

Análise molecular evolutiva do gene *MHC-DRB1* em primatas

Douglas Marinho Costa¹, Felesmar Rodrigues¹, Juliana de Moraes¹ e Karen Krejcik¹

¹Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo, São José dos Campos, Brasil

Resumo. O *MHC* é um complexo gênico de grande importância na resposta imunológica e apresenta alto grau de polimorfismo, sendo muito versátil diante dos desafios impostos pelas infecções que afetam os animais de maneira frequente. O *MHC-DRB1* é o gene responsável por codificar a cadeia β do *MHC* de classe II, que responde a infecções intravesiculares e as apresenta aos linfócitos T CD4. Inúmeras doenças autoimunes já foram ligadas a alelos do *MHC-DRB1*. Dado a importância desse locus gênico, procuramos analisar sua evolução ao longo dos primatas e avaliar como o seu desenvolvimento pode contribuir para a apresentação dessas doenças.

Palavras-chave: evolução, filogenia, imunologia, *MHC*, primatas

Evolutionary molecular analysis of the MHC-DRB1 gene in primates

Douglas Marinho Costa¹, Felesmar Rodrigues¹, Juliana de Moraes¹ e Karen Krejcik¹

¹Science and Technology Institute, Sao Paulo Federal University, Sao Jose dos Campos, Brazil

Abstract. MHC is a genetic complex of great importance in immune response and presents a high degree of polymorphism. It is also very versatile in front of the challenges imposed by the frequent infections that affect animals. MHC-DRB1 is the gene responsible for encoding the class II MHC β chain, which responds to intravesicular infections and presents them to CD4 T lymphocytes. Also, numerous autoimmune diseases have already been linked to MHC-DRB1 alleles. Given the importance of this genetic locus, this project analyses its evolution along the primates and evaluate how its development could contribute to the presentation of these diseases.

Keywords: evolution, immunology, MHC, phylogeny, primates

INTRODUÇÃO

O *MHC* (*Major Histocompatibility Complex*) é um complexo gênico multi-espécie altamente polimórfico e de expressão codominante responsável por codificar uma série de genes encarregados da resposta imune e importantes na diferenciação entre antígenos próprios e não-próprios (GENETICS HOME REFERENCE, 2017).

Esse complexo está localizado no cromossomo 6 na posição p21.32 em humanos, como mostrado na Figura 1.

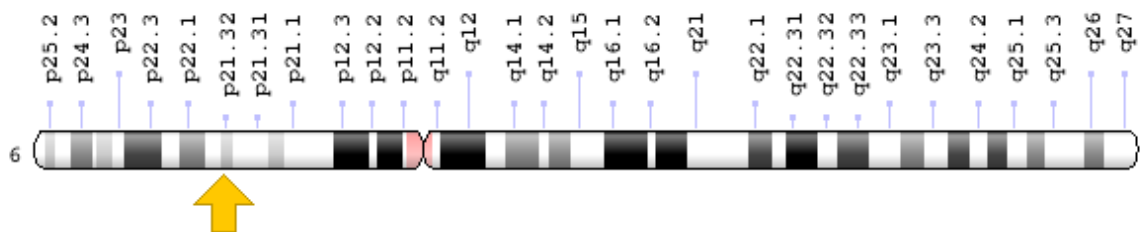


Figura 1: Localização do complexo do *MHC* em humanos (GENETICS HOME REFERENCE, 2017)

A principal função do *MHC* está em apresentar antígenos para as células do sistema imune (ABBAS *et al*, 2016). A efetividade de seu mecanismo de reconhecimento repousa sobre o seu polimorfismo e sua capacidade de se adaptar de maneira dinâmica em resposta às infecções constantes enfrentadas pelos organismos (ALVAREZ-BUSTO *et al*, 2007).

O *MHC-DRB1* é o loci gênico do *MHC* responsável pela codificação da cadeia β da molécula de *MHC* de classe II, responsável por apresentar antígenos intravesiculares para as células T CD4 (ABBAS *et al*, 2016). A polimorfia do *MHC* II ocorre na porção amino-terminal da cadeia β (OLIVEIRA *et al*, 2009). A estrutura da molécula é apresentada na Figura 2.

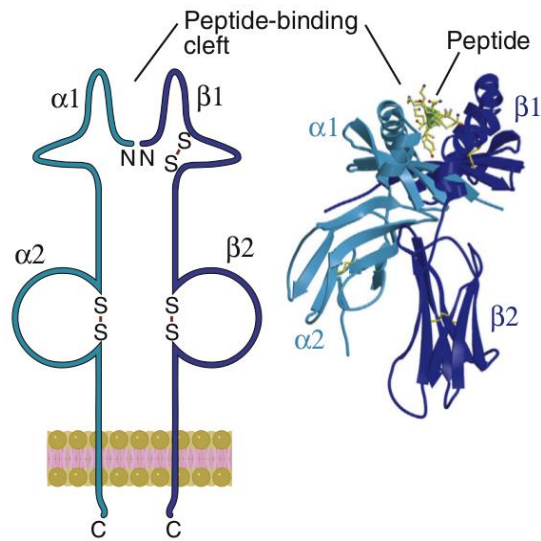


Figura 2: Estrutura da molécula de *MHC II* com a cadeia β codificada pelo *MHC-DRB1* (ABBAS *et al*, 2016)

Estudos já mostraram correlação entre alelos de *MHC-DRB1* e o risco de diversas doenças autoimunes, como diabetes do tipo I (NOBLE *et al*, 2011; STECK *et al*, 2011), artrite reumatóide (STAHL *et al*, 2010; EYRE *et al*, 2012; RAYCHAUDHURI *et al*, 2012), esclerose múltipla (ALCINA *et al*, 2012; CREE *et al*, 2014), entre outras.

