UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
PROGRAMA DE PÓS-GRADUAÇÃO EM CLÍNICA MÉDICA

BIBIANA SGORLA DE ALMEIDA

Regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G*: associação com doenças e com a magnitude de expressão da proteína

Ribeirão Preto

2019

BIBIANA SGORLA DE ALMEIDA

Regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G*: associação com doenças e com a magnitude de expressão da proteína

Tese apresentada ao Programa de Pós-Graduação em Clínica Médica da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para a obtenção do título de Doutora em Ciências.

Área de Concentração: Medicina (Clínica Médica)

Orientador: Prof. Dr. Eduardo Antônio Donadi

Ribeirão Preto

2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Almeida, Bibiana Sgorla de

Regiões regulatórias transcricionais e pós-transcricionais do gene HLA-G: associação com doenças e com a magnitude de expressão da proteína. Ribeirão Preto, 2019.

180 p. : il. ; 30 cm

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Medicina (Clínica Médica).

Orientador: Donadi, Eduardo Antônio.

1. 3'NT. 2. Promotora. 3. Polimorfismos. 4. Haplótipos. 5. HLA-G solúvel.

Nome: ALMEIDA, Bibiana Sgorla de

Título: Regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G*: associação com doenças e com a magnitude de expressão da proteína.

Tese apresentada ao Programa de Pós-Graduação em Clínica Médica da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para a obtenção do título de Doutora em Ciências, área de concentração Medicina (Clínica Médica).

Aprovado em: _____

Banca Examinadora

Prof. Dr _____ Instituição _____

Julgamento _____ Assinatura _____

Prof. Dr _____ Instituição _____

Julgamento _____ Assinatura _____

Prof. Dr _____ Instituição _____

Julgamento _____ Assinatura _____

AGRADECIMENTOS

À Universidade de São Paulo, pela infraestrutura cedida.

Ao Professor Dr Eduardo Antonio Donadi, pela oportunidade, orientação e compreensão durante todo este período.

Ao Programa de Pós-graduação em Clínica Médica, pela oportunidade de realização deste doutorado, especialmente ao secretário Emerson Quirino de Oliveira pela atenção e disponibilidade.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro e concessão da bolsa para a realização do doutorado-sanduíche.

Ao *Service de recherches en Hémato-Immunologie* do *Commissariat à l'Energie Atomique et aux Energies Alternatives* (CEA), especialmente ao Dr Philippe Moreau, pela orientação, atenção e oportunidade de realização do estágio de doutorado-sanduíche.

À Universidade Federal de Santa Catarina (UFSC), pela oportunidade e apoio para a realização deste doutorado.

Aos professores da banca examinadora, que aceitaram contribuir com este trabalho.

Aos professores Dr Celso Teixeira Mendes Jr e Dr Erick Castelli pelo auxílio na execução dos experimentos e avaliação dos dados.

Aos integrantes do Laboratório de Imunopatologia Molecular pelo auxílio na realização dos experimentos desta tese, especialmente à Dra Juliana Massaro e ao Dr Fabrício Dias.

Aos meus amigos, especialmente aos queridos Bruna Bertol e Guilherme Debortoli, por toda a ajuda, paciência, receptividade e disponibilidade. À minha Emiliana Weiss, pelas conversas e por todo o auxílio nas análises dos dados. E à Yara Muniz, pela amizade, pelos conselhos e pelas infindáveis ajudas ao longo deste doutorado.

Aos meus colegas de trabalho do Laboratório Multiusuário de Estudos em Biologia (UFSC), pela compreensão da minha ausência durante esse período, em especial aos meus grandes amigos Denis Dall Agnolo e Vanessa Silva.

Aos meus pais, Fernando e Luci, e ao meu irmão, Lucas, por sempre terem estimulado minha formação acadêmica e acreditarem que a educação é o caminho.

Ao meu marido, Achilles, pelo incentivo, compreensão, companheirismo e por ter me dado, junto com a Olívia e o Achilles, uma família linda e cheia de amor.

E, ao meu bebê Achilles, por trazer novos e lindos sentidos e sentimentos para a vida.

Muito obrigada!

Se o que eu sou é também o que eu escolhi ser, aceito a condição.

Rodrigo Amarante

(Los Hermanos – O velho e o moço)

RESUMO

ALMEIDA, Bibiana Sgorla de. **Regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G*: associação com doenças e com a magnitude de expressão da proteína.** 2019. 180 f. Tese de Doutorado - Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo, Ribeirão Preto, 2019.

O gene *HLA-G* apresenta considerável variabilidade nucleotídica nas suas regiões reguladoras, mas limitada variabilidade proteica e restrita expressão tecidual, em condições não-patológicas. A proteína HLA-G modula a atividade de células do sistema imune inato e adquirido e tem sido explorada em várias condições fisiológicas e patológicas, nas quais essas respostas podem ser benéficas ou deletérias, dependendo das características dessas condições. Considerando que a magnitude da expressão de HLA-G pode ser regulada pelas regiões promotora e 3'NT e que a expressão diferencial de HLA-G foi descrita mesmo em indivíduos saudáveis, neste trabalho avaliamos a diversidade das regiões regulatórias do gene *HLA-G* e sua possível influência nas variações de expressão da proteína em condições fisiológicas e patológicas. Para isso, realizamos: (i) uma revisão sistemática seguida de meta-análise sobre o polimorfismo mais estudado da região 3'NT do gene (presença ou ausência de 14-pb) em doenças de diversas etiologias; (ii) a tipificação das regiões regulatórias do *HLA-G* (sequenciamento de nova geração) em pacientes com doença de Chagas, para verificação de possíveis associações entre os polimorfismos e a doença, e a avaliação dos níveis plasmáticos (ELISA) de HLA-G e da expressão tecidual (em coração, colón e esôfago) da proteína HLA-G (imunoistoquímica); (iii) a análise da diversidade das regiões promotoras (segmentos distal, proximal e *enhancer* L) e 3'NT do gene *HLA-G* nas diversas populações mundiais que compõem a base de dados *1000 Genomes* - Fase 3, para identificação dos principais haplótipos, suas frequências e relações com as diferentes regiões geográficas; e (iv) a associação da variabilidade das regiões promotoras (distal e proximal) e 3'NT do *HLA-G* com os níveis plasmáticos de sHLA-G em indivíduos saudáveis, para identificar possíveis interferências das variações nucleotídicas dessas regiões na expressão basal da proteína HLA-G. De maneira geral, a meta-análise revelou que a variabilidade dos 14 pares de bases na região 3'NT não é um marcador genético eficiente para estudos de associação de doenças. Embora os níveis plasmáticos de sHLA-G e a expressão tecidual de HLA-G estivessem mais elevados pacientes

com doença de Chagas quando comparados com respectivos controles, não observamos associação entre os sítios polimórficos das regiões promotora e 3'NT do gene com a doença. Nas diversas populações mundiais, os alelos mais frequentes observados nas regiões reguladoras do gene *HLA-G* apresentam-se com frequências similares e verificou-se indícios de seleção balanceadora atuando nesses segmentos gênicos. Além disso, foi verificado um alto desequilíbrio de ligação entre os pares de SNP que compreendem toda essa região. Na população brasileira do Nordeste do Estado de São Paulo, não observamos associação entre os níveis de sHLA-G com a variabilidade das regiões reguladoras do gene. Por fim, considerando que a modulação da expressão de HLA-G pode ter possível aplicação clínica, o conhecimento da diversidade das regiões reguladoras pode contribuir para o entendimento acerca da complexidade da regulação da expressão e da magnitude de expressão do gene.

Palavras-chave: 3'NT. Promotora-Proximal. Promotora-Distal. Polimorfismos. Haplótipos. HLA-G solúvel. Doença de Chagas. Caso-controle. Metanálise.

ABSTRACT

ALMEIDA, Bibiana Sgorla de. **Transcriptional and post-transcriptional regulatory regions of *HLA-G* gene: association with diseases and the magnitude of protein expression**. 2019. 180 pages. Tese de Doutorado - Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo, Ribeirão Preto, 2019.

HLA-G gene has considerable nucleotide variability in its regulatory regions, but limited protein variability and restricted tissue expression under non-pathological conditions. *HLA-G* protein modulates the activity of innate and acquired immune system cells and has been exploited under various physiological and pathological conditions, where these responses may be beneficial or deleterious, depending on disease conditions. Considering that the magnitude of *HLA-G* expression can be regulated by the promoter and 3' untranslated (3'UTR) regions and that differential expression of *HLA-G* was described even in healthy individuals, in this work we evaluated the diversity of the *HLA-G* gene regulatory regions and its possible influence on protein expression variation under physiological and pathological conditions. For this, we performed: (i) a systematic review followed by meta-analysis of the most studied polymorphism of the *HLA-G* 3'UTR (presence or absence of 14-bp) in diseases of several etiologies; (ii) the typing of *HLA-G* regulatory regions (next generation sequencing) in patients with Chagas disease, to verify possible associations between polymorphisms and disease, and the evaluation of *HLA-G* plasma levels (ELISA) and tissue expression (in heart, colon and esophagus) of *HLA-G* protein (immunohistochemistry); (iii) the analysis of the promoter regions (distal, proximal and enhancer segments L) and 3'UTR diversity on worldwide populations, using the 1000 Genomes - Phase 3 database, to identify the main haplotypes, their frequencies and relations with different geographical regions; and (iv) the association of the variability of *HLA-G* promoter (distal and proximal) and 3'UTR segments with s*HLA-G* plasma levels in healthy individuals, to identify possible interference of nucleotide variations on basal *HLA-G* protein expression. Overall, the meta-analysis revealed that 14 base pair variability at the 3'UTR is not an efficient genetic marker for disease association studies. Although s*HLA-G* plasma levels and *HLA-G* tissue expression were higher in patients with Chagas disease compared to their controls, we found no association between the polymorphic promoter and 3'UTR gene regions with disease. In the world population evaluation, the most

frequent alleles observed at the regulatory regions of the *HLA-G* gene are similar in frequency and there was evidence of balancing selection acting on these gene segments. In addition, a high linkage disequilibrium was found between SNP pairs that comprise this entire region. In the Brazilian population of Northeastern São Paulo State, we did not observe any association between sHLA-G levels and the variability of gene regulatory regions. Finally, considering that the modulation of HLA-G expression may have possible clinical application, the knowledge of the diversity of regulatory regions may contribute to the understanding the complexity of *HLA-G* gene regulation.

