Research Article

Molecular Evolution of the Glycosyltransferase 6 Gene Family in Primates

Eliane Evanovich,1 Patricia Jeanne de Souza Mendonça-Mattos,1 and Maria Lúcia Harada2

1Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brazil
2Laboratório de Biologia Molecular “Francisco Mauro Salzano”, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brazil

Correspondence should be addressed to Eliane Evanovich; lianevanovich@gmail.com

Received 12 July 2016; Accepted 20 October 2016

Academic Editor: Stefano Pascarella

Copyright © 2016 Eliane Evanovich et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Glycosyltransferase 6 gene family includes ABO, Ggta1, iGb3S, and GBGT1 genes and by three putative genes restricted to mammals, GT6m6, GTm6, and GT6m7, only the latter is found in primates [3].

1. Introduction

Glycosyltransferases catalyze the biosynthesis of glycoconjugates and polysaccharides by addition of sugar residues to an acceptor substrate that produce important antigens in the signaling process and recognition by the immune system [1], consisting of more than 90 carbohydrate-active enzymes, grouped by protein similarities in the CAZY database (http://www.cazy.org/GlycosylTransferases.html). Glycosyltransferase 6 are type II transmembrane proteins localized in the Golgi complex. They have a general structure with a cytoplasmic tail, a transmembrane domain, and a large catalytic domain [1–3]. The family includes functional and nonfunctional proteins encoded by ABO, Ggtal, iGb3S, and GBGT1 genes and by three putative genes restricted to mammals: GT6m6, GTm6, and GT6m7; only the latter is found in primates [3].

ABO gene encodes A and B transferases that add N-acetyl-D-galactosamine (GlcNAc) or D-galactose (Gal) to H substance that produce A or B antigens, respectively. O allele is nonfunctional [4, 5]. Individuals with the O phenotype exhibit more resistance to severe malaria caused by Plasmodium falciparum than others with A and/or B types [6–8]. Nonetheless, analysis of this locus in primates showed that the alleles were maintained by balanced selection [9].

Ggtal gene encodes the enzyme α1,3-galactosyltransferase (α1,3-GT). It transfers UDP-Gal to N-acetyllactosamine and produces α-Gal epitope. Catarhini (human, apes, and Old World monkeys) produce anti-Gal antibody [10, 11]. According to Galili [12] the inactivation of that gene assured the emergence of this lineage. However, the functionality of Ggtal is maintained in noncatarrhine mammals [13]. The synthesis of α-Gal epitopes was also attributed to iGb3S gene in mouse. However, this gene encodes isogloboside b3 synthase by adding UDP-Gal to lactosylceramide that produces isogloboside b3 [14].

GBGT1 gene encodes the Forssman synthase (FS) via GlcNAc addition to globotriaosylceramide. The result of this
reaction is Forssman antigen, absent in catarrhine primates due to two nonsynonymous substitutions at residues 230 (G > S) and 269 (Q > R) [5, 15].

GT6 member putative genes do not appear to produce an enzyme with catalytic activity due to lack of six motifs LBR-B, LBR-C, LBR-F, LBR-G, LBR-H, and LBR [3].

Casals et al. [16] analyzed the molecular evolution of GBGT1, iGb3S, and GT6m7 in two human populations and surprisingly described that these supposed pseudogenes evolve under positive selection. Disagreeing with what would be expected for a nonfunctional sequence, although initially described as pseudogenes, some studies have shown the possibility of GT6m7 being functional. Schaefer et al. [17] found an association between this pseudogene and aggressive periodontitis in three European populations. The same association was strengthened by Hashim et al. [18].

According to Svensson et al. [19], three criteria are important to establish a pseudogene: (1) detrimental mutation (mutations, stop codon, frameshift, splice-site alterations, etc.); (2) number of nonsynonymous substitutions per nonsynonymous site (Ka/Ks ratio) that indicate absence of selective pressure; and (3) when the sequences are not overlapping any known gene.