O estudo da estrutura de genes imunológicos e sua evolução pode nos levar a um entendimento de como as patologias surgiram e caminhos para curá-las, ao se verificar como esses genes se comportam em diferentes espécies e como as doenças são apresentadas.

MÉTODOS E RESULTADOS

CONTEÚDO CG

Quanto maior o conteúdo CG do DNA, mais estável será o gene, pois o conteúdo CG possui três ligações de hidrogênio, enquanto o conteúdo AT possui apenas duas ligações, sendo este o mais instável (ALBERTS *et al*, 2010). O conteúdo CG¹ do gene *MHC-DRB1* foi calculado e este valor é de aproximadamente 8%, mostrando uma grande instabilidade do gene, contudo, por tratar-se de uma região polimórfica e de alta incidência de mutações, o valor encontrado é coerente com o que é explicitado na literatura.

Na Figura 3 é possível observar a distribuição dos quatro nucleotídeos (Figura 3.A) e dos conteúdos CG e AT do gene *MHC-DRB1* (Figura 3.B). Observa-se que há regiões

¹ Para o cálculo do conteúdo CG foi utilizado a DNA/RNA GC Content Calculator, disponível em: <<http://www.endmemo.com/bio/gc.php>>. Acesso: 13 de novembro de 2017.

específicas de predomínio do conteúdo CG no início da sequência, indicando a prevalência de regiões mais estáveis e não codificantes. Percebe-se na Figura 3.B que há um predomínio do conteúdo AT em todo gene, o que é equivalente com as informações encontradas no NCBI².

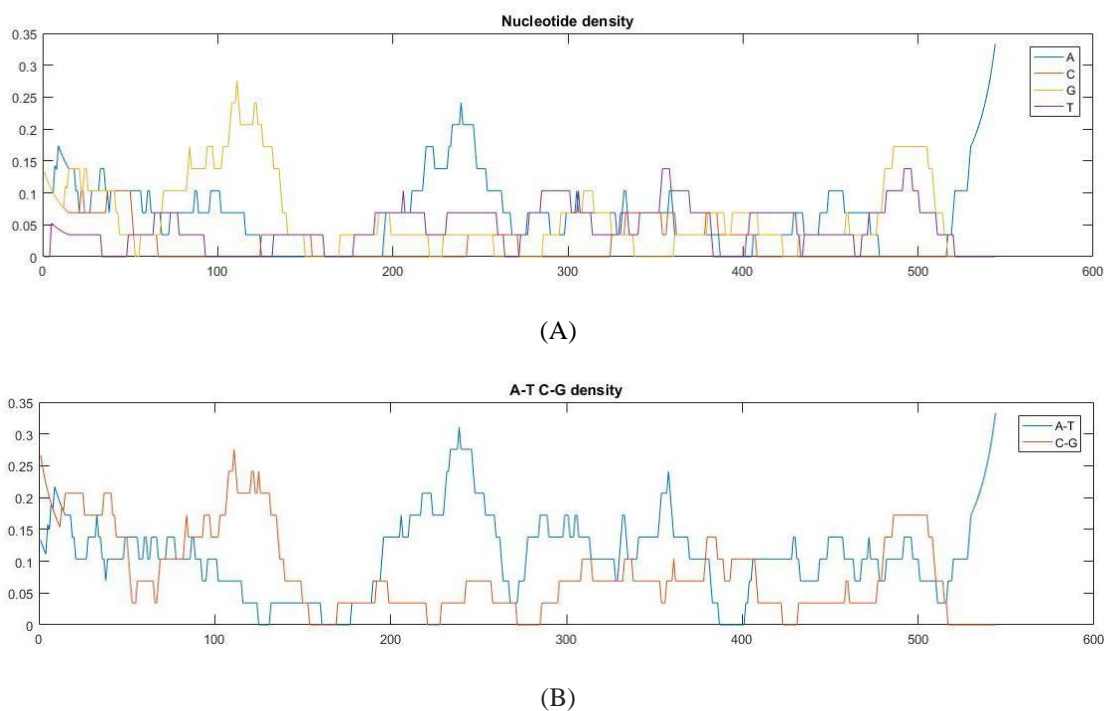


Figura 3: Distribuição dos quatro nucleotídeos em (A) e distribuição dos conteúdos AT e GC em (B)

ALINHAMENTO DE ORTÓLOGOS DO GENE *MHC-DRB1*

Para o alinhamento foram escolhidos genes ortólogos ao gene *MHC-DRB1* de sete espécies diferentes do grupo dos primatas (**Tabela 1**). Em todas as espécies havia a presença de inúmeros alelos diferentes, no entanto, quando possível, o alelo *DRB1 *03:01* foi escolhido, por também estar presente no ser humano (DA SILVA, 2010). A única exceção foi o primata gibão, por não haver dados disponíveis do alelo *DRB1 *03:01*, portanto foi usado o alelo *DRB1 *04:01* para o sequenciamento, pelo critério de proximidade com o alelo *DRB1 *03:01*.

