Its discov-
350 erers used only human sequences, finding HLA-OLI was most similar (88%) to HLA-P (Alexandrov
351 *et al.* 2023). Our inclusion of non-human primate genes revealed that HLA-OLI is actually most
352 similar in both structure and sequence to MHC-TL, a gene not found in humans (Figure 5—figure
353 Supplement 1). Furthermore, since MHC-Y and -OLI are fully linked and are located in close prox-
354 imity, it is likely that they duplicated as a unit. MHC-Y's similarity to MHC-AL/OKO and HLA-OLI's
355 similarity to MHC-TL is consistent with them duplicating together from those genes, which are ad-
356 jacent to each other on non-human haplotypes.

357 With these findings, in addition to many other observations from our trees and results from
358 past literature (references in Figure 6—figure Supplement 1), we propose a new hypothesis for the
359 evolution of the Class I α -block. Figure 6 shows a possible evolutionary path for α -block haplotypes
360 that could have led to the currently observed haplotypes. Haplotypes found so far in each species
361 are at the bottom of the figure (with additional never-before-reported haplotypes from our BLAST
362 search shown in Figure 1—figure Supplement 2). Notably, our work has revealed that MHC-V is
363 an old fragment, three MHC-W-like genes were already established at the time of the ape/OWM
364 ancestor, MHC-U is closely related to African ape MHC-A, and MHC-OLI is closely related to MHC-
365 TL. Additionally, the OWM MHC-A fragment pseudogene is actually more similar to the ape MHC-A
366 genes than to the other OWM MHC-A genes, supporting the existence of two MHC-A-like genes in
367 the ape/OWM ancestor.

368 Evolution of the MHC-DRB Region

369 The Class I vs. Class II division reflects a major functional distinction within the MHC gene family,
370 but even within these subfamilies evolution is not homogeneous. Among the Class II genes, there
371 are few duplicated genes and generally only one way for the protein products to pair, e.g. MHC-
372 DPA1 with MHC-DPB1. However, the MHC-DR genes are a notable exception to the general pattern;
373 MHC-DRA can pair with any of the multiple functional MHC-DRB genes. In addition, MHC-DRB has
374 many more related pseudogenes compared to the rest of the Class II genes, making the region's
375 pattern of evolution more reminiscent of Class I (see Figure 2 to see the varied landscape of MHC-
376 DRB genes in different species). We explored the evolution of the MHC-DRB region in greater detail
377 by creating focused trees with a larger set of MHC-DRB sequences.

378 In exon 2, which codes for the binding site, the MHC-DRB genes group mostly by name (e.g.
379 MHC-DRB3) across apes, OWM, and NWM (Figure 4—figure Supplement 8A). The exon 2 tree con-
380 sidered alone thus suggests that the genes are orthologous across apes, OWM, and NWM—which
381 is how the genes were named in the first place. Exon 3 does not encode the binding site and
382 is less likely to be affected by convergent evolution; its tree is shown in Figure 4—figure Supple-
383 ment 8B. In this tree, all NWM sequences group together (clade with blue boxes about halfway
384 up the tree), suggesting that the genes expanded separately in NWM and apes/OWM. Additionally,
385 OWM MHC-DRB1 and -DRB3 form their own clade (green boxes near the top of the tree) and OWM
386 MHC-DRB4 sequences group outside of several ape and NWM clades. We see that only three MHC-
387 DRB genes/pseudogenes (MHC-DRB5, -DRB2/6, and -DRB9) form monophyletic clades of ape and

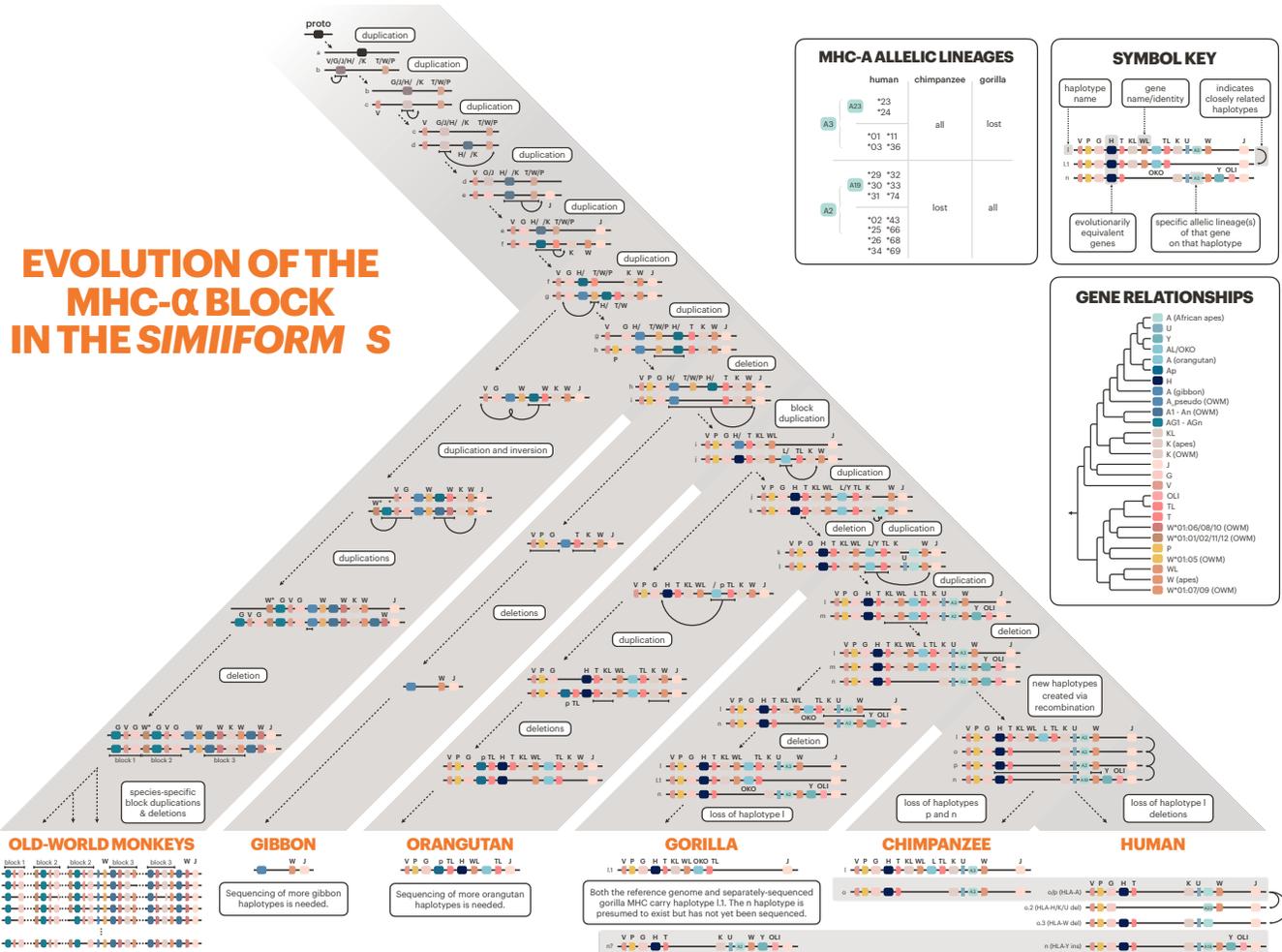


Figure 6. Evolution of the Class I α -block. The primate evolutionary tree is shown in gray (branches not to scale). The bottom of the tree shows currently known haplotypes in each species or species group. Horizontal gray bars indicate haplotypes shared among the African apes. The history of the genes/haplotypes in the α -block is overlaid on the tree, synthesizing previous work with our own observations (see Methods and Figure 8). Genes are represented by colored rectangles, while haplotypes are shown as horizontal lines containing genes. MHC-F—indicating the telomeric end of the α -block—was fixed early on and is located immediately to the left on all haplotypes shown, but is not pictured due to space constraints. Dashed arrows with descriptive labels represent evolutionary events. In the upper right, the Symbol Key explains the icons and labels. The Gene Relationships panel shows the relationships between the loci shown on the tree, without the layered complexity of haplotypes and speciation events. The MHC-A Allelic Lineages panel shows which MHC-A allele groups are present in human, chimpanzee, and gorilla.

Figure 6—figure supplement 1. A version of Figure 6 with references.

388 OWM sequences, indicating that these three are the only orthologous MHC-DRB genes. Further,
389 none are orthologous to any particular NWM gene. Thus, the longevity of individual MHC-DRB
390 genes in the primates appears to be less than 38 million years.

391 The longevity of MHC-DRB haplotypes is even shorter. Only one haplotype is shared between
392 human and chimpanzee, and none are shared with gorilla (*de Groot et al. 2009; Heijmans et al.*
393 *2020; Hans et al. 2015*). This shows that the region is evolving even more rapidly than Class I
394 (where haplotypes are shared among human, chimpanzee, and gorilla; *Figure 6*). These haplotypes,
395 combined with past literature (*Figure 7—figure Supplement 1*) and our trees, allowed us to trace
396 backward and propose a hypothesis for the evolution of the region, shown in *Figure 7*.

397 Limited data exists for the orangutan, and in our exon 3 tree (*Figure 4—figure Supplement 8B*)
398 orangutan alleles do not group definitively with any other ape lineages. Furthermore, the orangutan
399 MHC-DRB3 gene groups with orangutan MHC-DRB1, suggesting that it may not be orthologous to
400 the African ape MHC-DRB3 gene. We also found an orangutan sequence that groups with the hu-
401 man HLA-DRB2 pseudogene, suggesting that this gene has an ortholog in the orangutan. Several
402 haplotypes have been previously identified in the gibbon, but since they rely on exon 2 sequence
403 alone, it is unclear how these alleles relate to the known ape lineages (*de Groot et al. 2017a*). Anal-
404 ysis of more orangutan and gibbon haplotypes will be essential for understanding how the region
405 has evolved in the apes.

406 Overall, the MHC-DRB genes are not evolving in the same fashion as the rest of the Class II
407 genes, even though they have a shared structure and function. This peculiar case illustrates that
408 there are multiple ways to achieve a functional immune response from the same basic parts.

409 **Differences between MHC Subfamilies**

410 We explored the evolution of the Class I and Class II genes separately and noticed several differ-
411 ences between the classes. First, sequences group by gene rather than species group in the Class
412 II gene trees (*Figure 4, Figure 4—figure Supplement 1, and Figure 4—figure Supplement 2*). Our
413 inclusion of RefSeq sequences from distant groups of placental mammals confirms that most of
414 the primate Class II genes have maintained orthology at least since the ancestor of placentals, 105
415 million years ago (*Foley et al. 2023*). In contrast, our Class I trees (*Figure 3 and Figure 3—figure*
416 *Supplement 2*) showed sequences more often grouping by species group than by gene, indicating
417 that the genes turn over quickly and 1:1 orthology is often lost. Only non-classical MHC-F (and
418 possibly MHC-E) are truly orthologous among the apes, OWM, and NWM, consistent with previous
419 findings (*Piontkivska and Nei 2003; Adams and Parham 2001b; Sawai et al. 2004*). Additionally,
420 our tarsier and *Strepsirrhini* sequences group outside of all *Simiiformes* Class I sequences, setting
421 an upper bound on the maintenance of Class I orthology of 58 million years (*Kuderna et al. 2023;*
422 *Flügge et al. 2002*).

423 This turnover of genes at the MHC—rapid for Class I and slower for Class II—is generally be-
424 lieved to be due to host-pathogen co-evolution (*Radwan et al. 2020*). Although this means the
425 MHC genes are critically important for survival, no single gene is so vital that its role must be pre-
426 served. For example, in the apes the MHC-G gene is non-classical, but in the OWM it has been
427 inactivated and its role largely replaced by an MHC-A-related gene called MHC-AG (*Heijmans et al.*
428 *2020*). This process of turnover ultimately results in different sets of MHC genes being used in dif-
429 ferent lineages. For instance, separate expansions generated the classical Class I genes in NWM (all
430 called MHC-G) and the α -block genes in apes/OWM. Similarly, separate expansions generated the
431 MHC-DRB genes of the NWM and of the apes/OWM. Aside from MHC-DRB, the Class II genes have
432 been largely stable across the mammals, although we do see some lineage-specific expansions
433 and contractions (*Figure 2 and Figure 2—figure Supplement 2*).