Nevertheless, there are some controversial facts surrounding these criteria: some discoveries have found that some putative pseudogenes participate in DNA, RNA, and protein regulation and that their gene sequences were conserved for millions of years, when they should be evolving neutrally. Therefore, reviewing the pseudogene condition seemed to be important.

Like some of the GT6 genes are related to the susceptibility or resistance to various parasites, their evolutionary and functional analysis is important. Thus, we decided to investigate (i) the selective pressure on the GT6 paralogs genes in primates; (ii) the cause of the conservation of iGb3S pseudogenes in human, chimpanzee, and bonobo; and (iii) the functional potential of the GBGT1 and GT6m7 in catarrhines based on the literature and annotated protein database.

2. Material and Methods

2.1. Database and Alignment. The sequences used were retrieved from NCBI (https://www.ncbi.nlm.nih.gov/) and/or Ensembl databases (http://www.ensembl.org/index.html). The identification of the species used for each database and their accession numbers are presented in Additional file (Table S1 in Supplementary Material available online at http://dx.doi.org/10.1155/2016/9051727).

We analyzed only exons 6 and 7 of the ABO gene which encode the catalytic domain of the protein (comprising 823 of 1062 bp, resp.). Exons 1 to 5 are small (their all length is 237 bp) and they are absent or removed by alternative splicing of transcripts in many different tissues of mammals. The sequences were aligned using the MUSCLE of the MEGA 6 [20]. The best-fit model of evolution for each gene was estimated by Modeltest 3.7 [21] and the phylogenetic tree for 100 replicates was built using PhyML 3.0 [22]. Nucleotide diversity (π) was calculated by MEGA 6 [20].

2.2. Detection of Recombination Breakpoints. Recombination was performed by GARD methods available at the http://www.datamonkey.org/ server [23] under HKY85 model using the Kishino-Hasegawa test (KH). The rate variation was implemented by a general discrete distribution with three rate classes [23].

2.3. Analysis of Molecular Evolution. Molecular evolution was verified by Codeml program implemented in PAML 4 package [24] and by HyPhy in Datamonkey server [23]. Codeml was initially running the M0 model to obtain phylogenetic trees which subsequently were inputted in site models and “free-ratio” branch model [25, 26].

According to Svensson et al. [19], three criteria are important to establish a pseudogene: (1) detrimental mutation (mutations, stop codon, frameshift, splice-site alterations, etc.); (2) number of nonsynonymous substitutions per nonsynonymous site (Ka/Ks ratio) that indicate absence of selective pressure; and (3) when the sequences are not overlapping any known gene.

Nevertheless, there are some controversial facts surrounding these criteria: some discoveries have found that some putative pseudogenes participate in DNA, RNA, and protein regulation and that their gene sequences were conserved for millions of years, when they should be evolving neutrally. Therefore, reviewing the pseudogene condition seemed to be important.

Like some of the GT6 genes are related to the susceptibility or resistance to various parasites, their evolutionary and functional analysis is important. Thus, we decided to investigate (i) the selective pressure on the GT6 paralogs genes in primates; (ii) the cause of the conservation of iGb3S pseudogenes in human, chimpanzee, and bonobo; and (iii) the functional potential of the GBGT1 and GT6m7 in catarrhines based on the literature and annotated protein database.

2. Material and Methods

2.1. Database and Alignment. The sequences used were retrieved from NCBI (https://www.ncbi.nlm.nih.gov/) and/or Ensembl databases (http://www.ensembl.org/index.html). The identification of the species used for each database and their accession numbers are presented in Additional file (Table S1 in Supplementary Material available online at http://dx.doi.org/10.1155/2016/9051727).