Keywords: 3'UTR. Promoter Polymorphisms. Haplotypes. Soluble HLA-G. Chagas Disease. Case-control. Meta-analysis.

LISTA DE FIGURAS

Figura 4.1.1 - Fluxograma ilustrando a estratégia de busca usada para identificar os artigos relevantes para a meta-análise.	38
Figura 4.1.2 - Resultados da meta-análise do grupo Doenças Reprodutivas	50
Figura 4.1.3 - Resultados das meta-análises de Eclâmpsia – quando os alelos dos filhos foram avaliados (A) e Lúpus Eritematoso Sistêmico (B)	51
Figura 4.1.4 - Resultados das meta-análises de Câncer de Mama (A) e Infecção por Citomegalovírus Humano (B).....	52
Figura 4.2.1 – Comparação dos níveis de expressão de sHLA-G nos grupos Casos e Controles	54
Figura 4.2.2 – Comparação dos níveis de expressão de sHLA-G nos grupos Controles e Casos, subdivididos por formas da doença de Chagas.....	55
Figura 4.2.3 - Representação dos três tipos de tecidos avaliados quanto à expressão de HLA-G	56
Figura 4.2.4 – Comparação dos níveis de expressão tecidual de HLA-G em amostras de tecidos de pacientes acometidos pela Doença de Chagas.....	57
Figura 4.3.1 - Representação dos haplótipos da Região Promotora-Distal com frequências maiores que 1% nas populações avaliadas.....	73
Figura 4.3.2 - Rede <i>Median-Joining</i> dos haplótipos da região Promotora-Distal do <i>HLA-G</i>	74
Figura 4.3.3 - Dispersão em Escala Multidimensional considerando a Região Promotora-Distal do <i>HLA-G</i>	75
Figura 4.3.4 - Representação dos haplótipos da Região Promotora-Proximal com frequências maiores que 1% nas populações avaliadas.....	92
Figura 4.3.5 - Rede <i>Median-Joining</i> dos haplótipos da região Promotora-Proximal do <i>HLA-G93</i>	
Figura 4.3.6 - Dispersão em Escala Multidimensional considerando a Região Promotora-Proximal do <i>HLA-G</i>	94

Figura 4.3.7 - Representação dos haplótipos da Região <i>Enhancer L</i> nas populações avaliadas	100
Figura 4.3.8 - Rede <i>Median-Joining</i> dos haplótipos da região <i>Enhancer L</i> do <i>HLA-G</i>	101
Figura 4.3.9 - Dispersão em Escala Multidimensional considerando a Região <i>Enhancer L</i>	102
Figura 4.3.10 - Representação dos haplótipos da Região 3'NT com frequências maiores que 1% nas populações avaliadas	111
Figura 4.3.11 - Rede <i>Median-Joining</i> dos haplótipos da região 3'NT do <i>HLA-G</i>	112
Figura 4.3.12 - Dispersão em Escala Multidimensional considerando a Região 3'NT do <i>HLA-G</i>	113
Figura 4.3.13 - Desequilíbrio de Ligação entre as regiões <i>Enhancer L</i> , promotora-distal, promotora-proximal e 3'NT do gene <i>HLA-G</i>	115
Figura 4.4.1 – Expressão de sHLA-G de acordo com os haplótipos da região Promotora-Distal	118
Figura 4.4.2 – Expressão de sHLA-G de acordo com os grupos haplótipos da região Promotora-Distal.....	119
Figura 4.4.3 – Expressão de sHLA-G de acordo com os haplótipos da região Promotora-Proximal.....	120
Figura 4.4.4 – Expressão de sHLA-G de acordo com os grupos de haplótipos da região Promotora-Proximal	121
Figura 4.4.5 – Expressão de sHLA-G de acordo com os haplótipos da região 3'NT	122
Figura 4.4.6 – Expressão de sHLA-G de acordo com a presença dos dois haplótipos mais frequentes da região 3'NT	123

LISTA DE TABELAS

Tabela 4.1.1 - Relação, em ordem alfabética, do número de artigos por país em que foram realizados	39
Tabela 4.1.2 - Relação do número de artigos considerando o método de genotipagem utilizado	40
Tabela 4.1.3 - Resultados de todas as meta-análises realizadas no presente trabalho	45
Tabela 4.1.4 – Principais associações entre os alelos do polimorfismo de 14-pb e os haplótipos mais frequentes das regiões promotora-proximal, codificadora e 3’NT do gene HLA-G e proteínas.....	53
Tabela 4.2.1 – Classificação das amostras teciduais quanto ao número de células expressando HLA-G.....	57
Tabela 4.2.2 – Frequências e valores de OR e IC dos grupos de haplótipos da região promotora-proximal do gene HLA-G.....	59
Tabela 4.2.3 – Frequências e valores de OR e IC de pontos de variação da região promotora-proximal do gene HLA-G	59
Tabela 4.2.4 – Frequências e valores de OR e IC dos haplótipos da região 3’NT do gene <i>HLA-G</i>	60
Tabela 4.2.5 – Frequências e valores de OR e IC de pontos de variação da região 3’NT do gene <i>HLA-G</i>	60
Tabela 4.3.1 – Haplótipos da região Promotora-Distal do <i>HLA-G</i> e suas frequências na população mundial	63
Tabela 4.3.2 - Frequências dos haplótipos da região Promotora-Distal do <i>HLA-G</i> em diferentes populações do mundo.....	69
Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do <i>HLA-G</i> e suas frequências na população mundial	78
Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do <i>HLA-G</i> em diferentes populações do mundo.....	84
Tabela 4.3.5 – Haplótipos da região <i>Enhancer L</i> e suas frequências na população mundial ...	97

Tabela 4.3.6 - Frequências dos haplótipos da região <i>Enhancer L</i> em diferentes populações do mundo	98
Tabela 4.3.7 – Haplótipos da região 3’NT do <i>HLA-G</i> e suas frequências na população mundial	105
Tabela 4.3.8 - Frequências dos haplótipos da região 3’NT do <i>HLA-G</i> em diferentes populações do mundo.....	107
Tabela 4.4.1 – Haplótipos identificados nas regiões reguladoras Promotora-Distal, Promotora-Proximal e 3’NT do gene <i>HLA-G</i> em uma amostra de indivíduos saudáveis.....	116

LISTA DE ABREVIATURAS E SIGLAS

3'NT – 3' não-traduzida

AMOVA – Análise de Variância Molecular

HLA – Antígeno Leucocitário Humano

Indel – Inserção/deleção

LD – Desequilíbrio de Ligação

MAF – Frequência Alélica Mínima

MDS – Dispersão em Escala Multidimensional

MHC – Complexo de Histocompatibilidade Principal

miRNA - microRNA

s – desvio-padrão da média amostral

sHLA-G – HLA-G solúvel

SNV – variante nucleotídica única

$\bar{x}$ - média amostral

SUMÁRIO

1 INTRODUÇÃO	20
2 OBJETIVOS.....	25
2.1 Objetivo geral.....	25
2.2 Objetivos específicos	25
3 MATERIAL E MÉTODOS	26
3.1. Avaliação da associação dos alelos do polimorfismo de 14-pb do gene <i>HLA-G</i> em doenças de diversas etiologias.....	26
3.1.1. Estratégia de busca	26
3.1.2 Critérios de inclusão e exclusão.....	26
3.1.3 Extração dos dados.....	27
3.1.4 Análise estatística	27
3.1.5 Coleta de dados sobre polimorfismos e proteínas HLA-G	28
3.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene <i>HLA-G</i> , em um estudo caso-controle para a Doença de Chagas.....	28
3.2.1 Níveis de sHLA-G no plasma.....	29
3.2.2 Expressão de HLA-G em tecidos	30
3.2.3 Sequenciamento de DNA.....	31
3.3 Caracterização das regiões regulatórias do gene <i>HLA-G</i> na população mundial.....	33
3.3.1 Coleta de dados genéticos de <i>HLA-G</i> na população mundial	33
3.3.1.1 Extração dos dados do Projeto <i>1000 Genomes</i>	33
3.3.1.2 Revisão de literatura.....	34
3.3.2 Análise estatística.....	35
3.4 Relação entre as concentrações de HLA-G solúvel e os polimorfismos das regiões regulatórias do gene <i>HLA-G</i> , em amostras de indivíduos saudáveis	36
3.4.1 Quantificação de HLA-G Solúvel	36