Estes alelos foram escolhidos também por haver excesso de informação e uma limitação muito grande no equipamento utilizado para fazer o alinhamento. As sequências de aminoácidos dos alelos foram obtidas no banco de dados *IPD MHC NHP* (<http://www.ebi.ac.uk/ipd/mhc/nhp/>), exclusiva de primatas não humanos. A partir destes dados foram obtidos o alinhamento e a criação árvore filogenética. O código utilizado tanto para a

² Informações completas podem ser obtidas em: <<https://www.ncbi.nlm.nih.gov/gene/3123>>. Acesso: 13 de novembro de 2017.

obtenção dos dados e os gráficos do conteúdo CG, quanto para o alinhamento e a criação da árvore genética está no Anexo A.

Tabela 1. Relação entre espécie, nome comum e alelo escolhido das proteínas *DRB1* ortólogas do gene *MHC DRB1*

Espécie	Nome comum	Alelo escolhido
<i>Gorilla gorilla</i>	Gorila	DRB1*03:01
<i>Pan paniscus</i>	Bonobo	DRB1*03:01
<i>Pan troglodytes</i>	Chimpanzé	DRB1*03:01
<i>Hylobates moloch</i>	Gibão	DRB1*04:01
<i>Macaca fascicularis</i>	Macaco-cinomolgo	DRB1*03:01
<i>Macaca mulatta</i>	Macaco Rhesus	DRB1*03:01
<i>Chlorocebus sabaues</i>	Macaco-verde	DRB1*03:01

O alinhamento múltiplo foi obtido através do software *MATLAB*® e o resultado parcial deste alinhamento está na Figura 4.

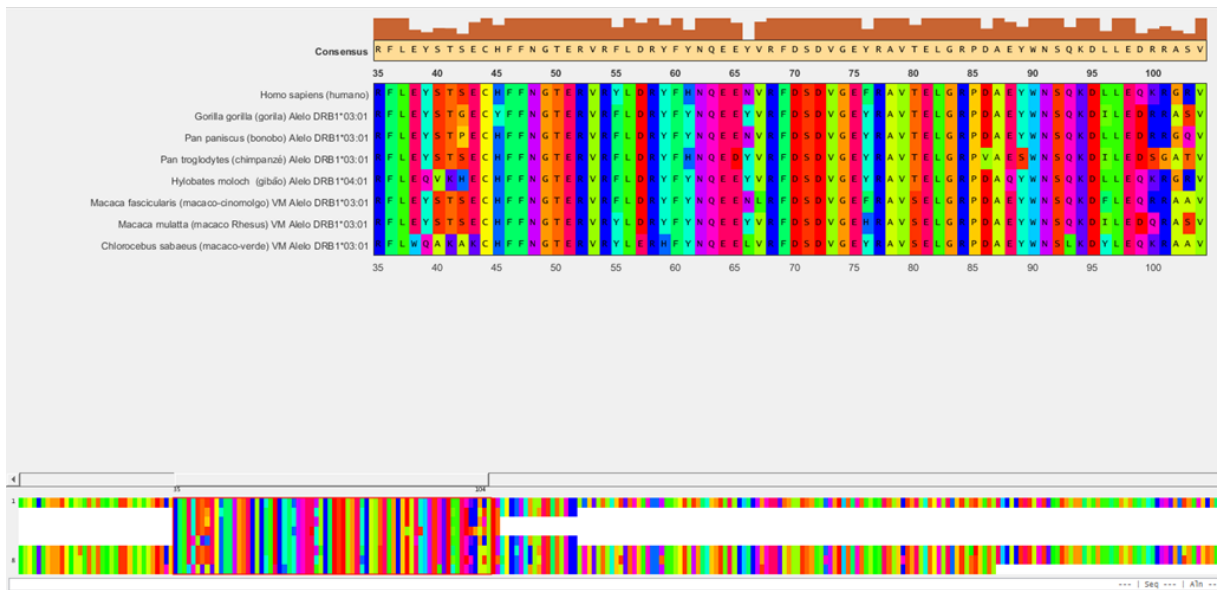


Figura 4: Imagem parcial do alinhamento dos ortólogos do gene *MHC-DRB1*

A maioria dos métodos de sequência múltipla tentam minimizar o número de inserções/deleções (gaps), para que o alinhamento seja o mais compacto possível. Estes gaps também são muito comuns em sequências produzidas recentemente e que são pobremente anotadas, podendo conter frameshifts, domínios errados ou splicing dos éxons não homólogos. Este foi o caso ocorrido com algumas proteínas dos genes ortólogos ao *MHC-DRB1* que foram escolhidos, pois atualmente o sequenciamento da estrutura *DRB1* ainda é pouco acurado (GROENEWEG, VOORTER & TILANUS, 2015).