We analyzed only exons 6 and 7 of the ABO gene which encode the catalytic domain of the protein (comprising 823 of 1062 bp, resp.). Exons 1 to 5 are small (their all length is 237 bp) and they are absent or removed by alternative splicing of transcripts in many different tissues of mammals. The sequences were aligned using the MUSCLE of the MEGA 6 [20]. The best-fit model of evolution for each gene was estimated by Modeltest 3.7 [21] and the phylogenetic tree for 100 replicates was built using PhyML 3.0 [22]. Nucleotide diversity (π) was calculated by MEGA 6 [20].

2.2. Detection of Recombination Breakpoints. Recombination was performed by GARD methods available at the http://www.datamonkey.org/ server [23] under HKY85 model using the Kishino-Hasegawa test (KH). The rate variation was implemented by a general discrete distribution with three rate classes [23].

2.3. Analysis of Molecular Evolution. Molecular evolution was verified by Codeml program implemented in PAML 4 package [24] and by HyPhy in Datamonkey server [23]. Codeml was initially running the M0 model to obtain phylogenetic trees which subsequently were inputted in site models and “free-ratio” branch model [25, 26].

According to Svensson et al. [19], three criteria are important to establish a pseudogene: (1) detrimental mutation (mutations, stop codon, frameshift, splice-site alterations, etc.); (2) number of nonsynonymous substitutions per nonsynonymous site (Ka/Ks ratio) that indicate absence of selective pressure; and (3) when the sequences are not overlapping any known gene.

Nevertheless, there are some controversial facts surrounding these criteria: some discoveries have found that some putative pseudogenes participate in DNA, RNA, and protein regulation and that their gene sequences were conserved for millions of years, when they should be evolving neutrally. Therefore, reviewing the pseudogene condition seemed to be important.

Like some of the GT6 genes are related to the susceptibility or resistance to various parasites, their evolutionary and functional analysis is important. Thus, we decided to investigate (i) the selective pressure on the GT6 paralogs genes in primates; (ii) the cause of the conservation of iGb3S pseudogenes in human, chimpanzee, and bonobo; and (iii) the functional potential of the GBGT1 and GT6m7 in catarrhines based on the literature and annotated protein database.
Table 1: Codons under positive selection or diversifying selection with values significant obtained by Codeml or HyPhy methods.

| Gene   | Codon position | Method       |
|--------|----------------|--------------|
| ABO    | 266            | M7-M8        |
|        | 268            | MEME         |
|        | 109            | M7-M8        |
| Ggtal  | 250            | MEME, SLAC   |
|        | 361            | FEL, MEME    |
|        | 96             | FEL, MEME    |
| iGb3S  | 135, 228, 233  | MEME         |
|        | 18, 21, 237, 266 | MEME   |
|        | 12             | FEL, MEME    |
| GT6m7  | 96             | SLAC, FEL, MEME, FUBAR |

268 (p value of 0.002). “Free-ratio” test showed that some branches evolve on positive selection. The results obtained corroborate to balancing selection described by Ségurel et al. [9]. According to some authors, certain alleles may act against pathogens and are maintained during the evolution [31–34]. In human populations, individuals with A or B types seem to be less susceptible to infections caused by Helicobacter pylori and Vibrio cholerae compared to the carriers of O allele [32, 35]; in contrast, greater resistance to severe malaria caused by Plasmodium falciparum in blood group O individuals due to reduced rosette formation in erythrocytes was observed [6, 36]. Results are shown in additional file: Tables S2–S4 and Figure S1.

3.2. Ggtal. None of the sites found by algorithms are involved in the protein activity in accordance with the inferred structure by Gastinel et al. [37] for bovine 1,3-Gal. Positive selection was indicated by comparison of M7 versus M8 models (LTR = 6.367; p < 0.05) in position 109 (E > K); however, it substitution is restricted to Gtta1 pseudogene of Angola colobus, thus expected.