3.4.2 Identificação das variantes das regiões reguladoras do <i>HLA-G</i>	36
3.4.3 Análise estatística	36
4. RESULTADOS	38
4.1 Avaliação da associação dos alelos do polimorfismo de 14-pb do gene <i>HLA-G</i> em doenças	38
4.1.1 Identificação e seleção dos artigos	38
4.1.2 Características dos artigos incluídos	38
4.1.3 Dados gerais das meta-análises	42
4.1.4 Associações entre os alelos do polimorfismo de 14-pb com os haplótipos das regiões promotora-proximal, codificadora e 3'NT do gene <i>HLA-G</i> e proteínas	44
4.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene <i>HLA-G</i> , em um estudo caso-controle para a Doença de Chagas	54
4.2.1 Expressão de sHLA-G no plasma	54
4.2.2 Expressão de HLA-G em tecidos	56
4.2.3 Variabilidade genética do <i>HLA-G</i>	58
4.3 Caracterização das regiões regulatórias do gene <i>HLA-G</i> na população mundial	61
4.3.1 Região Promotora-Distal	61
4.3.2 Região Promotora-Proximal	76
4.3.3 Região <i>Enhancer L</i>	95
4.3.4 Região 3'NT	103
4.3.5 Desequilíbrio de ligação	114
4.4 Relação entre as concentrações de HLA-G solúvel e os polimorfismos das regiões regulatórias do gene <i>HLA-G</i> , em amostras de indivíduos saudáveis	116
5 DISCUSSÃO	124
5.1 Avaliação da associação dos alelos do polimorfismo de 14-pb do gene <i>HLA-G</i> em doenças de diversas etiologias	124
5.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene <i>HLA-G</i> , em um estudo caso-controle para a Doença de Chagas	127
5.3 Caracterização das regiões regulatórias do gene <i>HLA-G</i> na população mundial	128

5.4 Relação entre as concentrações de HLA-G solúvel e os polimorfismos das regiões regulatórias do gene <i>HLA-G</i>	131
6 CONCLUSÃO	135
BIBLIOGRAFIA	136
ANEXOS	150
A – Artigo Publicado	150
B – Processo Comitê de Ética em Pesquisa	151
APÊNDICES	152
A – Dados completos das análises populacionais para a região Promotora-Distal	152
B – Dados completos das análises populacionais para a região Promotora-Proximal	158
C – Dados completos das análises populacionais para a região <i>Enhancer L</i>	164
D - Dados completos das análises populacionais para a região 3'NT	170

1 INTRODUÇÃO

O gene que codifica o Antígeno Leucocitário Humano G (*HLA-G*) está localizado na região cromossômica 6p21.3, dentro do agrupamento do Complexo Principal de Histocompatibilidade (CPH ou *Major Histocompatibility Complex*-MHC) (GERAGHTY; KOLLER; ORR, 1987). A HLA-G é uma proteína de histocompatibilidade não clássica de classe I que apresenta propriedades biológicas diferentes das proteínas HLA de classe I clássicas (-A, -B e -C), incluindo: i) limitada variabilidade proteica, expressão tecidual restrita em condições não patológicas e capacidade de modulação da atividade de células envolvidas na resposta imune. Em contraste com as proteínas HLA clássicas, HLA-G não tem papel importante na apresentação de antígenos, mas tem a propriedade de inibir: (i) a atividade de células TCD8+, (ii) a proliferação de células TCD4+, (iii) a apresentação de antígenos pelas células dendríticas, (iv) a função de células B e, ainda, pode (v) induzir células T regulatórias, por intermédio da interação direta com os receptores leucocitários ILT-2, ILT-4, KIR2DL4 e CD160. Esses receptores também interagem com as proteínas HLA clássicas, mas possuem maior afinidade pela proteína HLA-G (DONADI et al., 2011; CAROSELLA et al., 2015). Face às suas atividades de inibição das células do sistema imune inato e adquirido, a proteína HLA-G tem sido reconhecida como molécula de tolerância.

Embora descritas primariamente em células de coriocarcinoma, as propriedades tolerogênicas da HLA-G também têm sido associadas à proteção do feto contra a resposta imune materna (na gestação, o feto pode ser considerado um enxerto semi-alogênico no organismo materno). Atualmente, a HLA-G é considerada uma proteína de *checkpoint* imune e suas funções imunoreguladoras têm sido exploradas em uma variedade de condições fisiológicas e patológicas, nas quais essas respostas podem ser benéficas ou deletérias, dependendo das características das doenças. Em tumores, por exemplo, a expressão de HLA-G pode ser entendida como um mecanismo de escape da resposta imune, facilitando a progressão da doença; enquanto que, em pacientes transplantados, a expressão de HLA-G pode reduzir a resposta imune contra o enxerto (DONADI et al., 2011; CAROSELLA et al., 2015; MORANDI et al., 2016).

A variabilidade observada nas regiões regulatórias do gene *HLA-G* é relativamente maior do que a observada na sua região codificadora e, também, quando comparada com a mesma região dos genes HLA clássicos de classe I, os quais possuem grande variabilidade nas

suas regiões codificadoras (CASTELLI et al., 2014; SABBAGH et al., 2014). Para o *HLA-G*, estão descritas pelo menos quatro regiões reguladoras da sua expressão, incluindo: (i) região promotora-distal – recentemente descrita, possui em torno de 1,3 Kb (-2635 a -1406), na qual foram identificados 29 variações nucleotídicas simples (*single nucleotide variation* – SNV) e 19 haplótipos (OLIVEIRA et al., 2018); (ii) região promotora-proximal (convencional) – é a região que tem sido mais estudada, estendendo-se 1,4 Kb a montante do códon iniciador (ATG), que traduz o primeiro aminoácido (metionina) da proteína (entre as posições -1406 e -1) e possui pelo menos 36 SNVs, compondo 25 diferentes haplótipos (CASTELLI, et al., 2017); (iii) região *enhancer L* - região amplificadora da expressão do *HLA-G*, que está situada 12 Kb a montante do gene e possui 121 pb, cuja variabilidade genética ainda não foi estudada (FERREIRA et al; 2016); e (iv) região 3' não-traduzida (3'NT) - descrita por Geraghty (1987), é composta por 754 nucleotídeos, apresenta mais de 15 SNVs descritos e mais de 40 haplótipos, sendo pelo menos 8 deles com frequência maior do que 1% na população mundial (SABBAGH et al., 2014).

Os SNVs da região promotora-proximal, localizados nas regiões de interação com elementos regulatórios, podem alterar a capacidade ligação dos fatores de transcrição e, conseqüentemente, a expressão da proteína HLA-G. Alguns desses fatores de transcrição induzem a expressão de HLA-G (dentre eles: PRE – elemento de resposta à progesterona; HSE – elemento de choque térmico; ISRE – elemento de resposta ao estímulo pelo interferon; CRE – elemento de resposta ao AMP cíclico), enquanto outros atuam reprimindo a expressão da proteína (entre os quais estão: RREB-1- proteína-1 ligante ao elemento responsivo Ras; GLI3 – oncogene associado ao glioma homólogo 3) (CAROSELLA et al., 2015). Uma vez que a região promotora-distal foi recentemente descrita, ainda não existem estudos funcionais sobre a ação de fatores de transcrição nessa região. No entanto, estudos *in silico* mostram que os sítios -2257, -2218, -2012, -1972 e -1573 são alvos de fatores de transcrição, muitos deles relacionados com a regulação de células do sistema imune (OLIVEIRA et al, 2018).

Outro elemento de regulação da expressão é o *enhancer L*, que apresenta ação diferencial na expressão de HLA-G em uma linhagem celular de coriocarcinoma humano (JEG3) e em células trofoblásticas isoladas de placenta humana, as quais apresentam expressão constitutiva da proteína. Dados apresentados por Ferreira e colaboradores (2016) demonstraram que a deleção da sequência de 121 pb (sequência que corresponde à região *enhancer L* completa), utilizando a técnica CRISPR/Cas9, promoveu a completa ausência de

expressão de HLA-G nas superfícies celulares avaliadas. Ainda na região do *enhancer L*, foram detectados *motifs* de interação com pelo menos duas famílias de fatores de transcrição (CEBP e GATA), importantes para o processo de placentação (FERREIRA et al, 2016).

Na quarta região reguladora do gene, a região 3'NT, os polimorfismos de 14-pb, +3142C/G e 3187A/G foram funcionalmente estudados (TAN et al., 2007) e as diferenças dos níveis de expressão de HLA-G têm sido associados com instabilidade e degradação do RNAm pela ação de microRNAs (miRNA) (DONADI et al., 2011; CASTELLI et al., 2014; CAROSELLA et al., 2015). A presença/ausência (inserção/deleção; INS/DEL, indel) dos 14-pb, localizada entre as posições +2961 e +2974 (HARRISON et al., 1993; CASTELLI et al., 2014), foi o primeiro sítio polimórfico descrito e é o mais estudado deles. Em amostras de trofoblastos, a presença (inserção/INS) dos 14-pb foi associada com baixa produção de RNAm, tanto para as isoformas de HLA-G solúveis como para as de membrana (HVIID et al., 2003, 2006). Adicionalmente, a presença dos 14-pb é relacionada à ocorrência de edição alternativa no transcrito primário de *HLA-G*, que resulta em um RNAm menor e mais estável. Entretanto, aparentemente, esse aumento da estabilidade não compensa a baixa produção de HLA-G, previamente associada ao alelo de inserção dos 14-pb (CASTELLI et al., 2014). De maneira geral, a inserção dos 14-pb tem sido relacionada com baixa produção de HLA-G, enquanto que a ausência dos 14-pb (deleção/DEL) com produção de níveis mais altos de HLA-G (HARRISON et al., 1993; ROUSSEAU et al., 2003).