DISCUSSÃO

PROPOSTA DE FILOGENIA

O estudo da filogenia permite inferir uma relação evolutiva entre as espécies a partir de dados moleculares e morfológicos. O alinhamento proposto neste trabalho teve como referência sete genes ortólogos ao *MHC-DRB1* de primatas não humanos.

A partir dos alinhamentos dos sete ortólogos foi possível construir a árvore filogenética que se encontra na Figura 5.

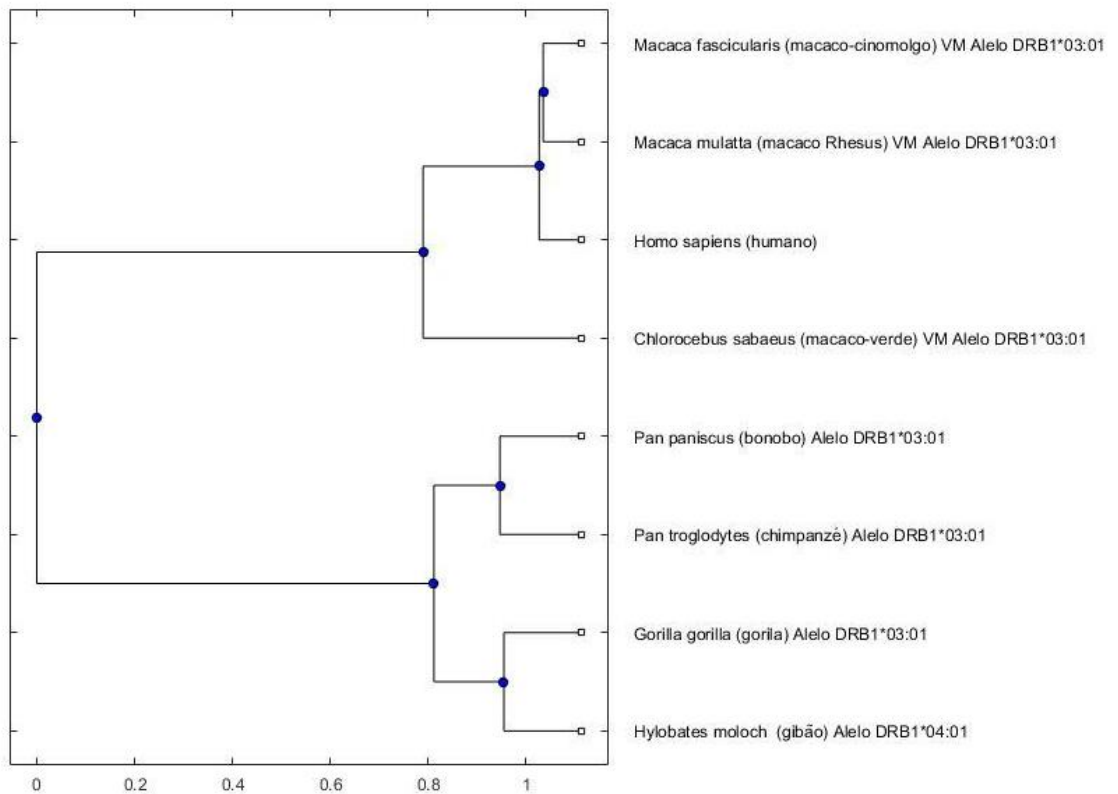


Figura 5: Árvore filogenética obtida a partir do alinhamento dos sete ortólogos utilizados neste trabalho

Analisando a filogenia proposta na Figura 5, percebe-se que há variações conflitantes com os dados da literatura, como é observado na Figura 6. Enquanto a literatura (Figura 6) mostra que há maior proximidade entre os humanos e os grandes símios, os dados obtidos através do alinhamentos dos ortólogos do gene *MHC-DRB1* mostram que há uma maior proximidade entre os humanos e os macacos do velho mundo, o que não é corroborado pela literatura.