250 and 351 site codons were found on pervasive diversifying selection by MEME and, respectively, by SLAC and FEL. Codon 250 is localized in α1,3-GalT catalytic pocket. It is involved in the stability of the C-terminal segment [37] and is a binding site for disaccharide acceptor substrates [38]. The squirrel monkey α1,3-GalT presents substitution of tryptophan to phenylalanine in this residue, but this change does not seem to have a large impact on the catalytic activity of the enzyme [38]. Codon 351 is near to C-terminal portion and in this analysis it presents very polymorphism, having polar and nonpolar amino acids in different mammals, but they do not seem to result in major changes in protein structure.

Results of “free-ratio” test indicate different selective pressures between the several lineages, especially in howler that shows the highest ω values (0.64), similar to results found by Koike et al. [13] (additional file: Tables S5–S7 and Figure S2).

3.3. iGb3S. Only the site 97 was pointed out as being under positive selection by FEL. It was also indicated by the MEME, such as sites 136, 229, and 275. Among them, 97, 229, and 275 positions are most relevant information, because amino acids are preserved in different taxa showing a strong selective pressure on protein in mammals (Figure 1) [39]. Substitutions in these residues are not involved in important enzymatic processes. However, the polarity changes involving residue 275 as observed in iGb3S protein of the Old World monkeys (G275R) may result in some change in conformational entropy of the backbone in the local coil of the protein (see [40]), because this residue is in a hydrophobic site on α-helix (α5).

This gene is inactive in humans due to multiple mutations, including Y252N and L187P substitutions, but the high conservation of the putative pseudogene in human populations (nucleotide diversity, π = 0.0010) was reported by Casals et al. [16].

iGb3S has no evidence of recombination and it is syntenically conserved in mammals. Moreover, it is flanked by two conserved genes (ZNF362, Zinc finger protein 362 and PHC2, polyhomeotic homolog 2 genes), suggesting that their conservation in humans may result from their chromosomal position that is under strong selective pressure (additional file: Tables S8–S10).

3.4. GBGT1. PARRIS indicates 82% sites under purifying selection (ω mean = 0.270) and particularly in Catarrhini that present low diversity (π = 0.030). (Results are shown in additional file: Table S11–S13.) These primates are Forssman-negative, instead of Forssman glycolipid. Some cells of these primates show the precursor glycolipids globotriaosylceramide and globoside that are used as the binding site for bacteria, viruses, and toxins [15, 41–44]. This is apparently an adaptive disadvantage, but it can be offset.

For instance, individuals whose expression is Forssman antigen (ABO subgroup A_pac) [19] present protection against Shiga toxin 1 (Stx1) and vulnerability to Stx2a toxin, variant of Stx2 though [45, 46]. Louise and Obrig [47] analyze patients with hemolytic uremic syndrome (HUS) caused by Escherichia coli O157 :H7 and observed that Stx2a toxin is 1000 times more cytotoxic than Stx1. Other epidemiological data using baboon and mouse models also noticed it [47–49]. It is possible that it is an evolutionary advantage, presenting a key role in the evolution of Catarrhini.

3.5. GT6m7. Our results indicated high conservation of nucleotides and purifying selection (ω mean of 0.475) as found by Casals et al. [16], although positive selection was detected at codon 126 by SLAC, FEL, MEME, and FUBAR (additional file: Tables S14–S16).

There is missing information on the protein structure, although its higher expression in the testis, leukocytes, and gingiva [17] is quite instigating. The most relevant information obtained was related to its regulatory potential, which is explained below.

In a genome-wide association study (GWAS), Schaefer et al. [17] found relation between an allele of SNP rs1537415 (localized at intron 2 of GT6m7) to aggressive periodontitis (AgP) in three European populations. More recently, this
association was found in Sudanese population [18]. Schaefer et al. [17] using computational analysis reported that substitution from C to G in SNP rs1537415 may decrease the affinity between the transcriptional factor GATA-3 and its binding site.

We submitted intron 2 to BLASTN Ensembl against Human GRCh38 and we found a fragment of 353 bp (121 bp upstream to SNP rs1537415) with greater identity than 90% to the following regulatory sequences: ADGRL3-AS1 (adhesion G protein-coupled receptor L3 antisense RNA), RUNX1 (runt-related transcription factor 1), SRBD1 (S1 RNA binding domain 1 gene), CTND2 (catenin delta 2), and FEZ2 (Fasci- culation and Elongation Protein Zeta 2). The values obtained for the five best hits are in additional file: Figure S3.