A partir dos estudos relatando a INS/DEL de 14-pb na região 3'NT do *HLA-G* e sua associação diferencial com a estabilidade do RNAm, diversos estudos foram conduzidos para avaliar o impacto desse sítio polimórfico nos níveis proteicos. Níveis plasmáticos mais altos de HLA-G solúvel (sHLA-G) foram observados em doadores de sangue de cordão com genótipo 14-pb INS/INS (CAPITTINI et al., 2014). Uma avaliação dos níveis de sHLA-G em pacientes submetidas à fertilização *in vitro*, mostrou que em pacientes apresentando o genótipo 14-pb INS/INS, não era possível detectar sHLA-G (HVIID et al., 2004). Foi relatada ainda, uma associação da presença do alelo 14-pb com níveis altos de sHLA-G materno durante a gravidez (DAHL et al., 2015). Além do polimorfismo de 14-pb, outros sítios, individualmente ou sob a forma de haplótipos da região, também foram estudados em associação com os níveis de sHLA-G. Os haplótipos UTR-2, UTR-5 e UTR-7 (que contêm o alelo 14-pb INS) apresentaram níveis mais baixos de sHLA-G do que os haplótipos que contêm o alelo 14-pb DEL (UTR-1, UTR-

3, UTR-4 e UTR-18) (MARTELLI-PALOMINO et al., 2013). Em outro estudo, em que uma população de Mali teve seus haplótipos da região 3'NT identificados e as concentrações de sHLA-G quantificadas, os indivíduos com a presença da UTR-2 eram os que tinham os níveis mais baixos de sHLA-G no plasma (CARLINI et al., 2013). Além da região 3'NT, variações nucleotídicas individuais da região promotora proximal também têm sido estudadas em associação com os níveis de sHLA-G. Nesse contexto, recentemente, nosso grupo de pesquisa avaliou os níveis de sHLA-G em associação com SNVs ou haplótipos da região promotora-proximal, não evidenciando diferenças significantes (DIAS et al., 2018). No entanto, não existem artigos abordando, na mesma população, as associações dos diversos haplótipos de cada região reguladora do gene *HLA-G* com os níveis de sHLA-G.

Devido à possibilidade desses polimorfismos afetarem a expressão da proteína por intermédio das interações com fatores de transcrição e miRNAs e assim resultarem em distintos perfis de expressão de *HLA-G*, alguns sítios de variabilidade das regiões promotora-proximal e 3'NT têm sido associados com a susceptibilidade, o desenvolvimento e a morbidade de doenças de diversas etiologias, incluindo, dentre outras, neoplasias, transplantes, aborto espontâneo, eclampsia, doenças autoimunes, infecciosas, inflamatórias crônicas. Também nesses estudos de associação com doenças, o polimorfismo de 14-pb, na região 3'NT, foi o primeiro e mais estudado ponto de variação. No entanto, resultados obtidos nas diferentes pesquisas para uma mesma doença mostraram dados discrepantes e muitas vezes conflitantes, sendo, portanto, inconclusivos. Diante dessa observação, algumas meta-análises foram realizadas e publicadas considerando algumas doenças como: lúpus (ZHANG et al., 2014; CATAMO et al., 2015), pré-eclampsia (PABALAN et al., 2015), aborto espontâneo (WANG et al., 2013; SHI et al., 2017) e carcinoma hepatocelular (COELHO et al., 2016). Apesar de essas meta-análises terem sido realizadas, um estudo mais amplo englobando doenças e grupos de doenças de diversas etiologias ainda é necessário. Além desse, outros SNVs da região 3'NT e da região promotora-proximal têm sido associados com diversas doenças (CAROSELLA et al., 2015).

A variabilidade do gene *HLA-G* e os níveis plasmáticos e a expressão tecidual da proteína em doenças e em alotransplantes têm sido estudados por nosso grupo de pesquisa. Recentemente, desenvolvemos uma estratégia para a tipificação dos diversos segmentos do *HLA-G*, usando sequenciamento de nova geração. Essa abordagem tem sido inovadora, mas

ainda são poucos os estudos de associação com doenças que foram publicados usando essa nova estratégia. Nesse contexto, escolhemos uma doença endêmica no Brasil para estudar o gene/proteína HLA-G: a doença de Chagas.

A doença de Chagas (tripanossomíase americana) é causada pelo protozoário *Trypanosoma cruzi*, presente nas fezes de insetos da família *Triatominae*, em especial os conhecidos como barbeiros, transmitindo o protozoário através de suas picadas. A doença é caracterizada por uma fase aguda e uma fase crônica que pode apresentar quatro variações clínicas principais: (i) forma indeterminada (casos que evoluem sem sinais clínicos); (ii) forma cardíaca (casos em que ocorrem arritmias cardíacas, eventos tromboembólicos e organomegalia); (iii) forma digestiva (caracterizada pelo desenvolvimento de megaesôfago e megacólon); e (iv) forma cardiodigestiva ou mista (casos que apresentam os sintomas das formas cardíaca e digestiva) (RASSI et al., 2010).

Considerando que a expressão de HLA-G é regulada pelas regiões promotora-distal, promotora-proximal e 3'NT e que a expressão diferencial de HLA-G tem sido descrita até mesmo em indivíduos saudáveis (MARTELLI-PALOMINO et al., 2013), nos parece importante ampliar o estudo das regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G* em associação com doenças e com os níveis de sHLA-G. Em relação à associação com doenças, duas abordagens foram realizadas para esclarecer aspectos importantes da regulação transcricional e pós-transcricional do gene. Posto que a dicotomia presença ou ausência de 14-pb na região 3'NT do gene *HLA-G* ainda suscita grande debate, a realização de estudo de meta-análise englobando grande grupo de doenças pode contribuir para a decisão de avaliar somente este sítio ou a região 3'NT inteira. Em paralelo, a avaliação dos diversos segmentos do gene *HLA-G* na doença de Chagas, junto com os estudos dos níveis plasmáticos de sHLA-G e da expressão de HLA-G nos tecidos com desnervação (coração, colón e esôfago) e das associações de genótipos/fenótipos, podem contribuir para o entendimento da patogenia dessa doença. Acerca do segundo componente do estudo, a associação entre os sítios polimórficos das regiões reguladoras do *HLA-G* com os níveis plasmáticos de sHLA-G é também justificada, uma vez que não existem estudos, na mesma população, avaliando essas associações com as diferentes regiões.

2 OBJETIVOS

2.1 Objetivo Geral

Estudar as regiões regulatórias do gene *HLA-G* em relação à associação com doenças e avaliar a influência dos polimorfismos dessas regiões gênicas na expressão da proteína HLA-G.

2.2 Objetivos Específicos

1. Realizar meta-análise para avaliar a associação dos alelos do polimorfismo de 14-pb do gene *HLA-G* com doenças de diversas etiologias;
2. Identificar os polimorfismos e os haplótipos das regiões regulatórias do gene *HLA-G* em amostras de pacientes com doença de Chagas;
3. Quantificar os níveis de sHLA-G em pacientes com doença de Chagas;
4. Avaliar a expressão de HLA-G em amostras de tecidos de pacientes com a doença de Chagas;
5. Caracterizar as regiões regulatórias do gene *HLA-G*, identificando polimorfismos e haplótipos, em diferentes populações mundiais;
6. Identificar os polimorfismos e os haplótipos das regiões regulatórias do gene *HLA-G* em uma amostra de indivíduos saudáveis da região Nordeste do Estado de São Paulo e compará-los com aqueles observados nas populações mundiais;
7. Correlacionar as concentrações de sHLA-G e os haplótipos das regiões regulatórias do gene *HLA-G* identificados na amostra de indivíduos saudáveis da região Nordeste do Estado de São Paulo.

3 MATERIAL E MÉTODOS

3.1. Avaliação da associação dos alelos do polimorfismo de 14-pb do gene *HLA-G* em doenças de diversas etiologias

Para avaliar a associação dos alelos do polimorfismo de 14-pb do gene *HLA-G* em doenças de diversas etiologias, foi realizada revisão sistemática de todos os estudos de associação publicados seguida de meta-análise, que foi conduzida da seguinte forma¹:

3.1.1. Estratégia de busca

Uma revisão sistemática de literatura foi realizada, usando as bases de dados PubMed e SciELO. Na estratégia de busca utilizada no presente trabalho, foram utilizados os seguintes descritores controlados: "*HLA-G Antigens*", "*Polymorphism, Genetic*", "*3' Untranslated Regions*" e "*14-bp*". Adicionalmente, foram incluídos os descritores não-controlados (palavras textuais e seus sinônimos, variações de siglas e correlatos). Entre os descritores controlados foi utilizado o operador booleano AND e entre os descritores não-controlados o operador booleano OR, conforme descrito a seguir: "*HLA-G*" OR "*HLA-G Antigens*" OR "*HLA G Antigens*" OR "*HLA-G Antigen*" OR "*HLA G Antigen*" OR "*HLA G*" OR "*HLA-G*" OR "*HLA-G2 Isoform*" OR "*HLA G2 Isoform*" OR "*HLA-G2*" OR "*HLA-G7 Isoform*" OR "*HLA G7 Isoform*" OR "*HLA-G7*" OR "*HLA-G7 Antigen*" OR "*HLA G7 Antigen*" OR "*HLA-G4 Isoform*" OR "*HLA-G4*" OR "*HLA-G5 Isoform*" OR "*HLA-G5*" OR "*HLA-G6 Isoform*" OR "*HLA G6 Isoform*" OR "*HLA-G6*" OR "*HLA-G1 Isoform*" OR "*HLA G1 Isoform*" OR "*HLA-G1*" OR "*HLA-G3 Isoform*" OR "*HLA G3 Isoform*" OR "*HLA-G3*" AND "*Polymorphism*" OR "*Polymorphism, Genetic*" OR "*Genetic Polymorphism*" OR "*Polymorphism (Genetics)*" OR "*Genetic Polymorphisms*" AND "*3'NTR*" OR "*3' Untranslated Regions*" OR "*3'NTR*" OR "*3'NTRs*" OR "*3' Untranslated Region*" OR "*3' UTR*" OR "*3' UTRs*" AND "*14bp*" OR "*14-bp*" OR "*14 bp*" OR "*14bp*" OR "*indel*" OR "*insdel*".