434 Class I and Class II also differ in their degree of gene conversion. Our *GENECONV* analysis re-
435 vealed two types of gene conversion events: 1) specific, more-recent events involving paralogous
436 genes or particular allelic lineages and 2) broad-scale, very-old events involving two dissimilar loci
437 (*Figure 3—source data 1 and Figure 4—source data 1*). We discovered far more specific events

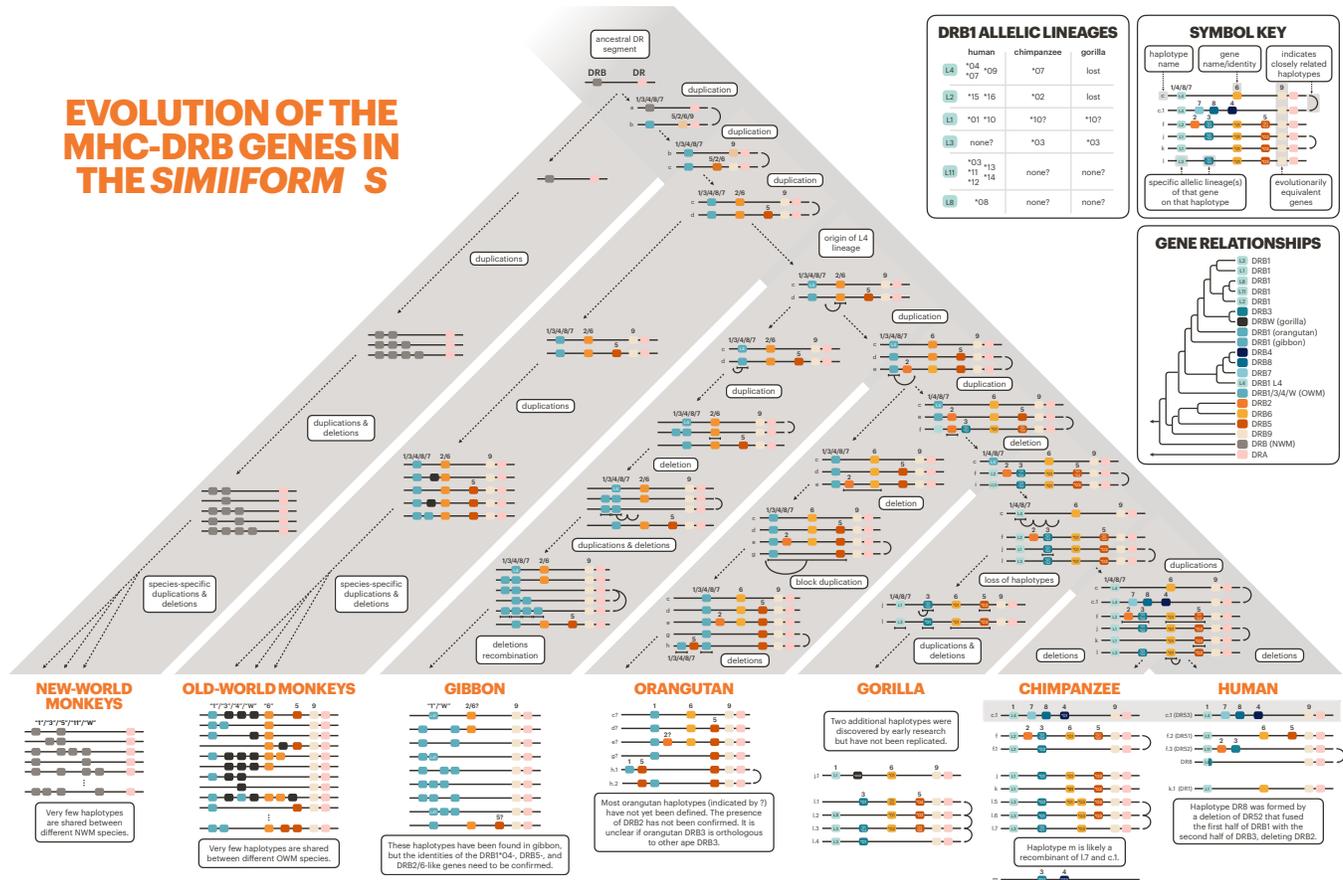


Figure 7. Evolution of MHC-DRB. The bottom of the tree shows current haplotypes in each species or species group; human, chimpanzee, gorilla, and old-world monkey haplotypes are well characterized, while orangutan, gibbon, and new-world monkey haplotypes are partially known. The history of the genes/haplotypes in the MHC-DRB region is overlaid on the tree, synthesizing previous work with our own observations (see Methods and [Figure 8](#)). The rest of the figure design follows that of [Figure 6](#).

Figure 7—figure supplement 1. A version of [Figure 7](#) with references.

438 in Class I, while broad-scale events were predominant in Class II. This could reflect the differ-
439 ent age of these gene groups: Class I genes turn over more rapidly and allelic lineages are less
440 diverged from each other, making gene conversion more likely. In contrast, Class II genes have
441 much longer-lived allelic lineages, potentially explaining why we mainly picked up older events in
442 the Class II *GENECONV* analysis.

443 Even within a class, evolutionary patterns are not homogeneous. The non-classical vs. classical
444 distinction is one functionally meaningful way to partition the genes. The classical genes perform
445 peptide presentation to T-cells, making them direct targets of host-pathogen co-evolution. In con-
446 trast, the non-classical genes are involved in innate immune surveillance or niche roles, and may
447 be less directly affected by this co-evolution. In our trees, sequences from non-classical genes of
448 both classes often group by gene with topology matching the species tree, while sequences from
449 classical genes do neither (*Figure 3—figure Supplement 2, Figure 4—figure Supplement 1, and Fig-
450 ure 4—figure Supplement 2*). This shows that classical genes experience more turnover and are
451 more often affected by long-term balancing selection or convergent evolution. Ultimately, selec-
452 tion acts upon functional differences between classical and non-classical genes in a manner that
453 is largely independent of whether they belong to Class I or Class II.

454 Discussion

455 The MHC proteins serve diverse roles in innate and adaptive immunity (*Adams and Luoma 2013*).
456 They are critically important to infection resistance, autoimmune-disease susceptibility, and organ
457 transplantation success, and can provide insight into human evolution, inform disease studies,
458 and improve upon non-human-primate disease models (*Kennedy et al. 2017*). Despite their varied
459 functions, all Class I and Class II MHC genes are derived from a common ancestor, allowing
460 us to compare genes to learn more about the evolution of the gene family as a whole (*Hansen
461 et al. 2007; Kupfermann et al. 1999; Kaufman 2022; Adams and Luoma 2013*). A few ~20-year-
462 old studies addressed the overall evolution of the MHC gene family via multi-gene alignment and
463 phylogenetics, but the trees had many polytomies (*Adams and Parham 2001b; Sawai et al. 2004;
464 Cardenas et al. 2005; Piontkivska and Nei 2003; Takahashi et al. 2000*). Since then, most work
465 has focused on particular genes or small sets of species, meaning our knowledge of primate MHC
466 evolution is scattered across hundreds of papers (*Urvater et al. 2000; Van Der Wiel et al. 2013;
467 Geller et al. 2002; Hans et al. 2017; Maibach et al. 2017; Wroblewski et al. 2017, 2019; Lafont
468 et al. 2004; Flügge et al. 2002; Go et al. 2003, 2005; Shiina et al. 2017; Abi-Rached et al. 2010;
469 Gleimer et al. 2011; de Groot et al. 2015; De Groot et al. 2012; de Groot et al. 2022; Cao et al.
470 2015; Otting et al. 2020; Fukami-Kobayashi et al. 2005; Lugo and Cadavid 2015; Averdam et al.
471 2011; Figueroa et al. 1994; Doxiadis et al. 2012; Buckner et al. 2021; Doxiadis et al. 2006; Diaz
472 et al. 2000; Gongora et al. 1997; Kasahara et al. 1992; de Groot et al. 2009; Satta et al. 1996*). In
473 this project, we revisited primate MHC evolution with more data from a wider range of species and
474 a coherent analysis framework. We confirm and unify past findings, as well as contribute many
475 new insights into the evolution of this complex family.

476 We found that the Class I genes turn over rapidly, with only the non-classical gene MHC-F being
477 clearly orthologous across the *Simiiformes*. In the rest of the Class I α -block, genes expanded en-
478 tirely separately in the ape/OWM and NWM lineages. This process of expansion generated many
479 full-length and fragment pseudogenes, which we found were equally important as the functional
480 genes to understanding the evolution of the region as a whole. Notably, we found that MHC-U
481 is an MHC-A-related pseudogene, MHC-V is not closely related to MHC-P, and that there were at
482 least three genes of the MHC-W/P/T/OLI family present in the ape/OWM ancestor. Generally, Class
483 II genes do not turn over as rapidly, although there were exceptions. The classical MHC-DRB genes
484 were even shorter-lived than the Class I genes, with most human genes lacking 1:1 orthologs be-
485 yond the great apes. We also found that the classical MHC-DQA and -DQB genes likely expanded
486 separately in the ape/OWM and NWM lineages. In contrast, the classical MHC-DPA and -DPB genes
487 were orthologous across the *Simiiformes*, and the non-classical Class II genes were 1:1 orthologous

488 across most of the mammals. In both Class I and Class II, classical genes turned over more rapidly
489 than non-classical genes and their trees exhibited more deviations from the expected species tree.
490 Overall, our treatment of the genes as related entities instead of distinct cases helped us under-
491 stand shared patterns of evolution across classes and species groups.

492 One concern when discussing gene families is the relative importance of birth-and-death and
493 concerted evolution by gene conversion (*Gu and Nei 1999a; Nei and Rooney 2005; Klein et al.*
494 *2007; Bergström and Gyllensten 1995; Gyllensten et al. 1991; Nei et al. 1997*). Gene conversion
495 can cause adjacent small sequence tracts to have wildly different evolutionary histories, making it
496 difficult to interpret a tree constructed from larger regions. Our trees reveal different topologies
497 depending on exon and our *GENECONV* analysis pulled out several different sequence pairs, re-
498 vealing that gene conversion has played a significant role in the evolution of the MHC genes. With
499 this in mind, comparing trees across exons helps us interpret the overall trees and strengthens
500 our conclusions. Neither birth-and-death nor concerted evolution can be ignored when discussing
501 gene families.

502 The MHC region is difficult to assemble owing to the large number of related genes, extreme
503 polymorphism, and abundant repetitive regions. Nonetheless, several committed researchers
504 have dedicated effort to sequencing and mapping the region in different species (*Wilming et al.*
505 *2013; Anzai et al. 2003; Karl et al. 2023; Okano et al. 2020; Hammer et al. 2020; Liao et al.*
506 *2023*). However, we now know that haplotype variation is just as important as nucleotide variation
507 to maintaining MHC diversity, and haplotypes are surprisingly short-lived (*de Groot et al. 2015,*
508 *2017a; Gleimer et al. 2011; Hans et al. 2015; Heijmans et al. 2020*). Therefore, more research ef-
509 fort should be dedicated to fully characterizing the breadth of MHC haplotypes in different species.
510 This is a difficult problem owing to the plethora of repetitive elements and recently-duplicated
511 genes in the region, and long-read sequencing will be invaluable for parsing these complex haplo-
512 types. Not only will this improve our understanding of health and disease in each species, but it
513 will also help us answer evolutionary questions with better precision.

514 We hope that our extensive set of trees incorporating classical genes, non-classical genes, pseu-
515 dogenes, gene fragments, and alleles of medical interest across a wide range of species will provide
516 context to future researchers. This work will provide a jumping-off-point for further exploration
517 of the evolutionary processes affecting different subsets of the gene family as well as necessary
518 context for studies of particular alleles or genes in disease.

519 **Methods and Materials**

520 **Data Collection**

521 We downloaded MHC allele nucleotide sequences for all human and non-human genes from the
522 IPD Database (collected January 2023) (*Barker et al. 2023; Maccari et al. 2017, 2020; Robinson*
523 *et al. 2024*). To supplement the alleles available in the database, we also collected nucleotide
524 sequences from NCBI using the Entrez E-utilities with query `histocompatibility AND txidX AND`
525 `alive[prop]`, where X is a taxon of interest. This resulted in a very large collection of sequences
526 from a large number of species. While Class II genes were generally assigned to loci, most Class
527 I sequences had ambiguous or no locus assignments. Therefore, we performed a refined search
528 for additional sequences by running *BLAST* on the available primate reference genomes.