In recent decades, the function of some mammalian pseudogenes was clarified, acting at different levels, from DNA to protein [50–55]. For instance, the transcription of the Makorin1 gene is regulated by its pseudogene paralogous Makarini-p1 [55]; and HMGA1 (high mobility group A1) presents regulatory action on the insulin receptor gene (INSR) [55]. The data presented suggest that GT6m7 also has some functionality, but experimental data are needed.

4. Conclusion

As shown in this paper, the purifying selection is prevalent in primate GT6 genes and has a low diversity (\(\omega\) mean = 0.342 and \(\pi\) mean = 0.0822). Surprisingly, their retention was found even in taxa where one of the genes is inactivated. ABO, Ggtal, and GBGT1 genes appear to be under strong selective pressures exerted by pathogens. ABO locus evolved in a balanced manner in primates with the maintenance of the three major alleles in at least ten species, including human, orangutan, and squirrel monkey. Ggtal gene also gives advantage against pathogens with expression of \(\alpha\)-Gal epitopes or anti-Gal antibody [12]. Similarly, the absence of expression of Forssman antigen is relevant against

---

**Figure 1**: Phylogenetic tree for \(iGb3\) gene indicated three sites under diversifying selection in accordance with MEME. The sites that are shared by different taxa are highlighted in bold. Tree constructed using PhyML, 100 replicates, under GTR + G model. Lnl = −4263.35. Values of branches sizes and bootstrap were omitted.
bacterial infections. However, GT6m7 and iGb3S seem to evolve differently. GT6m7, in primates, may have regulatory function. Otherwise, the maintenance of iGb3S in humans, chimpanzee, and bonobo seems to be the only result of the proximity to genes with important biological functions.

**Disclosure**

Part of this work was presented in Eliane Evanovich’s Master thesis [56].

**Competing Interests**

The authors have declared that no competing interests exist.

**Acknowledgments**

The work was supported by Coordenação de Aper-feiçoamento de Pessoal de Nível Superior (CAPES) from the Education Ministry, Brazil. Part of this work was presented in Eliane Evanovich’s Master thesis [56].

**References**

[1] A. Varki, "Biological roles of oligosaccharides: all of the theories are correct," Glycobiology, vol. 3, no. 2, pp. 97–130, 1993.

[2] J. C. Paulson and K. J. Colley, "Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation," Journal of Biological Chemistry, vol. 264, no. 30, pp. 17615–17618, 1989.

[3] A. L. Turcot-Dubois, B. Le Moullac-Vaidye, S. Despiau et al., "Long-term evolution of the CAZY glycosyltransferase 6 (ABO) gene family from fishes to mammals—a birth-and-death evolution model," Glycobiology, vol. 17, no. 5, pp. 516–528, 2007.

[4] F. Yamamoto, "Review: ABO blood group system—ABH oligosaccharide antigens, anti-A and anti-B, A and B glycosyltransferases and ABO genes," Immunohematology, vol. 20, no. 1, pp. 3–22, 2004.

[5] F. Yamamoto, E. Cid, M. Yamamoto, and A. Blancher, "ABO research in the modern era of genomics," Transfusion Medicine Reviews, vol. 26, no. 2, pp. 103–118, 2012.

[6] A. E. Fry, M. J. Griffiths, S. Auburn et al., "Common variation in the ABO glycosyltransferase is associated with susceptibility to severe Plasmodium falciparum malaria," Human Molecular Genetics, vol. 17, no. 4, pp. 567–576, 2008.

[7] A. K. Panda, S. K. Panda, A. N. Sahu, R. Tripathy, B. Ravindrann, and B. K. Das, "Association of ABO blood group with severe falciparum malaria in adults: case control study and meta-analysis," Malaria Journal, vol. 10, article no. 309, 2011.