3.1.2 Critérios de inclusão e exclusão

Todos os artigos identificados pelas buscas nas bases de dados passaram por processo de triagem e foram selecionados de acordo com os critérios de inclusão e exclusão. Na primeira etapa da triagem, os estudos foram avaliados pelos títulos e resumos e os estudos

¹ Os dados aqui apresentados, referentes à Revisão Sistemática e Meta-análise, compuseram o artigo *Genetic association between HLA-G 14-bp polymorphism and diseases: A systematic review and meta-analysis*, publicado na Revista Human Immunology, em agosto de 2018 (Anexo A).

considerados fora do objetivo de estudo desta meta-análise foram excluídos. Os artigos que permaneceram no estudo passaram por segundo estágio da triagem, no qual o conteúdo dos mesmos foi avaliado. Os artigos foram lidos na íntegra para averiguar se continham as informações de interesse para a realização da meta-análise. Foram incluídos os estudos que correspondessem aos seguintes critérios: (i) fosse estudo do tipo caso-controle; e (ii) possuisse dados originais, contendo as frequências dos alelos do polimorfismo de 14-pb. Foram excluídos os artigos que: (i) contivessem dados duplicados; (ii) fossem estudos familiares; e (iii) os dados de frequências dos alelos do polimorfismo de 14-pb estivessem incompletos.

3.1.3 Extração dos dados

Os dados extraídos dos artigos selecionados foram os seguintes: (i) título do artigo; (ii) nome do primeiro autor; (iii) ano de publicação; (iv) país de origem; (v) situação clínica/doença dos pacientes; (vi) método de genotipagem; (vii) Equilíbrio de Hardy-Weinberg (HWE) no grupo controle; (viii) número de casos e controles; (ix) frequências dos alelos do polimorfismo de 14-bp (INS/DEL) em casos e controles; (x) presença ou não de associação significativa e (xi) se significativa, qual o alelo associado. Os dados foram extraídos por dois pesquisadores de forma independente e as eventuais discordâncias foram resolvidas por consenso entre eles.

3.1.4 Análise estatística

Foram realizadas meta-análises para cada situação clínica/doença quando três ou mais artigos estavam disponíveis. Essas doenças também foram analisadas considerando grupos de doenças, categorizados da seguinte maneira: (i) câncer, (ii) doenças autoimunes, (iii) doenças inflamatórias, (iv) doenças infecciosas, (v) rejeição a transplantes, (vi) doenças reprodutivas. Além disso, as doenças foram classificadas em dois grandes grupos, considerando a atividade biológica da molécula HLA-G: (i) condições patológicas com hipotética alta produção de HLA-G (para esse grupo de doenças, a hipótese de associação esperada era com o alelo de deleção dos 14-pb, uma vez que esse alelo foi previamente associado com alta produção de HLA-G); e (ii) condições patológicas com hipotética baixa produção de HLA-G (para esse grupo de doenças, em contrapartida, a hipótese de associação era com o alelo de inserção dos 14-pb, anteriormente associado com baixa produção de HLA-G).

Todas as associações possíveis entre o polimorfismo de 14-pb e as doenças foram mensuradas calculando *Odds Ratio* (OR) e intervalo de confiança de 95% (IC) para os modelos

de efeito fixo (FEM) e randômico (REM). Os valores de p resultantes do Teste Q, os valores de τ^2 e os valores de I^2 foram utilizados para verificar a heterogeneidade entre os estudos analisados conjuntamente. Os efeitos de heterogeneidade I^2 variam de 0 a 100% e representam a proporção de variabilidade atribuída à heterogeneidade (valores de 25, 50 e 75% são considerados de baixa, moderada e alta heterogeneidade, respectivamente). O FEM assume que o fator genético tem efeito similar na susceptibilidade à doença entre todos os estudos analisados. Enquanto que o REM assume que a diferença entre os estudos mostra diversidade substancial. Quando os grupos de estudos são homogêneos, os dois modelos são similares. Em contraste, quando os grupos não são homogêneos, o REM usualmente fornece valores de intervalo de confiança (IC) mais confiáveis que o FEM. O REM é a melhor escolha na presença de heterogeneidade significativa entre os estudos. Por isso, foi assumido que, se $I^2 > 50\%$ ou no Teste Q $p < 0,01$, um REM foi aplicado nos dados. O viés de publicação foi avaliado usando o Teste de Regressão Linear de Egger, que mede a assimetria da distribuição dos dados, usando escala logarítmica natural de OR. Se o valor de $p < 0,05$ indica viés de publicação. Devido aos parâmetros definidos por testes estatísticos, somente os grupos com mais de 10 estudos tiveram os vieses de publicação avaliados. Se $I^2 > 50\%$ ou $p < 0,01$ no Teste Q, o REM foi aplicado naquele grupo de dados. A análise de dados foi realizada utilizando o pacote “meta” do *Software R Estatística*.

Esta revisão sistemática seguida de meta-análise foi conduzida de acordo com o Guia PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis - PRISMA guidelines*). O protocolo para esta revisão sistemática seguida de meta-análise foi registrado na Base de Dados de Revisões Sistemáticas PROSPERO (PROSPERO 2017:CRD42017067176).

3.1.5 Coleta de dados sobre polimorfismos genéticos e proteínas HLA-G

As relações entre os alelos do polimorfismo de 14-pb e as regiões promotora, codificadora e 3'NT do gene *HLA-G* e as proteínas mais frequentes foram adicionalmente analisadas. Os dados usados para comparação foram organizados baseados em banco de dados do nosso próprio laboratório e do banco de dados do Projeto *1000 Genomes* (<http://www.internationalgenome.org/>).

3.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene *HLA-G*, em um estudo caso-controle para a Doença de Chagas

Para avaliar as diferenças de variabilidade do gene *HLA-G* e de expressão da proteína HLA-G em situações patológicas e em indivíduos saudáveis, uma amostra de pacientes com Doença de Chagas (Casos) foi selecionada para ser comparada com os dados anteriormente apresentados de doadores de sangue (Controles). As análises foram realizadas do seguinte modo:

3.2.1 Níveis de sHLA-G no plasma

A concentração de sHLA-G no plasma foi avaliada em 80 amostras de pacientes com Doença de Chagas (64 Forma Cardíaca, 6 Forma Digestiva e 10 Forma Indeterminada) e em 161 indivíduos controles, utilizando ELISA (*Enzyme-Linked Immunosorbent Assay*). As amostras de plasma de pacientes chagásicos foram coletadas no ambulatório de Cardiopatia Chagásica do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, onde os mesmos recebem acompanhamento e tratamento para a doença. As 161 amostras de indivíduos controles, pessoas saudáveis doadoras de sangue, foram coletadas no Hemocentro do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (Processo Comitê de Ética em Pesquisa HCPR nº 11237/2009 - Anexo B).

Para realizar o ELISA, foram utilizados os anticorpos *mAb anti-HLA-G* (MEM-G/9 - EXBIO, República Tcheca) e *rabbit-anti-human β 2-microglobulin* (DAKO, Nova Iorque, EUA) como anticorpos de captura e detecção, respectivamente. O anticorpo de captura MEM-G/9 foi incubado em placas de 96 poços a 4°C, *overnight*. Após a saturação, os poços foram cobertos com 300 μ L do diluente (DAKO). Após duas horas, o diluente foi descartado e 50 μ L de plasma das amostras foram adicionados e a placa foi incubada à temperatura ambiente. Após duas horas, o anticorpo de detecção *anti- β 2-microglobulin* (DAKO) foi adicionado aos poços e a placa foi, novamente, incubada por mais uma hora à temperatura ambiente. Cada etapa do procedimento foi seguida de quatro lavagens, usando tampão contendo água, PBS 1X e 0,1% Tween. Por fim, os poços foram incubados com o substrato supersensível TMB (tetrametilbenzidina, SIGMA, Saint Louis, EUA) em um tampão moderadamente ácido, à temperatura ambiente, no escuro por 30 min. Após a incubação, foi adicionado HCl 1N e as densidades óticas foram medidas em 450 nm. Todas as amostras foram avaliadas em duplicata e o nível total de HLA-G solúvel foi determinado a partir de curva padrão, que foi construída

utilizando diluições de HLA-G5 como reagente padrão. Os resultados foram expressos em ng/mL (MARTELLI-PALOMINO et al., 2014).

Para a avaliação dos dados de sHLA-G obtidos foram, inicialmente, realizados testes de normalidade para verificar a distribuição das amostras analisadas. Como as amostras não mostraram distribuição normal ($p < 0,001$), foram utilizados os testes não-paramétricos Mann-Whitney (nos testes com dois grupos amostrais) e Kruskal-Wallis, seguido da aproximação de Dunn (nos testes com mais de dois grupos amostrais). Os testes foram realizados no software IBM SPSS.