529 *BLAST* Search

530 For the reference genomes, we downloaded human chromosome 6 (GenBank accession CM000668 . 2),
531 chimpanzee chromosome 5 (CM054439 . 2), bonobo chromosome 5 (CM055477 . 2), gorilla chromo-
532 some 5 (CM055451 . 2), Sumatran orangutan chromosome 5 (CM054684 . 2), Bornean orangutan chro-
533 mosome 5 (CM054635 . 2), pileated gibbon linkage group LG22 (CM038537 . 1), siamang chromosome
534 23 (CM054531 . 2), Northern white-cheeked gibbon chromosome 22a (CM016966 . 1), olive baboon
535 chromosome 6 (CM018185 . 2), Guinea baboon chromosome 6 (CM053423 . 1), gelada chromosome

536 4 (CM009953.1), Tibetan macaque chromosome 4 (CM045091.1), crab-eating macaque chromosome
537 4 (CP141358.1), Formosan rock macaque chromosome 4 (CM049490.1), mantled guereza chromo-
538 some 5 (CM058078.1), snub-nosed monkey chromosome 4 (CM017354.1), cotton-top tamarin chro-
539 mosome 4 (CM063172.1), golden-handed tamarin linkage group LG04 (CM038394.1), common mar-
540 moset chromosome 4 (CM021918.1), coppery titi chromosome 3 (CM080817.1), gray mouse lemur
541 chromosome 6 (CM007666.1), black-and-white ruffed lemur chromosome 6 (CM052441.1), mongoose
542 lemur chromosome 15 (CM052867.1), ring-tailed lemur chromosome 2 (CM036473.1), Bengal slow
543 loris linkage group LG08 (CM043617.1), Sunda slow loris chromosome 9 (CM050145.1), Philippine
544 flying lemur chromosome 5 (CM050031.1), and mouse chromosome 17 (CM001010.3).

545 To create the *BLAST* database, we first compiled all nucleotide MHC sequences from the IPD-
546 MHC and IPD-IMGT/HLA databases into three *fasta* files: one containing the Class I sequences, one
547 containing the Class II sequences, and one containing MHC-DRB9 sequences. We then constructed
548 three custom databases from these sets of sequences using the *makeblastdb* command in *BLAST*
549 version 2.11.0 (Camacho et al. 2009).

550 We then queried each of the three custom databases using the above reference genomes and
551 screened the hits manually. In most cases, we were able to identify loci unambiguously, resulting
552 in several newly-reported haplotypes (Figure 1—figure Supplement 2 and Figure 2—figure Supple-
553 ment 2). The discovery of various genes in various species also allowed us to fill in gaps in Figure 1
554 and Figure 2.

555 Sequence Selection

556 Because *BEAST2* is computationally limited by the number of sequences, it was necessary to prior-
557 itize certain sequences. To do this, we (very roughly) aligned as many exon 2 and 3 sequences as
558 possible (from both NCBI RefSeq and the IPD database) using *MUSCLE* (Edgar 2004) with default set-
559 tings. We then constructed UPGMA trees in R to visualize the sequences. We preferentially selected
560 sequences that were 1) in primate species not represented by the IPD database or 2) grouped with
561 genes not well represented by the IPD database, and which were not similar/identical to other se-
562 quences. We also included several non-primate species to provide context and explore orthology
563 beyond the primates. After choosing sequences with this preliminary screening method, we col-
564 lected the full-length sequences for inclusion in further analyses. We limited sequences to one per
565 species-gene pair for building the Class I, Class IIA, and Class IIB multi-gene trees (lists of alleles
566 provided as Supplementary Files).

567 For Class I, we then re-aligned all genes together for each exon separately using *MUSCLE* (Edgar
568 2004) with default settings (and manually adjusted). For Class II, alleles for each gene group (MHC-
569 DMA, -DMB, -DOA, -DOB, -DPA, -DPB, -DQA, -DQB, -DRA, and -DRB) were aligned separately for
570 each exon using *MUSCLE* (Edgar 2004) with default settings (and manually adjusted). Since some
571 Class II genes are too far diverged from one another to be reliably aligned automatically, the nu-
572 cleotide alignments were then combined manually based on published amino acid alignments
573 (Radley et al. 1994; Dijkstra et al. 2013; Dijkstra and Yamaguchi 2019; Cuesta et al. 2006; Chen
574 et al. 2006; Chazara et al. 2011). For Class IIA, exons 4 and 5 were concatenated together be-
575 fore this manual combination process because some analogous sites between genes are located
576 across exons. For the same reason, exons 5 and 6 were concatenated together for Class IIB before
577 combining. This produced three multi-gene alignments: Class I, Class IIA, and Class IIB.

578 We also aligned a larger set of sequences for each gene group to create our focused trees
579 that each zoomed in on a different subtree of the multi-gene trees. Details for this are located in
580 the Methods of our companion paper (Fortier and Pritchard 2024).

581 Bayesian Phylogenetic Analysis

582 We constructed phylogenetic trees using *BEAST2* (Bouckaert et al. 2014, 2019) with package *sub*
583 *stBMA* (Wu et al. 2013). *SubstBMA* implements a spike-and-slab mixture model that simultaneously
584 estimates the phylogenetic tree, the number of site partitions, the assignment of sites to partitions,

585 the nucleotide substitution model, and a rate multiplier for each partition. Since we were chiefly
586 interested in the partitions and their rate multipliers, we used the RDPM model as described by **Wu**
587 **et al. (2013)**. In the RDPM model, the number of nucleotide substitution model categories is fixed
588 to 1, so that all sites, regardless of rate partition, share the same estimated nucleotide substitution
589 model. This reduces the number of parameters to be estimated and ensures that only evolution-
590 ary rates vary across site partitions, reducing overall model complexity. We used an uncorrelated
591 lognormal relaxed molecular clock because we wanted evolutionary rates to be able to vary among
592 branches.

593 Priors

594 For the Dirichlet process priors, we used the informative priors constructed by **Wu et al. (2013)** for
595 their mammal dataset. This is appropriate because they include several of the same species and
596 their mammals span approximately the same evolutionary time that we consider in our study. We
597 also use their same priors on tree height, base rate distribution, and a Yule process coalescent
598 prior. We did not specify a calibration point—a time-based prior on a node—because we did not
599 expect our sequences to group according to the species tree.

600 Running *BEAST2*

601 We ran *BEAST2* on various subsets of the three alignments. Considering exons separately helped to
602 minimize the effects of recombination on the tree, while also allowing us to compare and contrast
603 tree topologies for exons encoding the binding site vs. exons encoding the other domains. For
604 Class I, we repeated the analysis for 1) exon 2 only (PBR), 2) exon 3 only (PBR), and 3) exon 4 only
605 (non-PBR). For Class IIA, we used 1) exon 2 only (PBR) and 2) exon 3 only (non-PBR). For Class IIB,
606 we analyzed 1) exon 2 only (PBR) and 2) exon 3 only (non-PBR). In the following, each analysis
607 refers to a collection of *BEAST2* runs using a particular subset of either the Class I, Class IIA, or Class
608 IIB alignment. The procedure is exactly the same for the focused trees, which each focus on a
609 particular gene group within the Class I, Class IIA, or Class IIB alignment. More detail about the
610 generation of the focused trees is located in the Methods of our companion paper (**Fortier and**
611 **Pritchard 2024**).

612 The xml files we used to run *BEAST2* were based closely on those used for the mammal dataset
613 with the RDPM model and uncorrelated relaxed clock in **Wu et al. (2013)** ([https://github.com/jessiewu/
614 substBM/blob/master/examples/mammal/mammal_rdpm_uc.xml](https://github.com/jessiewu/substBM/blob/master/examples/mammal/mammal_rdpm_uc.xml)). Running a model with per-site
615 evolutionary rate categories and a relaxed clock means there are many parameters to estimate.
616 Along with the large number of parameters, highly-polymorphic and highly-diverged sequences
617 make it difficult for *BEAST2* to explore the state space. Thus, we undertook considerable effort to
618 ensure good mixing and convergence of the chains. First, we employed coupled MCMC for all anal-
619 yses. Coupled MCMC is essentially the same as the regular MCMC used in *BEAST2*, except that it
620 uses additional heated chains with increased acceptance probabilities that can traverse unfavor-
621 able intermediate states and allow the main chain to move away from an inferior local optimum
622 (**Müller and Bouckaert 2020**). Using coupled MCMC both speeds up *BEAST2* runs and improves
623 mixing and convergence. We used four heated chains for each run with a delta temperature of
624 0.025. Second, we ran each *BEAST2* run for 40,000,000 states, discarding the first 4,000,000 states
625 as burn-in and sampling every 10,000 states. Third, we ran at least eight independent replicates
626 of each analysis. The replicates use the exact same alignment and coupled MCMC settings, but
627 explore state space independently and thus are useful for improving the effective sample size of
628 tricky parameters. As recommended by *BEAST2*, we examined all replicates in *Tracer* version 1.7.2
629 (**Rambaut et al. 2018**) to ensure that they were sampling from the same parameter distributions
630 and had reached convergence. We excluded replicates for which this was not true, as these chains
631 were probably stuck in suboptimal state space. Additionally, *Tracer* provides estimates of the effec-
632 tive sample size (ESS) for the combined set of states from all chosen replicates, and we required
633 that the combined ESS be larger than 100 for all parameters. If there were fewer than 4 acceptable

634 replicates or if the ESS was below 100 for any parameter, we re-ran more independent replicates
635 of the analysis until these requirements were satisfied. We obtained between 7 and 14 acceptable
636 replicates (median 8) per analysis for the Class I, Class IIA, and Class IIB runs.

637 For some analyses, computational limitations prevented *BEAST2* from being able to reach 40,000,000
638 states. In these situations, more replicates (of fewer states) were usually required to achieve good
639 mixing and convergence. Regardless of how far these *BEAST2* runs got, the first 4,000,000 states
640 from each run were still discarded as burn-in even though this represented more than 10% of
641 states. The xml files required to run all our analyses are provided as Supplementary Files.

642 This extremely stringent procedure ensured that all of the replicates were exploring the same
643 parameter space and were converging upon the same global optimum, allowing the 4 inde-
644 pendent runs to be justifiably combined. We combined the acceptable replicates (discarding the
645 first 4,000,000 states as burn-in) using *LogCombiner* version 2.6.7 (*Drummond and Rambaut 2007*),
646 which aggregates the results across all states. We then used the combined results for downstream
647 analyses.

648 **Phylogenetic Trees**

649 After combining acceptable replicates, we obtained 17,927 - 28,384 phylogenies per gene/sequence
650 subset for the Class I, Class IIA, and Class IIB trees (mean 25,154). We used *TreeAnnotator* version
651 2.6.3 (*Drummond and Rambaut 2007*) to summarize each set of possible trees as a maximum
652 clade credibility tree, which is the tree that maximizes the sum of posterior clade probabilities.
653 Since *BEAST2* samples trees from the posterior, one could in principle reduce the large set of trees
654 to a smaller 95% credible set of trees representing the true tree (*BEA 2024*). However, given the
655 high complexity of the model space, all our posterior trees were unique, meaning this was not
656 possible in practice. Throughout this paper, we rely on summary trees for our observations.

657 **Integration with Literature**

658 Hundreds of authors have contributed to the study of MHC evolution, and their myriad published
659 results played a key role in this project. *Figure 8* illustrates our approach to this project, including
660 how we used existing literature and how we divided results among this paper and its companion
661 (*Fortier and Pritchard 2024*). We first constructed large multi-gene trees encompassing all Class
662 I, Class IIA, and Class IIB genes. These provided a backbone for us to investigate subtrees in more
663 depth, adding more sequences and more species to construct focused trees for each gene group.
664 These, in combination with the literature, allowed us to create hypotheses about the evolution of
665 the Class I α -block (*Figure 6*) and Class II MHC-DRB region (*Figure 7*).

666 **Gene Conversion**

667 We inferred gene conversion fragments using *GENECONV* version 1.81a (*Sawyer 1999*) on each fo-
668 cused alignment. It is generally advisable to use only synonymous sites when running the program
669 on a protein-coding alignment, since silent sites within the same codon position are likely to be cor-
670 related. However, the extreme polymorphism in these MHC genes meant there were too few silent
671 sites to use in the analysis. Thus, we considered all sites but caution that this could slightly overes-
672 timate the lengths of our inferred conversion tracts. For each alignment, we ran *GENECONV* with
673 options `ListPairs`, `llouter`, `Numsims=10000`, and `Startseed=310`. We collected all inferred Global
674 Inner (GI) fragments with `sim_pval < 0.05` (this is pre-corrected for multiple comparisons by the
675 program). GI fragments indicate a stretch of similar sequence shared by two otherwise-dissimilar
676 sequences in the alignment. This suggests that a gene conversion event occurred between the
677 ancestors of the two sequences.