[8] M. J. Mackinnon, C. Ndila, S. Uyoga et al., "Environmental correlation analysis for genes associated with protection against malaria," Molecular Biology and Evolution, vol. 33, no. 5, pp. 1188–1204, 2016.

[9] L. Ségurel, E. E. Thompson, T. Flutre et al., "The ABO blood group is a trans-species polymorphism in primates," Proceedings of the National Academy of the Sciences of the United States of America, vol. 109, no. 45, pp. 18493–18498, 2012.

[10] W. M. Blanken and D. H. Van den Eijnden, "Bioformation of terminal GaIK1→3Galβ1→4GlcNAc-R oligosaccharide sequences on glycoconjugates. Purification and acceptor specificity of a UDP-Gal:α-N-acetyllactosaminidase1,3galactosyltransferase from calf thymus," The Journal of Biological Chemistry, vol. 260, pp. 12927–12934, 1985.

[11] F. Yamamoto, P. Q. H. Yang, and Y. Yang, "Characteristics of α-Gal epitope, anti-Gal antibody, α1,3-galactosyltransferase and its clinical exploitation (Review)," International Journal of Molecular Medicine, vol. 37, no. 1, pp. 11–20, 2016.

[12] U. Gallini, "Significance of the Evolutionary α1,3-galactosyltransferase (GGTA1) gene inactivation in preventing extinction of apes and old world monkeys," Journal of Molecular Evolution, vol. 80, no. 1, pp. 1–9, 2015.

[13] C. Koike, M. Uddin, D. E. Wildman et al., "Functionally important glycosyltransferase gain and loss during catarrhine primate emergence," Proceedings of the National Academy of Sciences of the United States of America, vol. 104, no. 2, pp. 559–564, 2007.

[14] J. Milland, D. Christiansen, B. D. Lazarus, S. G. Taylor, P. Xing, and M. S. Sandrin, "The molecular basis for Gal(1,3)Gal expression in animals with a deletion of the α1,3galactosyltransferase gene," The Journal of Immunology, vol. 176, no. 4, pp. 2448–2454, 2006.

[15] H. Xu, T. Storch, M. Yu, S. P. Elliott, and D. B. Haslam, "Characterization of the human Forsmansynthesetase gene. An evolving association between glycophillid synthesis and host-microbial interactions," The Journal of Biological Chemistry, vol. 274, no. 41, pp. 29390–29398, 1999.

[16] F. Casals, A. Ferrer-Admettla, M. Sikora et al., "Human pseudogenes of the ABO family show a complex evolutionary dynamics and loss of function," Glycobiology, vol. 19, no. 6, pp. 583–591, 2009.

[17] A. S. Schaef er, G. M. Richter, M. Nothnagel et al., "A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis," Human Molecular Genetics, vol. 19, no. 3, pp. 553–562, 2009.

[18] N. T. Hashim, G. J. Linden, M. E. Ibrahim et al., "Replication of the association of GLT6D1 with aggressive periodontitis in a Sudanese population," Journal of Clinical Periodontology, vol. 42, no. 4, pp. 319–324, 2015.

[19] Ø. Svensson, L. Arvestad, and J. Lagergren, "Genome-wide survey for biologically functional pseudogenes," PLoS Computational Biology, vol. 2, no. 5, p. e46, 2006.

[20] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, "MEGA6: molecular evolutionary genetics analysis version 6.0," Molecular Biology and Evolution, vol. 30, no. 12, pp. 2725–2729, 2013.

[21] D. Posada and K. A. Crandall, "MODELTEST: testing the model of DNA substitution," Bioinformatics, vol. 14, no. 9, pp. 817–818, 1998.

[22] S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and S. Gascuel, "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0," Systematic Biology, vol. 59, no. 3, pp. 307–321, 2010.