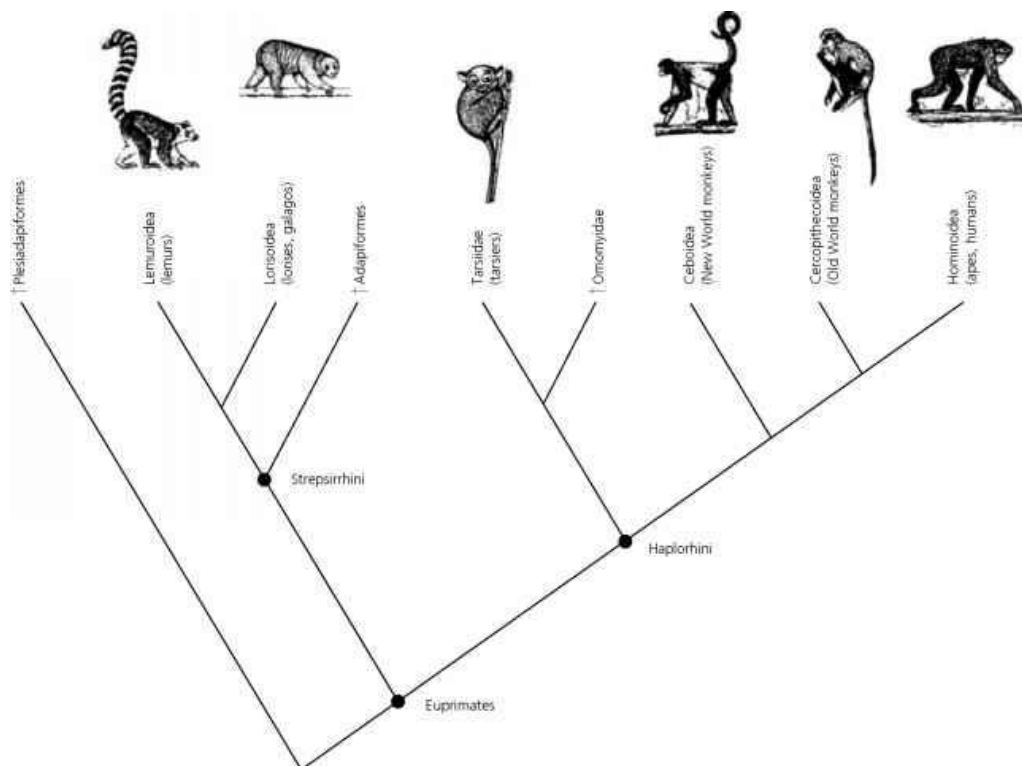


Figura 6: Cladograma dos primatas (adaptado de FLEAGLE, 1999)

O tempo de divergência evolutiva dos macacos do Velho Mundo e homínídeos foi calculado em aproximadamente 25 milhões de anos, enquanto dos grandes símios e dos seres humanos em aproximadamente 15 milhões de anos (DOXIADIS *et al*, 2012). Contudo, o gene *DRB1* é o gene mais polimórfico do locus de classe II (BERGSTROM & GYLLENSTEN, 1995) e isto infere diretamente na modificação do cladograma e em sua informação conflitante encontrados a partir do alinhamento proposto neste trabalho.

Sabe-se que se uma espécie ancestral é polimórfica segregando alelos A e B (Figura 7) e dividindo-se em duas linhagens descendentes, ambos os alelos podem ser retidos nas duas linhagens filhas. Se uma dessas linhagens se divide novamente relativamente em breve, então todas as três linhagens de espécies podem carregar ambos os alelos. Com o tempo, cada linhagem pode perder um ou mais alelos devido à deriva genética ou seleção natural. Neste caso, a espécie 1 retém o alelo A e a espécie 3 retém o alelo B. Neste segmento genômico, a espécie 2 se parecerá mais proximamente relacionada tanto com a espécie 1 ou 3, dependendo se irá reter o alelo A ou B. A retenção do alelo B significaria que este segmento genômico é coerente com a árvore filogenética mais comumente encontrada, no entanto, a retenção do alelo A levaria a uma discrepância. Análises de sequências genômicas de humanos, chimpanzés e gorilas indica que uma porção específica do genoma não é coerente com a árvore filogenética

mais comumente encontrada, na qual coloca os chimpanzés mais próximos evolutivamente dos humanos que os gorilas (ROGERS & GIBBS, 2014).

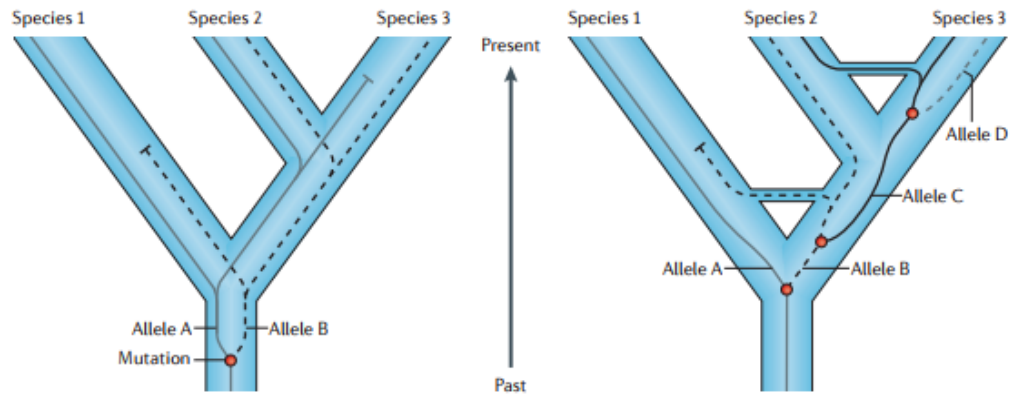


Figura 7: Separação incompleta da linhagem ancestral pode produzir uma discrepância na árvore filogenética para um gene ou segmento genômico específico (Adaptado de ROGERS & GIBBS, 2014)

A partir desta interpretação, pode-se então sugerir que o ancestral comum tanto dos macacos do Velho Mundo quanto dos hominídeos, que viveu há mais de 25 milhões de anos, já possuía um alto grau de polimorfismo no locus *DRB1*, como propomos no cladograma abaixo (Figura 8).