678 Many of the thousands of GI hits were redundant, involving very-closely-related alleles, slightly
679 different fragment bounds, or even a wide range of species all implicating the same gene. We
680 manually grouped and summarized these hits for *Figure 3—source data 1* and *Figure 4—source*
681 *data 1*. The `start` and `end` columns indicate the smallest start and largest end position (along

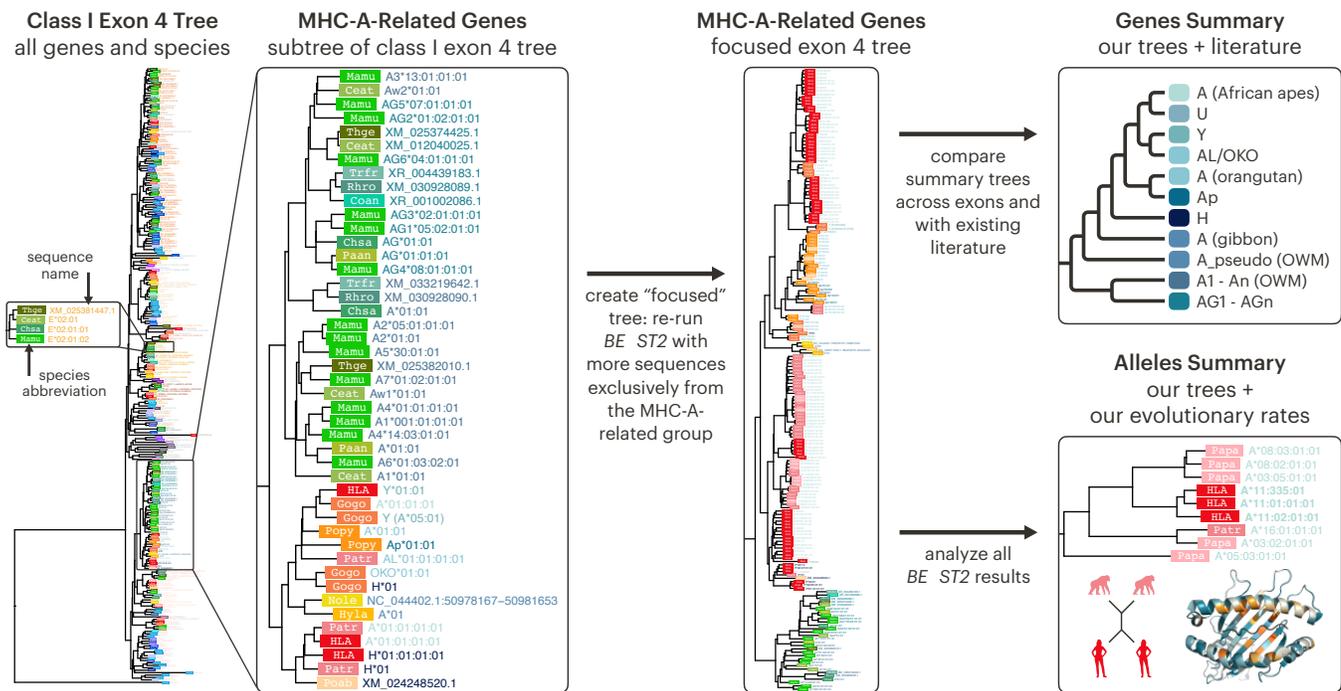


Figure 8. BEAST2 trees provide insight into MHC gene and allele relationships. We first created multi-gene Bayesian phylogenetic trees using sequences from all genes and species, separated into Class I, Class IIA, and Class IIB groups. We then focused on various subtrees of the multi-gene trees by adding more sequences for each subtree and running BEAST2 using only sequences from that group (in addition to the backbone sequences common to all trees). Our trees gave us insight into both overall gene relationships (this paper) and allele relationships within gene groups (see our companion paper, Fortier and Pritchard (2024)).

682 the alignment) for the group of redundant hits, and the sequences involved are summarized as
683 specifically as possible.

684 Acknowledgments

685 We acknowledge support from NIH grants R01 HG011432 and R01 HG008140. This material is
686 based upon work supported by the National Science Foundation Graduate Research Fellowship under
687 Grant No. DGE-1656518. We appreciate helpful comments from Jeffrey Spence, the Pritchard
688 lab, and the reviewers of the previous version of our companion paper, which jumpstarted this
689 project.

690 References

- 691 SUMMARIZING POSTERIOR TREES. Centre for Computational Evolution; 2024. <https://www.beast2.org/summarizing-posterior-trees/>.
- 692
- 693 **Abi-Rached L, Kuhl H, Roos C, ten Hallers B, Zhu B, Carbone L, de Jong PJ, Mootnick AR, Knaust F, Reinhardt R, Parham P, Walter L.** A Small, Variable, and Irregular Killer Cell Ig-Like Receptor Locus Accompanies the Absence of MHC-C and MHC-G in Gibbons. *The Journal of Immunology*. 2010; 184(3):1379–1391. doi: 10.4049/jimmunol.0903016.
- 694
- 695
- 696
- 697 **Adams EJ, Parham P.** Genomic analysis of common chimpanzee major histocompatibility complex class I genes. *Immunogenetics*. 2001; 53(3):200–208. doi: 10.1007/s002510100318.
- 698
- 699 **Adams EJ, Luoma AM.** The adaptable major histocompatibility complex (MHC) fold: Structure and function of nonclassical and mhc class I-like molecules. *Annual Review of Immunology*. 2013; 31:529–561. doi: 10.1146/annurev-immunol-032712-095912.
- 700
- 701
- 702 **Adams EJ, Parham P.** Species-specific evolution of MHC class I genes in the higher primates. *Immunological Reviews*. 2001; 183:41–64. doi: 10.1034/j.1600-065X.2001.1830104.x.
- 703

- 704 **Alexandrov N**, Wang T, Blair L, Nadon B, Sayer D. HLA-OL1: A new MHC class I pseudogene and HLA-Y are
705 located on a 60 kb indel in the human MHC between HLA-W and HLA-J. *Hla*. 2023; 102(5):599–606. doi:
706 [10.1111/tan.15180](https://doi.org/10.1111/tan.15180).
- 707 **Anderson JL**, Sandstrom K, Smith WR, Wetzel M, Klenchin VA, Evans DT. MHC Class I Ligands of Rhesus
708 Macaque Killer Cell Ig-like Receptors. *The Journal of Immunology*. 2023; 210(11):1815–1826. doi: [10.4049/jim-
munol.2200954](https://doi.org/10.4049/jim-
709 munol.2200954).
- 710 **Anzai T**, Shiina T, Kimura N, Yanagiya K, Kohara S, Shigenari A, Yamagata T, Kulski JK, Naruse TK, Fujimori Y,
711 Fukuzumi Y, Yamazaki M, Tashiro H, Iwamoto C, Umehara Y, Imanishi T, Meyer A, Ikeo K, Gojobori T, Bahram
712 S, et al. Comparative sequencing of human and chimpanzee MHC class I regions unveils insertions/deletions
713 as the major path to genomic divergence. *Proceedings of the National Academy of Sciences of the United
714 States of America*. 2003; 100(13):7708–7713. doi: [10.1073/pnas.1230533100](https://doi.org/10.1073/pnas.1230533100).
- 715 **Arden B**, Klein J. Biochemical comparison of major histocompatibility complex molecules from different sub-
716 species of *Mus musculus*: Evidence for trans-specific evolution of alleles. *Proceedings of the National
717 Academy of Sciences of the United States of America*. 1982; 79(7):2342–2346. doi: [10.1073/pnas.79.7.2342](https://doi.org/10.1073/pnas.79.7.2342).
- 718 **Averdam A**, Kuschal C, Otto N, Westphal N, Roos C, Reinhardt R, Walter L. Sequence analysis of the grey
719 mouse lemur (*Microcebus murinus*) MHC class II DQ and DR region. *Immunogenetics*. 2011; 63(2):85–93.
720 doi: [10.1007/s00251-010-0487-3](https://doi.org/10.1007/s00251-010-0487-3).
- 721 **Barker DJ**, Maccari G, Georgiou X, Cooper MA, Flicek P, Robinson J, Marsh SGE. The IPD-IMGT/HLA Database.
722 *Nucleic Acids Research*. 2023; 51(November 2022):1053–1060.
- 723 **Benton MJ**, Donoghue PCJ, Asher RJ, Friedman M, Near TJ, Vinther J. Constraints on the timescale of animal
724 evolutionary history. *Palaeontologia Electronica*. 2015; 18(1):1–107. doi: [10.26879/424](https://doi.org/10.26879/424).
- 725 **Bergström T**, Gyllensten U. Evolution of Mhc Class II Polymorphism: The Rise and Fall of Class II Gene Function
726 in Primates. *Immunological Reviews*. 1995; 143(1):13–31. doi: [10.1111/j.1600-065X.1995.tb00668.x](https://doi.org/10.1111/j.1600-065X.1995.tb00668.x).
- 727 **Biassoni R**, Malnati MS. Human natural killer receptors, co-receptors, and their ligands. *Current Protocols in
728 Immunology*. 2018; e47:1–52. doi: [10.1002/0471142735.im1410s84](https://doi.org/10.1002/0471142735.im1410s84).
- 729 **Bondarenko GI**, Burleigh DW, Durning M, Breburda EE, Grendell RL, Golos TG. Passive Immunization against
730 the MHC Class I Molecule Mamu-AG Disrupts Rhesus Placental Development and Endometrial Responses.
731 *The Journal of Immunology*. 2007; 179(12):8042–8050. doi: [10.4049/jimmunol.179.12.8042](https://doi.org/10.4049/jimmunol.179.12.8042).
- 732 **Bondarenko GI**, Dambaeva SV, Grendell RL, Hughes AL, Durning M, Garthwaite MA, Golos TG. Characterization
733 of cynomolgus and vervet monkey placental MHC class I expression: Diversity of the nonhuman primate AG
734 locus. *Immunogenetics*. 2009; 61(6):431–442. doi: [10.1007/s00251-009-0376-9](https://doi.org/10.1007/s00251-009-0376-9).
- 735 **Bouckaert R**, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. BEAST 2:
736 A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*. 2014; 10(4):1–6. doi:
737 [10.1371/journal.pcbi.1003537](https://doi.org/10.1371/journal.pcbi.1003537).
- 738 **Bouckaert R**, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Küh-
739 nert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, Du Plessis L, Poppinga A, Rambaut A,
740 Rasmussen D, Siveroni I, Suchard MA, et al. BEAST 2.5: An advanced software platform for Bayesian evolu-
741 tionary analysis. *PLoS Computational Biology*. 2019; 15(4):1–28. doi: [10.1371/journal.pcbi.1006650](https://doi.org/10.1371/journal.pcbi.1006650).
- 742 **Boyson JE**, McAdam SN, Gallimore A, Golos TG, Liu X, Gotch FM, Hughes AL, Watkins DI. The MHC E locus in
743 macaques is polymorphic and is conserved between macaques and humans. *Immunogenetics*. 1995; 41(2-
744 3):59–68. doi: [10.1007/BF00182314](https://doi.org/10.1007/BF00182314).
- 745 **Buckner JC**, Jack KM, Melin AD, Schoof VAM, Gutiérrez-Espeleta GA, Lima MGM, Lynch JW. Major histocompat-
746 ibility complex class II DR and DQ evolution and variation in wild capuchin monkey species (Cebinae). *PLoS
747 ONE*. 2021; 16(8):e0254604. doi: [10.1371/journal.pone.0254604](https://doi.org/10.1371/journal.pone.0254604).
- 748 **Budde ML**, Wiseman RW, Karl JA, Hanczaruk B, Simen BB, O'Connor DH. Characterization of Mauritian cynomol-
749 gus macaque major histocompatibility complex class I haplotypes by high-resolution pyrosequencing. *Im-
750 munogenetics*. 2010; 62(11-12):773–780. doi: [10.1007/s00251-010-0481-9](https://doi.org/10.1007/s00251-010-0481-9).
- 751 **Budeus B**, Álvaro-Benito M, Crivello P. HLA-DM and HLA-DO interplay for the peptide editing of HLA class
752 II in healthy tissues and leukemia. *Best Practice and Research: Clinical Haematology*. 2024; 37(2). doi:
753 [10.1016/j.beha.2024.101561](https://doi.org/10.1016/j.beha.2024.101561).