[23] S. L. K. Pond, S. D. W. Frost, and S. V. Muse, "HyPhy: hypothesis testing using phylogenies," Bioinformatics, vol. 21, no. 5, pp. 676–679, 2005.

[24] Z. Yang, "PAML 4: phylogenetic analysis by maximum likelihood," Molecular Biology and Evolution, vol. 24, no. 8, pp. 1586–1591, 2007.

[25] Z. Yang and R. Nielsen, "Synonymous and nonsynonymous rate variation in nuclear genes of mammals," Journal of Molecular Evolution, vol. 46, no. 4, pp. 409–418, 1998.
[26] Z. Yang, R. Nielsen, N. Goldman, and A.-M. K. Pedersen, “Codon-substitution models for heterogeneous selection pressure at amino acid sites,” *Genetics*, vol. 155, no. 1, pp. 431–449, 2000.

[27] W. J. Swanson, R. Nielsen, and Q. Yang, “Pervasive adaptive evolution in mammalian fertilization proteins,” *Molecular Biology and Evolution*, vol. 20, no. 1, pp. 18–20, 2003.

[28] W. S. W. Wong, Z. Yang, N. Goldman, and R. Nielsen, “Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites,” *Genetics*, vol. 168, no. 2, pp. 1041–1051, 2004.

[29] B. Murrell, S. Weaver, M. D. Smith et al., “Gene-wide identification of episodic selection,” *Molecular Biology and Evolution*, vol. 32, no. 5, pp. 1365–1371, 2015.

[30] E. Evanovich and M. L. Harada, “Primate ABO gene is under weak positive selection,” *Notulae Scientiae Biologicae*, vol. 4, pp. 7–12, 2012.

[31] N. Otting, N. G. De Groot, G. G. M. Doxiadis, and R. E. Bontrup, “Extensive Mhc-DQB variation in humans and non-human primate species,” *Immunogenetics*, vol. 54, no. 4, pp. 230–239, 2002.

[32] D. J. Anstee, “The relationship between blood groups and disease,” *Blood*, vol. 115, no. 23, pp. 4635–4643, 2010.

[33] E. M. Leffler, Z. Gao, S. Pfeifer et al., “Multiple instances of negative balancing selection shared between humans and chimpanzees,” *Science*, vol. 340, no. 6127, pp. 1578–1582, 2013.

[34] L. Azevedo, C. Serrano, A. Amorim, and D. N. Cooper, “Trans-species polymorphism in humans and the great apes is generally maintained by balancing selection that modulates the host immune response,” *Human Genetics*, vol. 9, article 21, 2015.

[35] J. B. Harris, A. I. Khan, R. C. LaRocque et al., “Blood group, immunity, and risk of infection with *Vibrio cholerae* in an area of endemicity,” *Infection and Immunity*, vol. 73, no. 11, pp. 7422–7427, 2005.

[36] T. N. Williams, “Human red blood cell polymorphisms and malaria,” *Current Opinion in Microbiology*, vol. 9, no. 4, pp. 388–394, 2006.

[37] L. N. Gastinel, C. Bignon, A. K. Misra, O. Hindsgaul, J. H. Shaper, and D. H. Joziase, “Bovine α1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases,” *The EMBO Journal*, vol. 20, no. 4, pp. 638–649, 2001.

[38] Y. Zhang, A. Deshpande, Z. Xie, R. Natesh, K. R. Acharya, and K. Brew, “Roles of active site tryptophans in substrate binding and catalysis by α-1,3 galactosyltransferase,” *Glycobiology*, vol. 14, no. 12, pp. 1295–1302, 2004.

[39] D. Christiansen, J. Milland, E. Mouhtouris et al., “Humans lack iGb3 due to the absence of functional iGb3-synthase: implications for NKT cell development and transplantation,” *PLoS Biology*, vol. 6, no. 7, article e172, 2008.

[40] C. N. Pace and J. M. Scholtz, “A helix propensity scale based on experimental studies of peptides and proteins,” *Biophysical Journal*, vol. 75, no. 1, pp. 422–427, 1998.