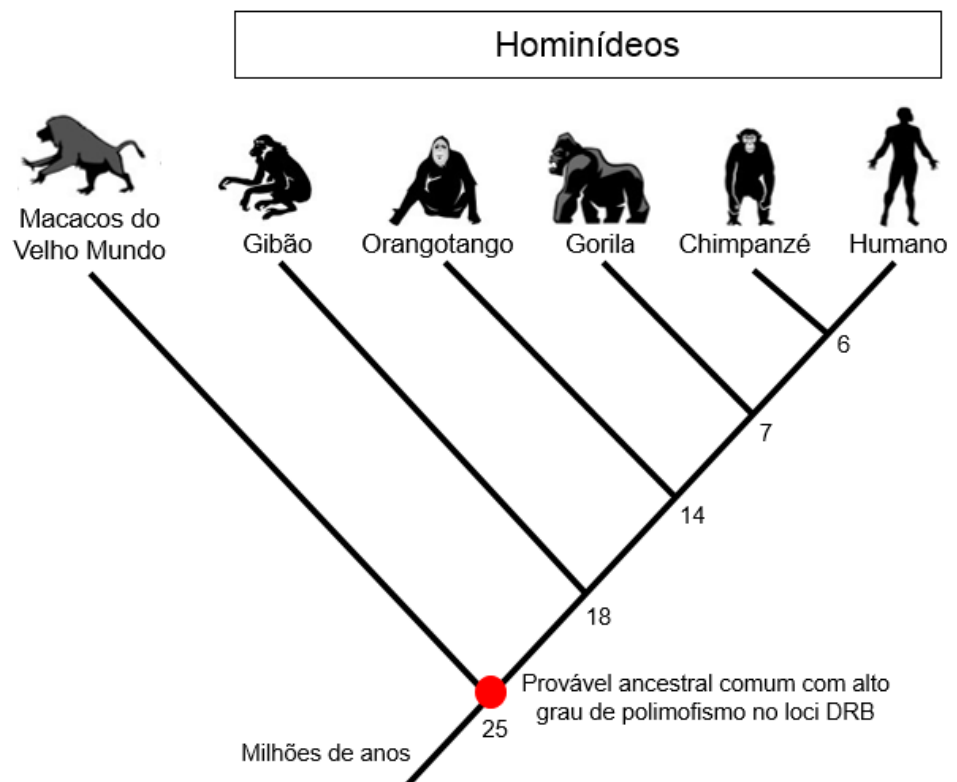


Figura 8: Cladograma proposto com o ancestral comum tanto dos macacos do Velho Mundo quanto dos homínídeos.

HIPÓTESES DE EVOLUÇÃO MOLECULAR DO GENE *MHC-DRB1*

O gene *DRB1* apresenta um polimorfismo de extrema importância para o sistema imunológico, na identificação de proteínas e moléculas próprias e não próprias, sendo uma proteína integrante do *MCH* de classe II, que é denominada cadeia β que se liga a outra proteína funcional denominada cadeia α , formando a proteína funcional chamada heterodímero de ligação ao antígeno relacionado D. O complexo *MHC* do tipo II desempenha papel fundamental no início de respostas imunológicas contra parasitas intracelulares, vírus e na auto-tolerância a antígenos próprios (ABBAS *et al*, 2012). Todavia, inúmeras doenças de caráter imunometabólicas vêm sendo associadas a variantes do gene *DRB1* envolvidas em resposta imunológicas exacerbadas ou falhas de função por diversos fatores ambientais e de caráter genético.

Durante a evolução das espécies, diversos eventos ocorreram como resposta a pressões evolutivas internas e pressões de co-evolução que ocasionaram a construção de diversos sistemas imunológicos característicos de cada grupo biológico. O desenvolvimento e refinamento do sistema imune foi modulado também por co-evolução a fatores de restrição celular, deriva gênica, metabolismo, aumento na complexidade reprodutiva entre outros (FLAJNIK & KASAHARA, 2010).

O sistema imune em mamíferos, segundo estudos na área de genômica, surgiu a cerca de 500 milhões de anos em peixes gnátostomados, sendo que esse sistema de defesa apresenta diversas moléculas, que não estão presentes em peixes agnatos. Este evento foi possível devido ao surgimento de um conjunto de genes denominado RAG e há eventos de duplicação do genoma (Figura 9). Essas descobertas sobre a evolução imunológica pode ajudar a elucidar como se sucedeu o surgimento de cada gene do sistema imune e assim identificar possíveis causas de patologias imunológicas, estudando grupos basais e espécies que possuem os genes e apresentam o fenótipo. O estudo da evolução imunológica é importante também para identificar anomalias fatais ou não para um maior entendimento do curso evolutivo da espécie estudada, sendo possível ser vista como uma perspectiva evolutiva mostrando que a anomalia não compromete a sobrevivência do indivíduo (DU PASQUIER, 1992).