- 754 **Buniello A**, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy
755 E, Sollis E, Suveges D, Vrousseau O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H,
756 Flicek P, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays
757 and summary statistics 2019. *Nucleic Acids Research*. 2019 1; 47(D1):D1005–D1012. <https://academic.oup.com/nar/article/47/D1/D1005/5184712>, doi: 10.1093/nar/gky1120.
- 759 **Cadauid LF**, Hughes AL, Watkins DI. MHC class I-processed pseudogenes in New World primates provide
760 evidence for rapid turnover of MHC class I genes. *Journal of immunology (Baltimore, Md : 1950)*. 1996;
761 157(6):2403–9. <http://www.ncbi.nlm.nih.gov/pubmed/8805639>.
- 762 **Camacho C**, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: Architecture and
763 applications. *BMC Bioinformatics*. 2009; 10:1–9. doi: 10.1186/1471-2105-10-421.
- 764 **Cao YH**, Fan JW, Li AX, Liu HF, Li LR, Zhang CL, Zeng L, Sun ZZ. Identification of MHC I class genes in two Platyrrhini
765 species. *American Journal of Primatology*. 2015; 77(5):527–534. doi: 10.1002/ajp.22372.
- 766 **Cardenas PP**, Suarez CF, Martinez P, Patarroyo ME, Patarroyo MA. MHC class I genes in the owl monkey: Mosaic
767 organisation, convergence and loci diversity. *Immunogenetics*. 2005; 56(11):818–832. doi: 10.1007/s00251-
768 004-0751-5.
- 769 **Carlini F**, Ferreira V, Buhler S, Tous A, Eliaou JF, René C, Chiaroni J, Picard C, Di Cristofaro J. Association of HLA-A
770 and non-classical HLA class I alleles. *PLoS ONE*. 2016; 11(10):1–17. doi: 10.1371/journal.pone.0163570.
- 771 **Carter AM**. Unique aspects of human placentation. *International Journal of Molecular Sciences*. 2021; 22(15).
772 doi: 10.3390/ijms22158099.
- 773 **Chazara O**, Tixier-Boichard M, Morin V, Zoorob R, Bed'Hom B. Organisation and diversity of the class II DM
774 region of the chicken MHC. *Molecular Immunology*. 2011; 48(9-10):1263–1271. <http://dx.doi.org/10.1016/j.molimm.2011.03.009>, doi: 10.1016/j.molimm.2011.03.009.
- 776 **Chen JM**, Cooper DN, Chuzhanova N, Férec C, Patrinos GP. Gene conversion: Mechanisms, evolution and
777 human disease. *Nature Reviews Genetics*. 2007; 8(10):762–775. doi: 10.1038/nrg2193.
- 778 **Chen SL**, Zhang YX, Xu MY, Ji XS, Yu GC, Dong CF. Molecular polymorphism and expression analysis of MHC
779 class II B gene from red sea bream (*Chrysophrys major*). *Developmental and Comparative Immunology*.
780 2006; 30(4):407–418. doi: 10.1016/j.dci.2005.06.001.
- 781 **Cuesta A**, Ángeles Esteban M, Meseguer J. Cloning, distribution and up-regulation of the teleost fish MHC
782 class II alpha suggests a role for granulocytes as antigen-presenting cells. *Molecular Immunology*. 2006;
783 43(8):1275–1285. doi: 10.1016/j.molimm.2005.07.004.
- 784 **De Groot NG**, Otting N, Robinson J, Blancher A, Lafont BAP, Marsh SGE, O'Connor DH, Shiina T, Walter L, Watkins
785 DI, Bontrop RE. Nomenclature report on the major histocompatibility complex genes and alleles of Great
786 Ape, Old and New World monkey species. *Immunogenetics*. 2012; 64(8):615–631. doi: 10.1007/s00251-012-
787 0617-1.
- 788 **Demuth JP**, Bie TD, Stajich JE, Christiani N, Hahn MW. The evolution of mammalian gene families. *PLoS ONE*.
789 2006; 1(1). doi: 10.1371/journal.pone.0000085.
- 790 **Dendrou CA**, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nature Reviews Immunology*. 2018;
791 18(5):325–339. <http://dx.doi.org/10.1038/nri.2017.143>, doi: 10.1038/nri.2017.143.
- 792 **Diaz D**, Naegeli M, Rodriguez R, Nino-Vasquez JJ, Moreno A, Patarroyo ME, Pluschke G, Daubenberger CA. Se-
793 quence and diversity of MHC DQA and DQB genes of the owl monkey *Aotus nancymae*. *Immunogenetics*.
794 2000; 51(7):528–537. doi: 10.1007/s002510000189.
- 795 **Dijkstra JM**, Grimholt U, Leong J, Koop BF, Hashimoto K. Comprehensive analysis of MHC class II genes in
796 teleost fish genomes reveals dispensability of the peptide-loading DM system in a large part of vertebrates.
797 *BMC Evolutionary Biology*. 2013; 13(1):1–14. doi: 10.1186/1471-2148-13-260.
- 798 **Dijkstra JM**, Yamaguchi T. Ancient features of the MHC class II presentation pathway, and a model for the
799 possible origin of MHC molecules. *Immunogenetics*. 2019; 71(3):233–249. doi: 10.1007/s00251-018-1090-2.
- 800 **Douillard V**, Castelli EC, Mack SJ, Hollenbach JA, Gourraud PA, Vince N, Limou S. Approaching Genetics Through
801 the MHC Lens: Tools and Methods for HLA Research. *Frontiers in Genetics*. 2021; 12(December):1–13. doi:
802 [10.3389/fgene.2021.774916](https://doi.org/10.3389/fgene.2021.774916).

- 803 **Doxiadis GGM**, Hoof I, De Groot N, Bontrop RE. Evolution of HLA-DRB genes. *Molecular Biology and Evolution*.
804 2012; 29(12):3843–3853. doi: 10.1093/molbev/mss186.
- 805 **Doxiadis GGM**, Rouweler AJM, De Groot NG, Louwense A, Otting N, Verschoor EJ, Bontrop RE. Extensive sharing
806 of MHC class II alleles between rhesus and cynomolgus macaques. *Immunogenetics*. 2006; 58(4):259–268.
807 doi: 10.1007/s00251-006-0083-8.
- 808 **Drummond AJ**, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary*
809 *Biology*. 2007; 7(1):214. <http://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-7-214>, doi:
810 10.1186/1471-2148-7-214.
- 811 **Ebert D**, Fields PD. Host–parasite co-evolution and its genomic signature. *Nature Reviews Genetics*. 2020 12;
812 21(12):754–768. <https://www.nature.com/articles/s41576-020-0269-1>, doi: 10.1038/s41576-020-0269-1.
- 813 **Edgar RC**. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Re-*
814 *search*. 2004 3; 32(5):1792–1797. <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkh340>, doi:
815 10.1093/nar/gkh340.
- 816 **Figuroa F**, O’huigin C, Tichy H, Klein J. The origin of the primate Mhc-DRB genes and allelic lineages as de-
817 duced from the study of prosimians. *The Journal of Immunology*. 1994; 152(9):4455–4465. doi: 10.4049/jim-
818 [munol.152.9.4455](https://doi.org/10.4049/jimmunol.152.9.4455).
- 819 **Flajnik MF**, Kasahara M. Comparative genomics of the MHC: Glimpses into the evolution of the adaptive im-
820 mune system. *Immunity*. 2001; 15(3):351–362. doi: 10.1016/S1074-7613(01)00198-4.
- 821 **Flügge P**, Zimmermann E, Hughes AL, Günther E, Walter L. Characterization and phylogenetic relationship of
822 prosimian MHC class I genes. *Journal of Molecular Evolution*. 2002; 55(6):768–775. doi: 10.1007/s00239-002-
823 2372-7.
- 824 **Foley NM**, Mason VC, Harris AJ, Bredemeyer KR, Damas J, Lewin HA, Eizirik E, Gatesy J, Karlsson EK, Lindblad-Toh
825 K, Springer MS, Murphy WJ, Andrews G, Armstrong JC, Bianchi M, Birren BW, Bredemeyer KR, Breit AM, Christ-
826 mas MJ, Clawson H, et al. A genomic timescale for placental mammal evolution. *Science*. 2023; 380(6643).
827 doi: 10.1126/science.abl8189.
- 828 **Fortier AL**, Pritchard JK. Ancient Trans-Species Polymorphism at the Major Histocompatibility Complex in Pri-
829 mates. *TBD*. 2024; .
- 830 **Friedman R**, Hughes AL. The temporal distribution of gene duplication events in a set of highly conserved
831 human gene families. *Molecular Biology and Evolution*. 2003; 20(1):154–161. doi: 10.1093/molbev/msg017.
- 832 **Fukami-Kobayashi K**, Shiina T, Anzai T, Sano K, Yamazaki M, Inoko H, Tateno Y. Genomic evolution of MHC
833 class I region in primates. *Proceedings of the National Academy of Sciences of the United States of America*.
834 2005; 102(26):9230–9234. doi: 10.1073/pnas.0500770102.
- 835 **Geller R**, Adams EJ, Guethlein LA, Little AM, Madrigal AJ, Parham P. Linkage of Patr-AL to Patr-A and-B in the
836 major histocompatibility complex of the common chimpanzee (*pan troglodytes*). *Immunogenetics*. 2002;
837 54(3):212–215. doi: 10.1007/s00251-002-0452-x.
- 838 **Geraghty DE**, Koller BH, Hansen JA, Orr HT. The HLA class I gene family includes at least six genes and twelve
839 pseudogenes and gene fragments. *The Journal of Immunology*. 1992; 149(6):1934–1946. doi: 10.4049/jim-
840 [munol.149.6.1934](https://doi.org/10.4049/jimmunol.149.6.1934).
- 841 **Gfeller D**, Bassani-Sternberg M. Predicting antigen presentation-What could we learn from a million peptides?
842 *Frontiers in Immunology*. 2018; 9(JUL):1–17. doi: 10.3389/fimmu.2018.01716.
- 843 **Gleimer M**, Wahl AR, Hickman HD, Abi-Rached L, Norman PJ, Guethlein LA, Hammond JA, Draghi M, Adams
844 EJ, Juo S, Jalili R, Gharizadeh B, Ronaghi M, Garcia KC, Hildebrand WH, Parham P. Although Divergent in
845 Residues of the Peptide Binding Site, Conserved Chimpanzee Patr-AL and Polymorphic Human HLA-A*02
846 Have Overlapping Peptide-Binding Repertoires. *The Journal of Immunology*. 2011; 186(3):1575–1588. doi:
847 10.4049/jimmunol.1002990.
- 848 **Go Y**, Rakotoarisoa G, Kawamoto Y, Shima T, Koyama N, Randrianjafy A, Mora R, Hirai H. Characterization and
849 evolution of major histocompatibility complex class II genes in the aye-aye, *Daubentonia madagascariensis*.
850 *Primates*. 2005; 46(2):135–139. doi: 10.1007/s10329-004-0101-0.