3.2.2 Expressão de HLA-G em tecidos

A expressão de HLA-G foi avaliada por imunoistoquímica em 93 amostras de tecidos de pacientes com Doença de Chagas, que foram a óbito entre os anos de 1997 e 2006, selecionados por meio dos registros do Banco de Necropsia do Departamento de Patologia da Faculdade de Medicina de Ribeirão Preto/USP. As amostras avaliadas foram de tecidos esofágico ($n = 7$), intestinal ($n = 17$) e cardíaco ($n = 69$), todas preservadas em blocos de parafina e armazenadas no Setor de Patologia do HC/FMRP. A técnica de imunoistoquímica foi realizada utilizando o sistema de detecção universal LSAB-2/HRP (DAKO), baseado no protocolo estreptavidina-biotina-peroxidase. Para isso, foram feitos dois cortes de cinco micrômetros de espessura para cada amostra, os quais foram montados em lâminas silanizadas, pré-tratadas com Poli L-lisina (StarFrost® Advanced Polycat Knittel, Alemanha). Essas lâminas passaram pelo processo de desparafinização, por intermédio de banhos consecutivos de xilol e álcool, e posterior hidratação. Os cortes foram imersos em tampão citrato (pH 6,0) e fervidos (96°C, 30 min), para a recuperação antigênica. Em seguida, foi realizado o bloqueio da peroxidase endógena com banho de peróxido de hidrogênio (H_2O_2) a 3% em PBS (temperatura ambiente, 30 min). Para evitar reações inespecíficas, as lâminas foram lavadas em PBS (2 vezes, 5 min) e incubadas com polímero (15 min). As lâminas foram incubadas em câmara úmida (temperatura ambiente, *overnight*) com o anticorpo primário monoclonal anti-HLA-G HLA-MEM/G-02 (EXBIO Antibodies, Praha, República Tcheca), que reconhece todas as isoformas descritas da proteína HLA-G, na diluição 1:100. Após esse período, as lâminas foram lavadas com PBS e incubadas com o anticorpo secundário do sistema de detecção universal LSAB-2/HRP (temperatura ambiente, 30 min). Em seguida, as lâminas foram lavadas com PBS, incubadas com estreptavidina e, novamente, lavadas com

PBS. Posteriormente, a revelação da reação foi realizada, na qual foi utilizado o substrato revelador 3,3 diaminobenzidina (DAB) (Biocare Medical, California, EUA), à temperatura ambiente, por 5 min. A reação foi interrompida com água destilada e as lâminas coradas com Hematoxilina de Mayer por 30 s. Após essa etapa, as lâminas foram lavadas com água destilada, desidratadas em baterias de álcool e xilol e montadas com Entellan® e lamínula. Neste experimento, cortes histológicos de citotrofoblasto humano do primeiro trimestre de gravidez, que apresentam expressão constitutiva de HLA-G, foram utilizados como controle positivo da reação. O controle negativo foi feito utilizando cortes das mesmas amostras, no entanto, nessas, o anticorpo primário (anti-HLA-G) foi substituído por PBS. Para avaliar os dados de expressão de HLA-G em tecidos, as lâminas foram fotografadas no Sistema de Digitalização de Imagens Axio Scan (Zeiss, Alemanha), em um aumento de 400x. A detecção e a quantificação da marcação com o anticorpo anti-HLA-G foram realizadas por análise computacional usando o Software ImageJ (National Institutes of Health, Maryland, EUA). Inicialmente, foram detectadas e definidas as intensidades de marcação que seriam consideradas positivas. Em seguida, as figuras foram convertidas de 256 cores (8 bits) para escala de cinza e as intensidades de marcação quantificadas. As amostras também foram avaliadas de acordo com o número de células marcadas e classificadas da seguinte forma: (i) 0 (zero) para os casos em que não houve identificação de expressão de HLA-G; (ii) 1 (um) para as amostras em que 1-25% das células expressavam a proteína; (iii) 2 (dois) para os casos em que 26-50% das células estavam marcadas; (iv) 3 (três) para as amostras que tinham de 51-75% das células positivas; e (v) 4 (quatro) para os casos com 76-100% das células marcadas para HLA-G.

3.2.3 Sequenciamento de DNA

Para a avaliação dos polimorfismos presentes nas regiões reguladoras do gene *HLA-G*, foi realizado o sequenciamento de 192 amostras de pacientes com Doença de Chagas (131 Forma Cardíaca, 34 Forma Digestiva e 27 Forma Indeterminada) pela técnica de NGS (*Next Generation Sequencing*) plataforma MiSeq® (Illumina, San Diego, EUA). Com os dados obtidos nesse sequenciamento, foi realizado um estudo de associação caso-controle para a Doença de Chagas. No qual, foram feitas comparações entre os dados dos pacientes com a Doença de Chagas (casos) e dados de indivíduos saudáveis (controles); esses últimos compõem um banco de dados e amostras do grupo de pesquisas do laboratório.

A amplificação do *HLA-G* foi realizada usando os iniciadores 5'-ACACTCATAATTCATTCATTGAGC-3' e 5'-TCTTCTGATAACACAGGAACTTC-3' e compreendeu o segmento de nucleotídeos entre 29.794.114 e 29.799.118 da sequência do cromossomo 6 do genoma humano h19. A amplificação dos fragmentos ocorreu em uma Reação em Cadeia da Polimerase (PCR) com volume final de reação de 50 µL, contendo 2,5mM de dNTPs (Invitrogen, Carlsbad, EUA), 10 pmol de cada iniciador, 1,25 unidades de DNA polimerase (PrimeStar GXL, Takara, Mountain View, EUA), tampão de PCR 1X e 50 ng de DNA genômico. As condições de ciclagem foram: 30 ciclos desnaturação a 98 °C por 10 s, anelamento a 60 °C por 15 s e extensão 68 °C por 6 min; a amplificação foi verificada em gel de agarose 1% corado com GelRed® (Biotium™, Fremont, EUA). Após a amplificação, as amostras foram purificadas usando a enzima Illustra ExoProStar® (GE Healthcare, Pittsburg, EUA), quantificadas usando Qubit dsDNA High-Sensitivity Assays (Invitrogen, Carlsbad, EUA) e normalizadas na concentração recomendada para a preparação das bibliotecas. Essas foram preparadas utilizando o Kit de Preparação Nextera XT (Illumina, San Diego, EUA) e identificadas utilizando o Kit Nextera XT Index (Illumina). Após a preparação, foram quantificadas por qPCR usando o kit Kapa (Kapa Biosystems, Willminton, EUA), e normalizadas para a concentração recomendada para o sequenciamento. O padrão de fragmentação das bibliotecas foi avaliado usando *chips* High-Sensitivity DNA BioAnalyzer (Agilent Technologies, Santa Clara, EUA). O sequenciamento foi realizado em um equipamento MiSeq (Illumina, San Diego, EUA) utilizando o Kit MiSeq V2, para 500 ciclos, 2x250 pb (Illumina, San Diego, EUA). Os dados obtidos nesse sequenciamento no formato FastQ foram processados e analisados utilizando uma metodologia desenvolvida pelo professor Dr. Erick Castelli (CASTELLI et al, 2017). Inicialmente, foram aplicados filtros para avaliar a qualidade das sequências produzidas, utilizando o programa Fastaqc. Em seguida, as sequências com boa qualidade (valores de Q acima de 20) foram mapeadas utilizando o Programa HLA-Mapper (CASTELLI et al., 2018), por meio do qual foram identificadas as sequências correspondentes ao gene *HLA-G*, utilizando o genoma referência h19. As sequências identificadas como sendo do gene *HLA-G* tiveram a cobertura e o balanceamento avaliados no Software IGV (*Integrative Genomics Viewer*, National Human Genome Research Institute, EUA). Posteriormente, as sequências foram submetidas a análises no Software GATK (*Genome Analysis Toolkit*, Broad Institute), utilizando as funções *HaplotypeCaller* e *ReadBackedPhasing*, por meio das quais foram detectados os genótipos e inferidos os haplótipos das amostras. A análise das frequências dos haplótipos foi

realizada no *Software* Arlequin (EXCOFFIER, LISCHER, 2010) e os cálculos de *Odds Ratio* (OR), para a avaliação de possíveis associações entre os polimorfismos/haplótipos do gene e a Doença de Chagas, foram realizados utilizando o pacote “meta” do *Software* R Estatística.

3.3 Caracterização das regiões regulatórias do gene *HLA-G* na população mundial

Para conhecer os principais sítios de variabilidade de quatro regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G* (promotora-distal, promotora-proximal, *enhancer L* e 3’NT) e seus haplótipos mais frequentes, foram avaliados dados de sequências de DNA de diferentes populações do mundo, disponíveis no banco de dados *1000 Genomes*, e os dados disponíveis em artigos, previamente publicados, que contivessem informações populacionais sobre alguma dessas regiões². A caracterização das regiões foi realizada da seguinte maneira:

3.3.1 Coleta de dados genéticos de *HLA-G* na população mundial

A coleta dos dados genéticos para identificação dos haplótipos mais frequentes das regiões promotora-distal, promotora-proximal, *enhancer L* e 3’NT foi realizada em duas etapas. Na primeira etapa, foram extraídas as sequências correspondentes a essas regiões do site do Projeto *1000 Genomes* – Fase 3 (1KG; <http://www.1000genomes.org>). Na segunda etapa, foi realizada revisão de literatura para a inclusão de todos os dados publicados e disponíveis dessas regiões em diferentes populações do mundo.