[41] D. B. Haslam and J. U. Baenziger, “Expression cloning of Forssman glycolipid synthetase: a novel member of the histo-blood group ABO gene family,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 20, pp. 10697–10702, 1996.

[42] G. T. Keusch, M. Jacewicz, M. Mobassaleh, and A. Donohue-Rolfe, “Shiga toxin: intestinal cell receptors and pathophysiology of enterotoxigenic effects,” *Reviews of Infectious Diseases*, vol. 13, supplement 4, pp. S304–S310, 1991.

[43] M. Ezzelarab, D. Ayares, and D. K. C. Cooper, “Carbohydrates in xenotransplantation,” *Immunology and Cell Biology*, vol. 83, no. 4, pp. 396–404, 2005.

[44] G. C. Hansson, K.-A. Karlsson, G. Larson, N. Strömberg, and J. Thurin, “Carbohydrate-specific adhesion of bacteria to thio-layer chromatograms: a rationalized approach to the study of host cell glycolipid receptors,” *Analytical Biochemistry*, vol. 146, no. 1, pp. 158–163, 1985.

[45] C. A. Lingwood, “Role of verotoxin receptors in pathogenesis,” *Trends in Microbiology*, vol. 4, no. 4, pp. 147–153, 1996.

[46] S. P. Elliott, M. Yu, H. Xu, and D. B. Haslam, “Forssman synthetase expression results in diminished Shiga toxin susceptibility: a role for glycolipids in determining host-microbe interactions,” *Infection and Immunity*, vol. 71, no. 11, pp. 6543–6552, 2003.

[47] C. B. Louise and T. G. Obrig, “Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells,” *Journal of Infectious Diseases*, vol. 172, no. 5, pp. 1397–1401, 1995.

[48] V. L. Tesh, J. A. Burris, J. W. Owens et al., “Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice,” *Infection and Immunity*, vol. 61, no. 8, pp. 3392–3402, 1993.

[49] R. L. Siegler, T. G. Obrig, T. J. Pysher, V. L. Tesh, N. D. Denkers, and F. B. Taylor, “Response to Shiga toxin 1 and 2 in a baboon model of hemolytic uremic syndrome,” *Pediatric Nephrology*, vol. 18, no. 2, pp. 92–96, 2003.

[50] S. N. Hirotsune, N. Yoshida, A. Chen et al., “An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene,” *Nature*, vol. 423, no. 6935, pp. 91–96, 2003.

[51] E. S. Balakirev and F. J. Ayala, “Pseudogenes: are they ‘junk’ or functional DNA?” *Annual Review of Genetics*, vol. 37, pp. 123–151, 2003.

[52] T. C. Roberts and K. V. Morris, “Not so pseudo anymore: pseudogenes as therapeutic targets,” *Pharmacogenomics*, vol. 14, no. 16, pp. 2023–2034, 2013.

[53] L. Korrodi-Gregório, J. Abrantes, T. Muller et al., “Not so pseudo: the evolutionary history of protein phosphatase 1 regulatory subunit 2 and related pseudogenes,” *BMC Evolutionary Biology*, vol. 13, no. 1, article no. 242, 2013.

[54] O. Podlaha and J. Zhang, “Nonneutral evolution of the transcribed pseudogene Makorin1-p1 in mice,” *Molecular Biology and Evolution*, vol. 21, no. 12, pp. 2202–2209, 2004.

[55] E.Chiefari,S.Iiritano,F.Paonessaetal.,”Pseudogene-mediatedposttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes,” *Nature Communications*, vol. 1, article 40, 2010.

[56] E. B. dos Santos, *Evolução molecular do sistema ABO em primatas do novo mundo (Master in Biological Sciences) [M.S. thesis]*, Universidade Federal do Pará, Centro de Ciências Biológicas, Departamento de Genética, Belém, Brazil, 2001.