- 851 **Go Y**, Satta Y, Kawamoto Y, Rakotoarisoa G, Randrianjafy A, Koyama N, Hirai H. Frequent segmental sequence
852 exchanges and rapid gene duplication characterize the MHC class I genes in lemurs. *Immunogenetics*. 2003;
853 55(7):450–461. doi: 10.1007/s00251-003-0613-6.
- 854 **Gongora R**, Figueroa F, O'Huigin C, Klein J. HLA-DRB9 - Possible remnant of an ancient functional DRB subregion.
855 *Scandinavian Journal of Immunology*. 1997; 45(5):504–510. doi: 10.1046/j.1365-3083.1997.d01-428.x.
- 856 **Goyos A**, Guethlein LA, Horowitz A, Hilton HG, Gleimer M, Brodsky FM, Parham P. A Distinctive Cytoplasmic
857 Tail Contributes to Low Surface Expression and Intracellular Retention of the Patr-AL MHC Class I Molecule.
858 *The Journal of Immunology*. 2015; 195(8):3725–3736. doi: 10.4049/jimmunol.1500397.
- 859 **Grimley C**, Mather KA, Ober C. HLA-H: A pseudogene with increased variation due to balancing selection
860 at neighboring loci. *Molecular Biology and Evolution*. 1998; 15(12):1581–1588. doi: 10.1093/oxfordjour-
861 nals.molbev.a025886.
- 862 **de Groot N**, Stanbury K, de Vos-Rouweler AJM, de Groot NG, Poirier N, Blancho G, de Luna C, Doxiadis GGM,
863 Bontrop RE. A quick and robust MHC typing method for free-ranging and captive primate species. *Immuno-*
864 *genetics*. 2017; 69(4):231–240. doi: 10.1007/s00251-016-0968-0.
- 865 **de Groot NG**, Blokhuis JH, Otting N, Doxiadis GGM, Bontrop RE. Co-evolution of the MHC class I and KIR gene
866 families in rhesus macaques: Ancestry and plasticity. *Immunological Reviews*. 2015; 267(1):228–245. doi:
867 10.1111/imr.12313.
- 868 **de Groot NG**, de Groot N, de Vos-Rouweler AJM, Louwense A, Bruijnesteijn J, Bontrop RE. Dynamic evolution of
869 Mhc haplotypes in cynomolgus macaques of different geographic origins. *Immunogenetics*. 2022; 74(4):409–
870 429. <https://doi.org/10.1007/s00251-021-01249-y>, doi: 10.1007/s00251-021-01249-y.
- 871 **de Groot NG**, Heijmans CMC, de Ru AH, Janssen GMC, Drijfhout JW, Otting N, Vangenot C, Doxiadis GGM, Koning
872 F, van Veelen PA, Bontrop RE. A Specialist Macaque MHC Class I Molecule with HLA-B*27-like Peptide-Binding
873 Characteristics. *The Journal of Immunology*. 2017; 199(10):3679–3690. doi: 10.4049/jimmunol.1700502.
- 874 **de Groot NG**, Heijmans CMC, van der Wiel MKH, Blokhuis JH, Mulder A, Guethlein LA, Doxiadis GGM, Claas FHJ,
875 Parham P, Bontrop RE. Complex MHC Class I Gene Transcription Profiles and Their Functional Impact in
876 Orangutans. *The Journal of Immunology*. 2016; 196(2):750–758. doi: 10.4049/jimmunol.1500820.
- 877 **de Groot NG**, Heijmans CMC, de Groot N, Doxiadis GGM, Otting N, Bontrop RE. The chimpanzee Mhc-DRB
878 region revisited: Gene content, polymorphism, pseudogenes, and transcripts. *Molecular Immunology*. 2009;
879 47(2-3):381–389. doi: 10.1016/j.molimm.2009.09.003.
- 880 **de Groot NG**, Otting N, Maccari G, Robinson J, Hammond JA, Blancher A, Lafont BAP, Guethlein LA, Wroblewski
881 EE, Marsh SGE, Shiina T, Walter L, Vigilant L, Parham P, O'Connor DH, Bontrop RE. Nomenclature report 2019:
882 major histocompatibility complex genes and alleles of Great and Small Ape and Old and New World monkey
883 species. *Immunogenetics*. 2020; 72(1-2):25–36. doi: 10.1007/s00251-019-01132-x.
- 884 **Gu X**, Nei M. Locus specificity of polymorphic alleles and evolution by a birth-and death process in mam-
885 malian MHC genes. *Molecular Biology and Evolution*. 1999; 16(2):147–156. doi: 10.1093/oxfordjour-
886 nals.molbev.a026097.
- 887 **Gu X**, Nei M. Locus specificity of polymorphic alleles and evolution by a birth-and death process in mam-
888 malian MHC genes. *Molecular Biology and Evolution*. 1999; 16(2):147–156. doi: 10.1093/oxfordjour-
889 nals.molbev.a026097.
- 890 **Gu X**, Wang Y, Gu J. Age distribution of human gene families shows significant roles of both large- and small-
891 scale duplications in vertebrate evolution. *Nature Genetics*. 2002; 31(2):205–209. doi: 10.1038/ng902.
- 892 **Guethlein LA**, Norman PJ, Hilton HHG, Parham P. Co-evolution of MHC class I and variable NK cell receptors
893 in placental mammals. *Immunological Reviews*. 2015; 267(1):259–282. doi: 10.1111/imr.12326.
- 894 **Gyllensten UB**, Sundvall M, Erlich HA. Allelic diversity is generated by intraexon sequence exchange at the
895 DRB1 locus of primates. *Proceedings of the National Academy of Sciences of the United States of America*.
896 1991; 88(9):3686–3690. doi: 10.1073/pnas.88.9.3686.
- 897 **Hahn MW**, De Bie T, Stajich JE, Nguyen C, Cristianini N. Estimating the tempo and mode of gene family evolution
898 from comparative genomic data. *Genome Research*. 2005; 15(8):1153–1160. doi: 10.1101/gr.3567505.

- 899 **Hammer SE**, Ho CS, Ando A, Rogel-Gaillard C, Charles M, Tector M, Tector AJ, Lunney JK. Importance of the Major
900 Histocompatibility Complex (Swine Leukocyte Antigen) in Swine Health and Biomedical Research. *Annual*
901 *Review of Animal Biosciences*. 2020; 8:171–198. doi: 10.1146/annurev-animal-020518-115014.
- 902 **Hans JB**, Bergl RA, Vigilant L. Gorilla MHC class I gene and sequence variation in a comparative context. *Im-*
903 *munogenetics*. 2017; 69(5):303–323. doi: 10.1007/s00251-017-0974-x.
- 904 **Hans JB**, Haubner A, Arandjelovic M, Bergl RA, Fünfstück T, Gray M, Morgan DB, Robbins MM, Sanz C, Vigilant L.
905 Characterization of MHC class II B polymorphism in multiple populations of wild gorillas using non-invasive
906 samples and next-generation sequencing. *American Journal of Primatology*. 2015; 77(11):1193–1206. doi:
907 [10.1002/ajp.22458](https://doi.org/10.1002/ajp.22458).
- 908 **Hansen TH**, Huang S, Arnold PL, Fremont DH. Patterns of nonclassical MHC antigen presentation. *Nature*
909 *Immunology*. 2007; 8(6):563–568. doi: 10.1038/ni1475.
- 910 **Heijmans CMC**, de Groot NG, Bontrop RE. Comparative genetics of the major histocompatibility complex
911 in humans and nonhuman primates. *International Journal of Immunogenetics*. 2020; 47(3):243–260. doi:
912 [10.1111/iji.12490](https://doi.org/10.1111/iji.12490).
- 913 **Heimbruch KE**, Karl JA, Wiseman RW, Dudley DM, Johnson Z, Kaur A, O'Connor DH. Novel MHC class I full-
914 length allele and haplotype characterization in sooty mangabeys. *Immunogenetics*. 2015; 67(8):437–445.
915 doi: 10.1007/s00251-015-0847-0.
- 916 **Hilton HG**, Parham P. Missing or altered self: human NK cell receptors that recognize HLA-C. *Immunogenetics*.
917 2017; 69(8-9):567–579. doi: 10.1007/s00251-017-1001-y.
- 918 **Horton R**, Gibson R, Coghill P, Miretti M, Allcock RJ, Almeida J, Forbes S, Gilbert JGR, Halls K, Harrow JL, Hart E,
919 Howe K, Jackson DK, Palmer S, Roberts AN, Sims S, Stewart CA, Traherne JA, Trevanion S, Wilming L, et al. Vari-
920 ation analysis and gene annotation of eight MHC haplotypes: The MHC Haplotype Project. *Immunogenetics*.
921 2008; 60(1):1–18. doi: 10.1007/s00251-007-0262-2.
- 922 **Hubert L**, Paganini J, Picard C, Chiaroni J, Abi-Rached L, Pontarotti P, Di Cristofaro J. HLA-H*02:07 Is a Membrane-
923 Bound Ligand of Denisovan Origin That Protects against Lysis by Activated Immune Effectors. *The Journal of*
924 *Immunology*. 2022; 208(1):49–53. doi: [10.4049/jimmunol.2100358](https://doi.org/10.4049/jimmunol.2100358).
- 925 **Hughes AL**, Nei M. Evolution of the major histocompatibility complex: independent origin of nonclassical
926 class I genes in different groups of mammals. *Molecular Biology and Evolution*. 1989; 6(6):559–579. doi:
927 [10.1093/oxfordjournals.molbev.a040573](https://doi.org/10.1093/oxfordjournals.molbev.a040573).
- 928 **Hughes AL**, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for over-
929 dominant selection. *Proceedings of the National Academy of Sciences of the United States of America*. 1989;
930 86(3):958–962. doi: [10.1073/pnas.86.3.958](https://doi.org/10.1073/pnas.86.3.958).
- 931 **Hughes AL**. Origin and evolution of HLA class I pseudogenes. *Molecular Biology and Evolution*. 1995; 12(2):247–
932 258. doi: [10.1093/oxfordjournals.molbev.a040201](https://doi.org/10.1093/oxfordjournals.molbev.a040201).
- 933 **Hughes AL**, Hughes MK. Natural selection on the peptide-binding regions of major histocompatibility complex
934 molecules. *Immunogenetics*. 1995; 42(4):233–243. doi: 10.1007/BF00176440.
- 935 **Hughes AL**, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveal
936 overdominant selection. *Nature*. 1988; 335(6186):167–170. doi: 10.1038/335167a0.
- 937 **Hurley CK**. Naming HLA diversity: A review of HLA nomenclature. *Human Immunology*. 2021; 82(7):457–465.
938 <https://doi.org/10.1016/j.humimm.2020.03.005>, doi: [10.1016/j.humimm.2020.03.005](https://doi.org/10.1016/j.humimm.2020.03.005).
- 939 **Karl JA**, Prall TM, Bussan HE, Varghese JM, Pal A, Wiseman RW, O'Connor DH. Complete sequencing of a
940 cynomolgus macaque major histocompatibility complex haplotype. *Genome Research*. 2023; 33(3):448–462.
941 doi: [10.1101/gr.277429.122](https://doi.org/10.1101/gr.277429.122).
- 942 **Kasahara M**, Klein D, Vincek V, Sarapata DE, Klein J. Comparative anatomy of the primate major histocompati-
943 bility complex DR subregion: Evidence for combinations of DRB genes conserved across species. *Genomics*.
944 1992; 14(2):340–349. doi: 10.1016/S0888-7543(05)80224-1.
- 945 **Kaufman J**. The new W family reconstructs the evolution of MHC genes. *Proceedings of the National Academy*
946 *of Sciences*. 2022; 119(5):119–121.