3.3.1.1 Extração dos dados do Projeto *1000 Genomes*

As sequências correspondentes às regiões promotora-distal, promotora-proximal, *enhancer L* e 3’NT foram diretamente extraídas do Projeto *1000 Genomes* – Fase 3. Os dados são compostos por um total de 2.504 indivíduos de 26 diferentes populações localizadas na Europa (EUR), na África (AFR), no Sul da Ásia (SAS), no Leste da Ásia (EAS) e nas Américas (AMR). Populações EUR: CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); populações AFR: ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD

(gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações SAS: BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo, residentes no Reino Unido), PJI (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); EAS: CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residents em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e populações AMR: MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru). Para realizar as comparações, foram considerados os seguintes segmentos de nucleotídeos para cada região: i) região promotora-distal – Chr6: 29.792.987 a 29.794.215; ii) região promotora-proximal – Chr6: 29.794.216 a 29.795.621; iii) região *Enhancer L* - Chr6: 29.783.026 a 29.783.146 e iv) região 3'NT – Chr6: 29.798.581 a 29.798.834. O arquivo .vcf foi convertido em um arquivo de dados compatível com o *Software* Arlequin 3.5 (EXCOFFIER, LISCHER, 2010), o qual foi usado para fazer as contagens e detectar os haplótipos compartilhados e exclusivos de cada uma das populações do Projeto *1000 Genomes* (CASTELLI et al., 2014). O programa vcfx (disponível em www.castelli-lab.net/apps/vcfx) foi utilizado para fazer as conversões dos arquivos vcf para os arquivos compatíveis com o *Software* Arlequin 3.5.

3.3.1.2 Revisão de literatura

Uma revisão de literatura foi realizada para a inclusão de todos os dados publicados e disponíveis para as regiões promotora-distal, promotora-proximal e 3'NT em diferentes populações do mundo. Para cada uma das regiões, as populações foram nomeadas e incluídas dentro dos grupos definidos pelo Projeto *1000 Genomes*, conforme descrito a seguir: promotora-distal – BSP (residentes no estado de São Paulo no Brasil DE OLIVEIRA et al., 2018); promotora-proximal - BSP (residentes no estado de São Paulo no Brasil DIAS et al., 2018); 3'NT - populações EUR: FRM (residentes em Marselha na França DICRISTOFARO et al., 2013), FRP (residentes em Paris na França MARTELLI-PALOMINO et al., 2013), POR (residentes em Coimbra em Portugal ALVAREZ et al., 2009); CYP (residentes no Chipre CASTELLI et al., 2017); ITA (residentes em Milão na Itália SIZZANO et al., 2012); populações AFR: MAL (pessoas dos

grupos étnicos Dogon, Peulh, Tamashek e Bambara, de Mali CARLINI et al., 2013), YAN (residentes na República do Congo SABBAGH et al., 2014), SER (residentes no Senegal SABBAGH et al., 2014), GUI (residentes nativos de Guinéa-Bissau ALVAREZ et al., 2009), TOR (residentes em Tori-Bossito no Benin SABBAGH et al., 2014); populações SAS: MDE (pessoas residentes em países do Oriente Médio SIZZANO et al., 2012); populações AMR: BRS (residentes no estado do Rio Grande do Sul no Brasil GRAEBIN et al., 2012; MICHITA et al., 2016; ZAMBRA et al., 2016), BPR (residentes no estado do Paraná no Brasil HANH et al., 2017), BSP (residentes no estado de São Paulo no Brasil CASTELLI et al., 2010; CASTELLI et al., 2017; DE OLIVEIRA et al., 2018; ALBUQUERQUE et al., 2016; FIGUEIREDO FEITOSA et al., 2016), BPE (residentes no estado de Pernambuco no Brasil LUCENA-SILVA et al., 2012). Nos casos em que havia mais de uma amostra populacional para uma mesma região analisada, foi utilizada a média ponderada das frequências dos haplótipos.

3.3.2 Análise estatística

Os haplótipos foram identificados, nomeados e suas frequências calculadas para cada população. As análises estatísticas de Diversidade Haplotípica, Teste de Neutralidade Genético-populacional de Ewens-Watterson, Distância Genética (F_{ST} de Reynolds), Diferenciação Populacional e Análise de Variância Molecular (AMOVA) foram realizados utilizando o software Arlequin 3.5 (EXCOFFIER, LISCHER, 2010). As relações genéticas entre as populações foram avaliadas pela Dispersão em Escala Multidimensional (MDS), a partir dos valores de F_{ST} obtidos pelo cálculo da matriz de distância (distância de Reynolds) utilizando o Software SPSS (IBM). As redes *Median-Joining* de haplótipos das regiões analisadas foram realizadas considerando as variações entre os haplótipos e suas frequências populacionais, utilizando o Software PopART (*Population Analysis with Reticulate Trees*) (BANDELT et al., 1999). O desequilíbrio de ligação (LD) foi calculado entre os pares de SNVs presentes entre a região do *Enhancer L* até a última variante antes do primeiro códon iniciador (ATG) do gene *HLA-G*, utilizando o Software Haploview 4.2 (BARRET et al., 2005). Para isso, foram usados os sítios variáveis com frequência alélica mínima (MAF) $\geq 40\%$ e ponto de corte de 0,05 para HWE. Nas figuras resultantes desta análise, as áreas em vermelho ou cinza escuro indicam LD forte [$\log$ de OR ($LOD \geq 2$, $D' = 1$)]; as áreas em vermelho claro ou cinza claro indicam LD moderada ($LOD \geq 2$, $D' \leq 1$); as áreas em azul ou quase branco indicam LD fraca ($LOD \leq 2$, $D' = 1$); e as áreas brancas indicam que não há LD ($LOD \leq 2$, $D' \leq 1$). D' valores diferentes de 1 são

representados dentro dos quadrados como porcentagens. As variantes do tipo INS/DEL foram ocultadas da análise.

3.4 Relação entre as concentrações de HLA-G solúvel e os haplótipos das regiões regulatórias do gene *HLA-G*

Após a identificação dos principais sítios de variação das regiões regulatórias transcricionais e pós-transcricionais do gene *HLA-G* (promotora-distal, promotora-proximal e 3'NT), as relações entre as concentrações de HLA-G solúvel (sHLA-G) e seus haplótipos mais frequentes foram avaliadas em uma população composta por amostras de indivíduos saudáveis (doadores de sangue), pertencentes a um banco de dados e amostras do nosso grupo de pesquisa em parceria com o laboratório do Professor Dr Celso Teixeira Mendes Junior (FFCLRP-USP). A quantificação das concentrações de sHLA-G, a identificação dos haplótipos e as relações entre eles foram realizadas da seguinte forma:

3.4.1 Quantificação de HLA-G Solúvel

A concentração de sHLA-G no plasma foi avaliada em 80 amostras de indivíduos saudáveis (doadores de sangue), pertencentes a um banco de amostras do grupo de pesquisa do laboratório. A quantificação foi realizada por ELISA de acordo com o protocolo descrito na seção 3.2.1 deste trabalho.

3.4.2 Identificação das variantes das regiões reguladoras do *HLA-G*

A identificação das variantes presentes nas regiões reguladoras promotora-distal, promotora-proximal e 3'NT do *HLA-G*, para as mesmas amostras que foram avaliadas quanto à concentração de sHLA-G, foi obtida no banco de dados anteriormente citado. Com esses dados, foi realizada a identificação dos haplótipos das regiões para cada uma das amostras e as comparações entre as diferentes regiões reguladoras e as concentrações de sHLA-G foram analisadas (DE OLIVEIRA et al., 2018). Os haplótipos identificados nas amostras foram nominados de acordo com publicações prévias (CASTELLI et al., 2010, CASTELLI et al., 2014, SABBAGH et al., 2014, DE OLIVEIRA et al., 2018).

3.4.3 Análise Estatística

Para a avaliação da relação dos dados de sHLA-G e dos haplótipos das regiões reguladoras, inicialmente, foram realizados testes de normalidade para verificar a distribuição

das amostras analisadas. Como as amostras não mostraram distribuição normal ($p < 0,001$), foi utilizado o teste não-paramétrico Kruskal-Wallis para as comparações, pois foram realizadas análises entre mais de dois grupos de haplótipos. Os testes foram realizados no *Software* SPSS (IBM).

4. RESULTADOS

4.1 Avaliação da associação dos alelos do polimorfismo de 14-pb do gene *HLA-G* em doenças de diversas etiologias

4.1.1 Identificação e seleção dos artigos

Os artigos foram selecionados conforme apresentado no fluxograma a seguir (Figura 4.1.1):

Figura 4.1.1 - Fluxograma ilustrando a estratégia de busca usada para identificar os artigos relevantes para a meta-análise.

4.1.2 Características dos artigos incluídos

Os artigos incluídos foram publicados entre 2001 e 2017 e todos eram escritos em língua inglesa. Os estudos foram realizados em populações de 28 diferentes países (Tabela 4.1.1). Dentre os países, o Brasil é o que aparece com o maior número de publicações sobre o assunto (28 artigos).

Tabela 4.1.1 - Relação, em ordem alfabética, do número de artigos por país em que foram realizados.

País	Número de Artigos Publicados
Alemanha	4
Arábia Saudita	1
África do Sul	1
Austrália e Nova Zelândia	1
Brasil	28
Canadá	5
China	18
Coreia	4
Dinamarca	6
Espanha	1
EUA	1
França	2
Grécia	1
Índia	9
Iran	9
Iraque	1
Irlanda	1
Israel	1
Itália	13
Japão	2
México	1
Países Baixos	1
Noruega	1

Polônia	5
Taiwan	1
Tunísia	6
Zâmbia	1

Nesta tabela, está destacado, em negrito, o Brasil, país com maior número de publicações sobre o tema.