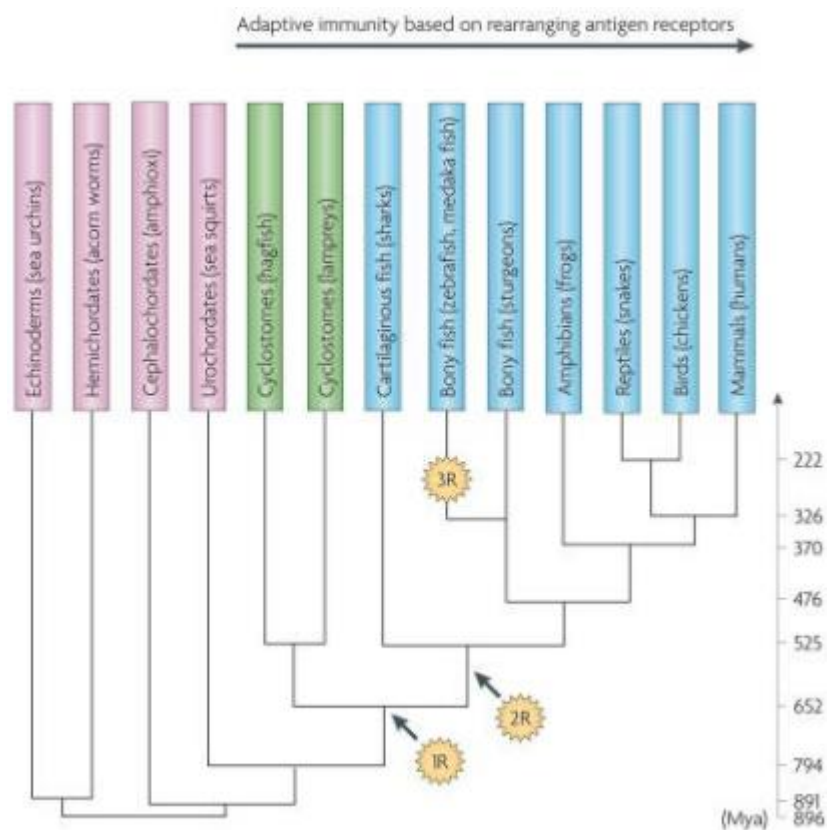


Figura 9: Uma visão geral da evolução do sistema imunológico. Moléculas restritas aos vertebrados com e sem mandíbula indicados em azul e verde respectivamente. A cor púrpura indica genes semelhantes ao gene RAG. 1R, 2R e 3R indicam eventos onde ocorreu rodadas de duplicação do genoma (Adaptado FLAJNIK & KASAHRA, 2010)

O complexo RAG é responsável pela codificação da enzima que desempenha o processo de rearranjo e recombinação de genes de imunoglobulinas e receptores de células linfocitárias, como de células T no processo de recombinação de *VDJ* (ABBAS *et al*, 2012).

Assim, é evidenciado que o aumento na complexidade genômica foi uma solução encontrada para adquirir mecanismos de proteção contra infecções. O *MHC* é um dos mecanismos desenvolvidos pelos vertebrados que lhe conferem uma barreira defensiva difícil de ser transpassada, caracterizado pela grande polimorfia, taxas de heterozigose maiores do que taxas esperadas sobre neutralidade e grande divergência nucleotídica (MURPHY *et al*, 2010).

Os primatas não humanos apresentam um sistema imunológico muito similar ao sistema imunológico humano, sendo algumas espécies com mapas genéticos na ordem de 98,7% de semelhança com o *Homo sapiens* (PRUFER *et al*, 2011). Conforme mostrado, os primatas não humanos têm em seu código genético genes homólogos aos dos seres humanos, demonstrando sua proximidade com ancestrais comuns.

O gene *DRBI* está presente nos primatas de maneira geral, com algumas diferenças em sua composição nucleotídica. Esse gene está relacionado a diversas doenças de caráter imunológico que é descrito como um fator importante para patogênese e eventos de intolerância. Em diversas patologias relacionadas ao metabolismo insulínico, como na diabetes do tipo 1, tem por causa inúmeros fatores, como infecções virais e fatores genéticos, e nos primatas estudados neste trabalho somente o gibão não apresenta o fenótipo da doença. Daí, constata-se que as diferenças nucleotídicas relacionadas as espécies estudadas, lhe conferem suscetibilidade para algumas doenças imunológicas e outras não.

No curso evolutivo, um ancestral comum tinha um gene *DBRI-X* e através de um processo de especiação, devido a mutações no código genético causada por eventos de inserção, deleção ou deslocamento no quadro de leitura, eventos no processo de recombinação genética na fusão de núcleos ou *crosstalk* na meiose, eventos de saltos transposômicos ou infecções virais (RIDLEY, 2006) , deram origem a duas espécies com um gene *DBRI-X1* e *DBRI-X2*. O indivíduo com o gene *DBRI-X1* sofreu as mesmas pressões evolutivas e surgiu duas espécies uma com um gene *DBRI-Y* e *DBRI-Z*. O indivíduo *DBRI-Y* sofreu pressões evolutivas e a partir de eventos de especiação surgiram as espécies bonobo e chimpanzé e o indivíduo *DBRI-Z* pelo mesmo mecanismo evolutivo deu origem às espécies Gorila e Gibão. Consequentemente, o indivíduo *DBRI-X2* sofreu o processo de especiação e deu origem a indivíduos *DBRI-H* e *DBRI-M*, onde o indivíduo *DBRI-M* deu origem a espécie macaco-verde e o indivíduo *DBRI-H* sofreu pressões evolutivas e surgiram indivíduos com genes e *DBRI-H1* e *DBRI-H2*. Os indivíduos *DBRI-H1* deram origem aos homínídeos que mais tarde deram origem aos humanos e os indivíduos *DBRI-H2* pelo mesmo mecanismo evolutivo deram origem às espécies *Macaca mulatta* e *Macaca fascicularis*.