- 947 **Kennedy AE**, Ozbek U, Dorak MT. What has GWAS done for HLA and disease associations? *International*
948 *Journal of Immunogenetics*. 2017; 44(5):195–211. doi: [10.1111/iji.12332](https://doi.org/10.1111/iji.12332).
- 949 **Klein J**, Sato A, Nagl S, Colm O. Molecular trans-species polymorphism. *Annual Review of Ecology and System-*
950 *atics*. 1998; 21(1):1–21.
- 951 **Klein J**, Sato A, Nikolaidis N. MHC, TSP, and the origin of species: From immunogenetics to evolutionary genetics.
952 *Annual Review of Genetics*. 2007; 41:281–304. doi: [10.1146/annurev.genet.41.110306.130137](https://doi.org/10.1146/annurev.genet.41.110306.130137).
- 953 **Knapp LA**, Cadavid LF, Watkins DI. The MHC-E Locus Is the Most Well Conserved of All Known Primate Class I His-
954 tocompatibility Genes . *The Journal of Immunology*. 1998; 160(1):189–196. doi: [10.4049/jimmunol.160.1.189](https://doi.org/10.4049/jimmunol.160.1.189).
- 955 **Kono A**, Brameier M, Roos C, Suzuki S, Shigenari A, Kametani Y, Kitaura K, Matsutani T, Suzuki R, Inoko H, Walter
956 L, Shiina T. Genomic Sequence Analysis of the MHC Class I G/F Segment in Common Marmoset (*Callithrix*
957 *jacchus*). *The Journal of Immunology*. 2014; 192(7):3239–3246. doi: [10.4049/jimmunol.1302745](https://doi.org/10.4049/jimmunol.1302745).
- 958 **Kuderna LFK**, Gao H, Janiak MC, Kuhlwil M, Orkin JD, Bataillon T, Manu S, Valenzuela A, Bergman J, Rousselle
959 M, Silva FE, Agueda L, Blanc J, Gut M, de Vries D, Goodhead I, Harris RA, Raveendran M, Jensen A, Chuma IS,
960 et al. A global catalog of whole-genome diversity from 233 primate species. *Science (New York, NY)*. 2023;
961 380(6648):906–913. doi: [10.1126/science.abn7829](https://doi.org/10.1126/science.abn7829).
- 962 **Kulski JK**, Anzai T, Inoko H. ERVK9, transposons and the evolution of MHC class I duplicons within the alpha-
963 block of the human and chimpanzee. *Cytogenetic and Genome Research*. 2005; 110(1-4):181–192. doi:
964 [10.1159/000084951](https://doi.org/10.1159/000084951).
- 965 **Kulski JK**, Anzai T, Shiina T, Inoko H. Rhesus macaque class I duplicon structures, organization, and evolu-
966 tion within the alpha block of the major histocompatibility complex. *Molecular Biology and Evolution*. 2004;
967 21(11):2079–2091. doi: [10.1093/molbev/msh216](https://doi.org/10.1093/molbev/msh216).
- 968 **Kulski JK**, Suzuki S, Shiina T. SNP-Density Crossover Maps of Polymorphic Transposable Elements and HLA
969 Genes Within MHC Class I Haplotype Blocks and Junction. *Frontiers in Genetics*. 2021; 11(January). doi:
970 [10.3389/fgene.2020.594318](https://doi.org/10.3389/fgene.2020.594318).
- 971 **Kupfermann H**, Satta Y, Takahata N, Tichy H, Klein J. Evolution of Mhc-DRB introns: Implications for the origin
972 of primates. *Journal of Molecular Evolution*. 1999; 48(6):663–674. doi: [10.1007/PL00006510](https://doi.org/10.1007/PL00006510).
- 973 **Lafont BAP**, Buckler-White A, Plishka R, Buckler C, Martin MA. Pig-tailed macaques [*Macaca nemestrina*] pos-
974 sess six MHC-E families that are conserved among macaque species: Implication for their binding to natural
975 killer receptors variants. *Immunogenetics*. 2004; 56(3):142–154. doi: [10.1007/s00251-004-0663-4](https://doi.org/10.1007/s00251-004-0663-4).
- 976 **Lampen MH**, Hassan C, Sluijter M, Geluk A, Dijkman K, Tjon JM, de Ru AH, van der Burg SH, van Veelen PA,
977 van Hall T. Alternative peptide repertoire of HLA-E reveals a binding motif that is strikingly similar to HLA-
978 A2. *Molecular Immunology*. 2013; 53(1-2):126–131. <http://dx.doi.org/10.1016/j.molimm.2012.07.009>, doi:
979 [10.1016/j.molimm.2012.07.009](https://doi.org/10.1016/j.molimm.2012.07.009).
- 980 **Lenormand C**, Bausinger H, Gross F, Signorino-Gelo F, Koch S, Peressin M, Fricker D, Cazenave JP, Bieber T,
981 Hanau D, de la Salle H, Tourne S. HLA-DQA2 and HLA-DQB2 Genes Are Specifically Expressed in Human
982 Langerhans Cells and Encode a New HLA Class II Molecule . *The Journal of Immunology*. 2012; 188(8):3903–
983 3911. doi: [10.4049/jimmunol.1103048](https://doi.org/10.4049/jimmunol.1103048).
- 984 **Li WH**, Gu Z, Wang H, Nekrutenko A. Evolutionary analyses of the human genome. *Nature*. 2001; 409(6822):847–
985 849. doi: [10.1038/35057039](https://doi.org/10.1038/35057039).
- 986 **Liao WW**, Asri M, Ebler J, Doerr D, Haukness M, Hickey G, Lu S, Lucas JK, Monlong J, Abel HJ, Buonaiuto S,
987 Chang XH, Cheng H, Chu J, Colonna V, Eizenga JM, Feng X, Fischer C, Fulton RS, Garg S, et al. A draft human
988 pangenome reference. *Nature*. 2023; 617(7960):312–324. doi: [10.1038/s41586-023-05896-x](https://doi.org/10.1038/s41586-023-05896-x).
- 989 **Lugo JS**, Cadavid LF. Patterns of MHC-G-like and MHC-B diversification in new world monkeys. *PLoS ONE*. 2015;
990 10(6):1–20. doi: [10.1371/journal.pone.0131343](https://doi.org/10.1371/journal.pone.0131343).
- 991 **Maccari G**, Robinson J, Ballingall K, Guethlein LA, Grimholt U, Kaufman J, Ho CS, de Groot NG, Flice P, Bontrop
992 RE, Hammond JA, Marsh SGE. IPD-MHC 2.0: an improved inter-species database for the study of the major
993 histocompatibility complex. *Nucleic Acids Research*. 2017 1; 45(D1):D860–D864. <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1050>, doi: [10.1093/nar/gkw1050](https://doi.org/10.1093/nar/gkw1050).

- 995 **Maccari G**, Robinson J, Hammond JA, Marsh SGE. The IPD Project: a centralised resource for the study of
996 polymorphism in genes of the immune system. *Immunogenetics*. 2020 2; 72(1-2):49–55. <http://link.springer.com/10.1007/s00251-019-01133-w>, doi: 10.1007/s00251-019-01133-w.
- 998 **Maibach V**, Hans JB, Hvilsom C, Marques-Bonet T, Vigilant L. MHC class I diversity in chimpanzees and bonobos.
999 *Immunogenetics*. 2017; 69(10):661–676. doi: 10.1007/s00251-017-0990-x.
- 1000 **Mao Y**, Harvey WT, Porubsky D, Munson KM, Hoekzema K, Lewis AP, Audano PA, Rozanski A, Yang X, Zhang S,
1001 Yoo DA, Gordon DS, Fair T, Wei X, Logsdon GA, Haukness M, Dishuck PC, Jeong H, del Rosario R, Bauer VL, et al.
1002 Structurally divergent and recurrently mutated regions of primate genomes. *Cell*. 2024; 187(6):1547–1562.
1003 <https://doi.org/10.1016/j.cell.2024.01.052>, doi: 10.1016/j.cell.2024.01.052.
- 1004 **Marsh SGE**, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernández-Viña M, Geraghty DE,
1005 Holdsworth R, Hurley CK, Lau M, Lee KW, MacH B, Maiers M, Mayr WR, Müller CR, Parham P, Petersdorf
1006 EW, Sasazuki T, Strominger JL, et al. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*.
1007 2010; 75(4):291–455. doi: 10.1111/j.1399-0039.2010.01466.x.
- 1008 **Mayer WE**, Jonker M, Klein D, Ivanyi P, van Seventer G, Klein J. Nucleotide sequences of chimpanzee MHC
1009 class I alleles: evidence for trans-species mode of evolution. *The EMBO journal*. 1988; 7(9):2765–2774. doi:
1010 [10.1002/j.1460-2075.1988.tb03131.x](https://doi.org/10.1002/j.1460-2075.1988.tb03131.x).
- 1011 **Messer G**, Zemmour J, Orr HT, Parham P, Weiss EH, Girdlestone J. HLA-J, a second inactivated class I HLA
1012 gene related to HLA-G and HLA-A. Implications for the evolution of the HLA-A-related genes. *The Journal of*
1013 *Immunology*. 1992; 148(12):4043–4053. doi: 10.4049/jimmunol.148.12.4043.
- 1014 **Moffett-King A**. Natural killer cells and pregnancy. *Nature Reviews Immunology*. 2002; 2(9):656–663. doi:
1015 [10.1038/nri886](https://doi.org/10.1038/nri886).
- 1016 **Müller NF**, Bouckaert RR. Adaptive metropolis-coupled MCMC for BEAST 2. *PeerJ*. 2020; 8:1–16. doi:
1017 [10.7717/peerj.9473](https://doi.org/10.7717/peerj.9473).
- 1018 **Neefjes J**, Jongsma MLM, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class
1019 II antigen presentation. *Nature Reviews Immunology*. 2011 12; 11(12):823–836. [http://www.nature.com/](http://www.nature.com/articles/nri3084)
1020 [articles/nri3084](http://www.nature.com/articles/nri3084), doi: 10.1038/nri3084.
- 1021 **Neehus AL**, Wistuba J, Ladas N, Eiz-Vesper B, Schlatt S, Müller T. Gene conversion of the major histocompati-
1022 bility complex class I C_α-G in common marmosets (*Callithrix jacchus*). *Immunology*. 2016; 149(3):343–352.
1023 doi: 10.1111/imm.12652.
- 1024 **Nei M**, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate
1025 immune system. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;
1026 94(15):7799–7806. doi: 10.1073/pnas.94.15.7799.
- 1027 **Nei M**, Rooney AP. Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*.
1028 2005; 39:121–152. doi: 10.1146/annurev.genet.39.073003.112240.
- 1029 **Nicholas RE**, Sandstrom K, Anderson JL, Smith WR, Wetzel M, Banerjee P, Janaka SK, Evans DT. KIR3DL05 and
1030 KIR3DS02 Recognition of a Nonclassical MHC Class I Molecule in the Rhesus Macaque Implicated in Pregnancy
1031 Success. *Frontiers in Immunology*. 2022; 13(March):1–12. doi: 10.3389/fimmu.2022.841136.
- 1032 **Okano M**, Miyamae J, Suzuki S, Nishiya K, Katakura F, Kulski JK, Moritomo T, Shiina T. Identification of Novel
1033 Alleles and Structural Haplotypes of Major Histocompatibility Complex Class I and DRB Genes in Domestic Cat
1034 (*Felis catus*) by a Newly Developed NGS-Based Genotyping Method. *Frontiers in Genetics*. 2020; 11(July):1–15.
1035 doi: 10.3389/fgene.2020.00750.
- 1036 **Olsson N**, Jiang W, Adler LN, Mellins ED, Elias JE. Tuning DO:DM Ratios Modulates MHC Class II Immunopep-
1037 tidomes. *Molecular and Cellular Proteomics*. 2022; 21(3):100204. [https://doi.org/10.1016/j.mcpro.2022.](https://doi.org/10.1016/j.mcpro.2022.100204)
1038 [100204](https://doi.org/10.1016/j.mcpro.2022.100204), doi: 10.1016/j.MCPRO.2022.100204.
- 1039 **Otting N**, de Groot NG, Bontrop RE. Evolution of HLA-F and its orthologues in primate species: a complex tale
1040 of conservation, diversification and inactivation. *Immunogenetics*. 2020; 72(9-10):475–487. [https://doi.org/](https://doi.org/10.1007/s00251-020-01187-1)
1041 [10.1007/s00251-020-01187-1](https://doi.org/10.1007/s00251-020-01187-1), doi: 10.1007/s00251-020-01187-1.
- 1042 **Paganini J**, Abi-Rached L, Gouret P, Pontarotti P, Chiaroni J, Di Cristofaro J. HLA-Ib worldwide genetic diversity:
1043 New HLA-H alleles and haplotype structure description. *Molecular Immunology*. 2019; 112(May):40–50. [https://](https://doi.org/10.1016/j.molimm.2019.04.017)
1044 doi.org/10.1016/j.molimm.2019.04.017, doi: 10.1016/j.molimm.2019.04.017.