CONCLUSÃO

Essa é apenas uma proposta de evolução se embasando somente no gene *DRBI*. Apesar de muita proximidade genética entre as espécies de primatas, podemos observar que a diversas doenças imunológicas afetam algumas espécies e outras não. Provavelmente isso ocorre por mecanismos de co-evolução, pois os microrganismos e vírus desenvolveram maneiras de evadir o sistema imune através de evasinas e mimetismo molecular, mostrando que apesar da simplicidade genômica, os patógenos têm estratégias refinadas para garantir a perpetuação do seu material genético e sua sobrevivência, e evidenciando a complexidade evolutiva que os

organismos desenvolveram para se proteger dos patógenos. Contudo, mesmo com esse aparato ainda ocorrem inúmeros falhas, que em sua maioria comprometem o sistema imune.

REFERÊNCIAS

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. *Imunologia celular e molecular*. 7. ed. Rio de Janeiro: Elsevier, 2012.

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. *Basic Immunology*. 5th ed. Missouri: Elsevier, 2016.

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. *Biologia Molecular da Célula*. Rio de Janeiro: Editora Artmed. 5ª Edição, 2010.

ALCINA, A. *et al.* Multiple sclerosis risk variant HLA-DRB1*1501 associates with high expression of DRB1 gene in different human populations. *PLoS One*, 2012; 7(1):e29819. doi:10.1371/journal.pone.0029819

ALVAREZ-BUSTO, J. *et al.* Diversity and evolution of the MHC-DRB1 gene in the two endemic Iberian subspecies of Pyrenean chamois, *Rupicapra pyrenaica*. *Heredity*, 2007; 99: 406-413. doi:0018-067X/07

BERGSTROM, T.; GYLLENSTEN, U. Evolution of Mhc Class II Polymorphism: The Rise and Fall of Class II Gene Function in Primates. *Immunological Reviews*, 1995; 145(1): 13-31. doi:10.1111/j.1600-065X.1995.tb00668.x

CREE, B.A.C. Multiple sclerosis genetics. *Handb Clin Neurol*, 2014; 122: 193-209. doi:10.106/B978-0-444-52001-2.00009-1

DOXIADIS, G. *et al.* Evolution of HLA-DRB Genes. *Molecular Biology and Evolution*. 2012; 29(12): 3843–3853. doi:10.1093/molbev/mss186

DU PASQUIER L. Origin and evolution of the vertebrate immune system. *APMIS*. 1992.

EYRE, S. *et al.* High density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet*, 2012; 44(12): 1336–1340. doi:10.1038/ng.2462

FLAJNIK, M. F., KASAHARA, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews*, 2010.

- GROENEWEG, M; VOORTER, C.E.; TILANUS, M.G.J. Genomic structure of the HLA complex on the short arm of chromosome 6. Conferência London Calling (2015).
- HLA-DRB1 gene. Genetics Home Reference (2017). Disponível em: <<https://ghr.nlm.nih.gov/gene/HLA-DRB1>>. Acesso em: 14 de novembro de 2017.
- LARSON, D. Fossil Hunters - Paleobiology (2017). Disponível em: <<https://www.fossilhunters.xyz/paleobiology-2/introduction-izd.html>>. Acesso: 15 de novembro de 2017.
- MACCARI, G. *et al.* IPD-MHC 2.0: an improved inter-species database for the study of the major histocompatibility complex. *Nucleic Acids Res*, 2017; 45: D860–D864.
- MURPHY, K.; TRAVERS, P.; WALPORT, M. *Imunobiologia de Janeway*. Porto Alegre: Artmed, 7ª edição. 2010.
- NOBLE, J.A; VALDES, A.M. Genetics of the HLA region in the prediction of type 1 diabetes. *Curr Diab Rep*, 2011; 11(6): 533-542. doi:10.1007/s11892-011-0223-x
- OLIVEIRA, R.A.F DE. Análise da diversidade genética dos loci HLA-A, -B, -DRB1 e da estrutura da população humana do estado de Minas Gerais, Brasil. Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais. Belo Horizonte: 2009.
- PRUFER, K. *et al.* The bonobo genome compared with the chimpanzee and human genomes, *Nature research journal*, 2011.
- RAYCHAUDHURI, S. *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet*, 2012; 44(3): 291-296. doi:10.1038/ng.1076
- RIDLEY, M. *Evolução*. 3. ed. Porto Alegre: Artmed, 2006. 752p.
- ROGERS, J; GIBBS, R.A. Comparative primate genomics: emerging patterns of genome content and dynamics. *Nature Reviews Genetics*, 2014; 15(5): 347–359. doi:10.1038/nrg3707
- STAHL, E.A. *et al.* Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*, 2010; 42(6): 508-514. doi:10.1038/ng.582

STECK, A.K; REWERS, M.J; Genetics of type 1 diabetes. *Clin Chem*, 2011; 57(2): 176–185.
doi:10.1373/clinchem.2010.148221

SVENSSON, A; ANDERSON, G. Presence of retroelements reveal the evolutionary history of the human DR haplotypes. *Hereditas*, 1997; 127(1): 124-131. doi/10.1111/j.1601-5223.1997.00113.x