- 1045 **Parham P**, Moffett A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and
1046 human evolution. *Nature Reviews Immunology*. 2013; 13(2):133–144. doi: 10.1038/nri3370.
- 1047 **Pierini F**, Lenz TL. Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection.
1048 *Molecular Biology and Evolution*. 2018 9; 35(9):2145–2158. [https://academic.oup.com/mbe/article/35/9/2145/](https://academic.oup.com/mbe/article/35/9/2145/5034935)
1049 [5034935](https://doi.org/10.1093/molbev/msy116), doi: 10.1093/molbev/msy116.
- 1050 **Piontkivska H**, Nei M. Birth-and-death evolution in primate MHC class I genes: Divergence time estimates.
1051 *Molecular Biology and Evolution*. 2003; 20(4):601–609. doi: 10.1093/molbev/msg064.
- 1052 **Radley E**, Alderton RP, Kelly A, Trowsdale J, Beck S. Genomic organization of HLA-DMA and HLA-DMB. Compari-
1053 son of the gene organization of all six class II families in the human major histocompatibility complex. *Journal*
1054 *of Biological Chemistry*. 1994; 269(29):18834–18838. [http://dx.doi.org/10.1016/S0021-9258\(17\)32242-1](http://dx.doi.org/10.1016/S0021-9258(17)32242-1), doi:
1055 10.1016/S0021-9258(17)32242-1.
- 1056 **Radwan J**, Babik W, Kaufman J, Lenz TL, Winternitz J. Advances in the Evolutionary Understanding of MHC
1057 Polymorphism. *Trends in Genetics*. 2020; 36(4):298–311. doi: 10.1016/j.tig.2020.01.008.
- 1058 **Rambaut A**, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior Summarization in Bayesian Phylogenetics
1059 Using Tracer 1.7. *Systematic Biology*. 2018 9; 67(5):901–904. [https://academic.oup.com/sysbio/article/67/5/](https://academic.oup.com/sysbio/article/67/5/901/4989127)
1060 [901/4989127](https://doi.org/10.1093/sysbio/syy032), doi: 10.1093/sysbio/syy032.
- 1061 **Riegert P**, Wanner V, Bahram S. Genomics, Isoforms, Expression, and Phylogeny of the MHC Class I-Related
1062 MR1 Gene . *The Journal of Immunology*. 1998; 161(8):4066–4077. doi: 10.4049/jimmunol.161.8.4066.
- 1063 **Robinson J**, Barker DJ, Marsh SGE. 25 years of the IPD-IMGT/HLA Database. *Hla*. 2024; 103(6):e15549. doi:
1064 [10.1111/tan.15549](https://doi.org/10.1111/tan.15549).
- 1065 **Salomonsen J**, Marston D, Avila D, Bumstead N, Johansson B, Juul-Madsen H, Olesen GD, Riegert P, Skjødt K,
1066 Vainio O, Wiles MV, Kaufman J. The properties of the single chicken MHC classical class II α chain (B-LA) gene
1067 indicate an ancient origin for the DR/E-like isotype of class II molecules. *Immunogenetics*. 2003; 55(9):605–
1068 614. doi: 10.1007/s00251-003-0620-7.
- 1069 **Satta Y**, Mayer WE, Klein J. Evolutionary relationship of HLA-DRB genes inferred from intron sequences.
1070 *Journal of Molecular Evolution*. 1996 6; 42(6):648–657. [http://link.springer.com/10.1007/](http://link.springer.com/10.1007/F02338798) [F02338798](https://doi.org/10.1007/BF02338798), doi:
1071 10.1007/BF02338798.
- 1072 **Sawai H**, Kawamoto Y, Takahata N, Satta Y. Evolutionary Relationships of Major Histocompatibility Complex
1073 Class I Genes in Simian Primates. *Genetics*. 2004; 166(4):1897–1907. doi: 10.1534/genetics.166.4.1897.
- 1074 **Sawyer SA**, GENECONV: a computer package for the statistical detection of gene conversion. Distributed by
1075 the author, Department of Mathematics, Washington University in St. Louis; 1999. [http://www.math.wustl.](http://www.math.wustl.edu/~sawyer)
1076 [edu/~sawyer](http://www.math.wustl.edu/~sawyer).
- 1077 **Schulze MSED**, Wucherpfennig KW. The mechanism of HLA-DM induced peptide exchange in the MHC
1078 class II antigen presentation pathway. *Current Opinion in Immunology*. 2012; 24(1):105–111. doi:
1079 [10.1016/j.coi.2011.11.004](https://doi.org/10.1016/j.coi.2011.11.004).
- 1080 **Shiina T**, Blancher A, Inoko H, Kulski JK. Comparative genomics of the human, macaque and mouse major
1081 histocompatibility complex. *Immunology*. 2017; 150(2):127–138. doi: 10.1111/imm.12624.
- 1082 **Shiina T**, Kono A, Westphal N, Suzuki S, Hosomichi K, Kita YF, Roos C, Inoko H, Walter L. Comparative genome
1083 analysis of the major histocompatibility complex (MHC) class I B/C segments in primates elucidated by ge-
1084 nomic sequencing in common marmoset (*Callithrix jacchus*). *Immunogenetics*. 2011; 63(8):485–499. doi:
1085 10.1007/s00251-011-0526-8.
- 1086 **Shiina T**, Tamiya G, Oka A, Takishima N, Yamagata T, Kikkawa E, Iwata K, Tomizawa M, Okuaki N, Kuwano Y,
1087 Watanabe K, Fukuzumi Y, Itakura S, Sugawara C, Ono A, Yamazaki M, Tashiro H, Ando A, Ikemura T, Soeda E,
1088 et al. Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I
1089 region. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96(23):13282–
1090 13287. doi: 10.1073/pnas.96.23.13282.
- 1091 **Shiina T**, Yamada Y, Aarnink A, Suzuki S, Masuya A, Ito S, Ido D, Yamanaka H, Iwatani C, Tsuchiya H, Ishigaki
1092 H, Itoh Y, Ogasawara K, Kulski JK, Blancher A. Discovery of novel MHC-class I alleles and haplotypes in Fil-
1093 ipino cynomolgus macaques (*Macaca fascicularis*) by pyrosequencing and Sanger sequencing: Mafa-class I
1094 polymorphism. *Immunogenetics*. 2015; 67(10):563–578. doi: 10.1007/s00251-015-0867-9.

- 1095 **Slierendregt BL**, van Noort JT, Bakas RM, Otting N, Jonker M, Bontrop RE. Evolutionary stability of transspecies
1096 major histocompatibility complex class II DRB lineages in humans and rhesus monkeys. *Human Immunol-*
1097 *ogy*. 1992 9; 35(1):29–39. <https://linkinghub.elsevier.com/retrieve/pii/0198885992900922>, doi: 10.1016/0198-
1098 8859(92)90092-2.
- 1099 **Takahashi K**, Rooney AP, Nei M. Origins and divergence times of mammalian class II MHC gene clusters. *Journal*
1100 *of Heredity*. 2000; 91(3):198–204. doi: [10.1093/jhered/91.3.198](https://doi.org/10.1093/jhered/91.3.198).
- 1101 **Thornton JW**, Desalle R. *G Ene F Amily E Volution and H Omology*. New York. 2000; 1(48):41–73. [http://www.](http://www.ncbi.nlm.nih.gov/pubmed/11701624)
1102 [ncbi.nlm.nih.gov/pubmed/11701624](http://www.ncbi.nlm.nih.gov/pubmed/11701624).
- 1103 **Urvater JA**, Otting N, Loehrke JH, Rudersdorf R, Slukvin II, Piekarczyk MS, Golos TG, Hughes AL, Bontrop RE,
1104 **Watkins DI**. Mamu-I: A Novel Primate MHC Class I B-Related Locus with Unusually Low Variability. *The*
1105 *Journal of Immunology*. 2000; 164(3):1386–1398. <https://doi.org/10.4049/jimmunol.164.3.1386>.
- 1106 **Van Der Wiel MK**, Otting N, De Groot NG, Doxiadis GGM, Bontrop RE. The repertoire of MHC class I genes
1107 in the common marmoset: Evidence for functional plasticity. *Immunogenetics*. 2013; 65(12):841–849. doi:
1108 10.1007/s00251-013-0732-7.
- 1109 **Van Lith M**, McEwen-Smith RM, Benham AM. HLA-DP, HLA-DQ, and HLA-DR have different requirements for
1110 invariant chain and HLA-DM. *Journal of Biological Chemistry*. 2010; 285(52):40800–40808. [http://dx.doi.org/](http://dx.doi.org/10.1074/jbc.M110.148155)
1111 [10.1074/jbc.M110.148155](http://dx.doi.org/10.1074/jbc.M110.148155), doi: [10.1074/jbc.M110.148155](https://doi.org/10.1074/jbc.M110.148155).
- 1112 **Vollmers S**, Lobermeyer A, Körner C. The new kid on the block: Hla-c, a key regulator of natural killer cells in
1113 viral immunity. *Cells*. 2021; 10(11). doi: [10.3390/cells10113108](https://doi.org/10.3390/cells10113108).
- 1114 **Walter L**. Nomenclature report on the major histocompatibility complex genes and alleles of the laboratory
1115 rat (*Rattus norvegicus*). *Immunogenetics*. 2020; 72(1-2):5–8. doi: [10.1007/s00251-019-01131-y](https://doi.org/10.1007/s00251-019-01131-y).
- 1116 **Watkins DI**, Chen ZW, Garber TL, Hughes AL, Letvin NL. Segmental exchange between MHC class I genes in a
1117 higher primate: recombination in the gorilla between the ancestor of a human non-functional gene and an
1118 A locus gene. *Immunogenetics*. 1991; 34(3):185–191. doi: [10.1007/BF00205822](https://doi.org/10.1007/BF00205822).
- 1119 **Watkins DI**, Chen ZW, Hughes AL, Evans MG, Tedder TF, Letvin NL. Evolution of the MHC class I genes of a New
1120 World primate from ancestral homologues of human non-classical genes. *Nature*. 1990; 346(6279):60–63.
1121 doi: [10.1038/346060a0](https://doi.org/10.1038/346060a0).
- 1122 **Welsh R**, Song N, Sadegh-Nasseri S. What to do with HLA-DO/H-2O two decades later? *Immunogenetics*. 2019;
1123 71(3):189–196. doi: [10.1007/s00251-018-01097-3](https://doi.org/10.1007/s00251-018-01097-3).
- 1124 **Welsh RA**, Sadegh-Nasseri S. The love and hate relationship of HLA-DM/DO in the selection of immunodomi-
1125 nant epitopes. *Current Opinion in Immunology*. 2020; 64(Mhc II):117–123. [https://doi.org/10.1016/j.coi.2020.](https://doi.org/10.1016/j.coi.2020.05.007)
1126 [05.007](https://doi.org/10.1016/j.coi.2020.05.007), doi: [10.1016/j.coi.2020.05.007](https://doi.org/10.1016/j.coi.2020.05.007).
- 1127 **van der Wiel MKH**, Doxiadis GGM, de Groot N, Otting N, de Groot NG, Poirier N, Blanco G, Bontrop RE. MHC
1128 class I diversity of olive baboons (*Papio anubis*) unravelled by next-generation sequencing. *Immunogenetics*.
1129 2018; 70(7):439–448. doi: [10.1007/s00251-018-1053-7](https://doi.org/10.1007/s00251-018-1053-7).
- 1130 **Wilming LG**, Hart EA, Coggill PC, Horton R, Gilbert JGR, Clee C, Jones M, Lloyd C, Palmer S, Sims S, Whitehead
1131 S, Wiley D, Beck S, Harrow JL. Sequencing and comparative analysis of the gorilla MHC genomic sequence.
1132 *Database*. 2013; 2013. doi: [10.1093/database/bat011](https://doi.org/10.1093/database/bat011).
- 1133 **Wroblewski EE**, Guethlein LA, Norman PJ, Li Y, Shaw CM, Han AS, Ndjanga JBN, Ahuka-Mundeke S, Georgiev AV,
1134 Peeters M, Hahn BH, Parham P. Bonobos Maintain Immune System Diversity with Three Functional Types
1135 of MHC-B. *The Journal of Immunology*. 2017; 198(9):3480–3493. doi: [10.4049/jimmunol.1601955](https://doi.org/10.4049/jimmunol.1601955).
- 1136 **Wroblewski EE**, Parham P, Guethlein LA. Two to tango: Co-evolution of hominid natural killer cell receptors
1137 and MHC. *Frontiers in Immunology*. 2019; 10(FEB):1–22. doi: [10.3389/fimmu.2019.00177](https://doi.org/10.3389/fimmu.2019.00177).
- 1138 **Wu CH**, Suchard MA, Drummond AJ. Bayesian Selection of Nucleotide Substitution Models and Their Site
1139 Assignments. *Molecular Biology and Evolution*. 2013 3; 30(3):669–688. [https://academic.oup.com/mbe/](https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/mss258)
1140 [article-lookup/doi/10.1093/molbev/mss258](https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/mss258), doi: [10.1093/molbev/mss258](https://doi.org/10.1093/molbev/mss258).
- 1141 **Yeager M**, Hughes AL. Evolution of the mammalian MHC: Natural selection, recombination, and convergent
1142 evolution. *Immunological Reviews*. 1999; 167:45–58. doi: [10.1111/j.1600-065X.1999.tb01381.x](https://doi.org/10.1111/j.1600-065X.1999.tb01381.x).
- 1143 **Zhou Y**, Song L, Li H. Full resolution HLA and KIR gene annotations for human genome assemblies. *Genome*
1144 *Research*. 2024; p. gr.278985.124. doi: [10.1101/gr.278985.124](https://doi.org/10.1101/gr.278985.124).