A genotipagem do polimorfismo de 14-pb da 3'NT do *HLA-G* foi realizada por diferentes metodologias, sendo a mais usada a PCR seguida de eletroforese em gel de agarose ou poliacrilamida, técnicas utilizadas em torno de 70% dos estudos avaliados (Tabela 4.1.2). Outras técnicas como PCR em Tempo Real (qPCR) e sequenciamento por eletroforese capilar também foram utilizadas.

Tabela 4.1.2 - Relação do número de artigos considerando o método de genotipagem utilizado.

Método de Genotipagem	Número de Artigos
PCR seguida de eletroforese em gel de agarose	48
PCR seguida de eletroforese em gel de poliacrilamida	38
Sequenciamento por eletroforese capilar	22
qPCR	3
PCR (não especificado)	3
HRM	1
<i>Gene Scan</i>	1
<i>Massarray</i>	1
Não informado	8

Todos os artigos incluídos compreendem estudos do tipo caso-controle, envolvendo o polimorfismo de 14-pb em variadas doenças. Os artigos foram agrupados de acordo com as características clínicas, patogênese e resposta imune envolvida, conforme descrito a seguir:

(1) câncer – Linfoma de células B (BIELSKA et al., 2015), Câncer de Mama (ESKANDARI-NASAB et al., 2013; JEONG et al., 2014b; RAMOS et al., 2014; HAGHI et al., 2015; ZIDI et al., 2016a), Câncer de Colo de Útero (FERGUSON et al., 2012; SILVA et al., 2013; BORTOLOTTI et al., 2014;

YANG et al., 2014), Câncer Colorretal (GARZIERA et al., 2016), Câncer de Esôfago (CHEN et al., 2012), Carcinoma Hepatocelular (JIANG et al., 2011; KIM et al., 2013; TEIXEIRA et al., 2013), Carcinoma de Células Escamosas de Cabeça e Pescoço (HNSCC) (AGNIHOTRI et al., 2017), Câncer de Pulmão Cancer (WISNIEWSKI et al., 2015), Neuroblastoma (LAU et al., 2011), Carcinoma de Tireoide Papilífero (DARDANO et al., 2016) e Câncer de Próstata (ZAMBRA et al., 2016); **(2) doenças autoimunes** – Tireoidite de Hashimoto (DARDANO et al., 2016), Artrite Idiopática Juvenil (VEIT et al., 2008), Esclerose Múltipla (KRONER et al., 2007; WIŚNIEWSKI A, BILIŃSKA M, KLIMCZAK A, WAGNER M, MAJORCZYK E, NOWAK I, POKRYSZKO-DRAGAN A, 2010; BEN FREDJ et al., 2016; MOHAMMADI et al., 2016), Vitiligo não-segmentar (JEONG et al., 2014a), Pênfigo Vulgar (GAZIT et al., 2004), Artrite Reumatoide (RIZZO et al., 2006; VEIT et al., 2008, 2014; CATAMO et al., 2014; MARIASELVAM et al., 2015; HASHEMI et al., 2016), Lúpus Eritematoso Sistêmico (LES) (RIZZO et al., 2008; VEIT et al., 2009; WU et al., 2009; CONSIGLIO et al., 2011; PEDROZA et al., 2011; CHEN et al., 2013; LUCENA-SILVA et al., 2013a; CATAMO et al., 2015a; HACHIYA et al., 2016) e Diabetes Tipo 1 (ALBUQUERQUE et al., 2016; GERASIMOU et al., 2016; SILVA et al., 2016); **(3) doenças inflamatórias** – Síndrome Coronariana Aguda (GARCÍA-GONZÁLEZ et al., 2014), Asma Atópica (ZHENG et al., 2010), Doença de Behcet (PARK et al., 2007; SAKLY et al., 2016), Doença Celíaca (FABRIS et al., 2011; CATAMO et al., 2015c), Doença Coronariana (LIN et al., 2007; ZIDI et al., 2016b), Doença de Crohn (GLAS et al., 2007; ZIDI et al., 2015), Doença Renal (MISRA et al., 2013; LIU et al., 2017), Diabetes Tipo 2 (GARCÍA-GONZÁLEZ et al., 2014), e Colite Ulcerativa (GLAS et al., 2007); **(4) doenças infecciosas** – Hepatite B Crônica (HBV) (FERREIRA et al., 2017), Infecção por *Helicobacter pylori* (GENRE et al., 2016), Infecção pelo Vírus da Imunodeficiência Humana (HIV) (SEGAT et al., 2010), Transmissão vertical de HIV (FABRIS et al., 2009; SEGAT et al., 2009, 2014; HONG et al., 2015), Vírus Linfotrópico de Células T Humanas (HTLV-1) (HADDAD et al., 2011), Citomegalovírus Humano (HCMV) (ZHENG et al., 2009; JIN et al., 2012; MISRA et al., 2013), Papilomavírus Humano (HPV) (SIMÕES et al., 2009; BORTOLOTTI et al., 2014; XU et al., 2014; YANG et al., 2014), Hanseníase (LUCENA-SILVA et al., 2013b) e Choque Séptico (GRAEBIN et al., 2012); **(5) rejeição a transplantes** – Rejeição Aguda ao Aloenxerto (JIN et al., 2012; MISRA et al., 2013), Transplante de Células Hematopoiéticas Alogênicas (WATERHOUSE et al., 2013), Transplante Alogênico de Células-tronco Hematopoiéticas (Allo-HSCT) (CHIUSOLO et al., 2011), Transplante de Medula Óssea (BOUKOUACI et al., 2011), Transplante Cardíaco (TORRES et al., 2009; LAZARTE et al., 2016), Transplante Hepático (AZARPIRA et al., 2014; THUDE et al., 2016)

e Transplante Renal (CRISPIM et al., 2008; AGHDAIE et al., 2011; CILIÃO ALVES et al., 2012); **(6) doenças reprodutivas** – Falha de Implantação (HVIID et al., 2004; SIPAK-SZMIGIEL et al., 2009; ENGHELABIFAR et al., 2014; LASHLEY et al., 2014; NARDI et al., 2016), Pré-eclâmpsia (BRIEN et al., 2001; HYLENIUS et al., 2004; LIN et al., 2006; VIANNA et al., 2007; IVERSEN et al., 2008; MOREAU et al., 2008; LARSEN et al., 2010; ZHANG et al., 2012; JAHAN et al., 2014; QUACH et al., 2014; MANDO et al., 2016; NILSSON et al., 2016), Aborto Recorrente (HVIID et al., 2002, 2004; TRIPATHI et al., 2004; YAN et al., 2006; KANO et al., 2007; XUE et al., 2007; LIU et al., 2008; SURYANARAYANA et al., 2008; ZHU et al., 2010; BERGER et al., 2010; SHANKARKUMAR et al., 2011; SHAO et al., 2011; VARGAS et al., 2011; ARUNA et al., 2011; AKHTER et al., 2012; CHRISTIANSEN et al., 2012; JASSEM et al., 2012; AFKHAMI et al., 2014; ARJMAND; SAMADI, 2015; MICHITA et al., 2016; AMODIO et al., 2016), Falha Reprodutiva (SIPAK-SZMIGIEL et al., 2007, 2009) e Aborto Espontâneo (MOREAU et al., 2008; SHANKARKUMAR et al., 2011; OMAR et al., 2015; NOWAK et al., 2016). Adicionalmente, dois grandes grupos de doenças foram avaliados: **(A) doenças associadas com hipotética alta expressão de *HLA-G***, compreendendo artigos sobre câncer e doenças infecciosas; e **(B) doenças associadas com hipotética baixa expressão de *HLA-G***, incluindo artigos sobre doenças autoimunes, doenças inflamatórias, rejeição a transplantes e doenças reprodutivas.

4.1.3 Dados gerais das meta-análises

Em todas as análises, foram utilizados somente os dados de frequência dos alelos do polimorfismo de 14-pb da 3'NT do gene *HLA-G*. Alguns artigos aparecem em mais de uma análise, pois esses estudos possuem dados de diferentes grupos de pacientes ou diferentes doenças. Inicialmente, foram analisados dois grandes grupos (A e B); em seguida, foram avaliados os seis grupos previamente descritos (1 a 6). As características de todos os artigos incluídos e os principais dados obtidos nessas meta-análises estão apresentados na Tabela 4.1.3. Na tabela, conforme descrito na seção Material e Métodos desta tese, nos casos em que $I^2 > 50\%$ ou Teste Q $p < 0,01$, somente o REM foi aplicado nos dados, pois esse modelo usualmente fornece valores de intervalo de confiança mais confiáveis que o FEM. Devido às características similares entre as doenças, os grupos de doenças autoimunes e inflamatórias foram também analisados conjuntamente. No entanto, não foram verificadas diferenças significativas entre as análises separadas e conjuntas desses dois grupos (-14-pb; OR = 1,04; 95%CI = 0,96-1,12; $p = 0,385$). Isso também aconteceu quando o grupo de Rejeição a

transplantes foi estratificado em quatro diferentes subgrupos para comparações: transplantes sólidos: (+14-bp; OR = 1,23; 95%CI = 0,77-1,98; p = 0,385), transplantes celulares (+14-bp; OR = 1,04; 95%CI = 0,84-1,29; p = 0,717), pacientes *versus* doadores (+14-pb; OR = 1,20; 95%CI = 0,69-2,11; p = 0,514) e transplantes com ou sem rejeição (+14-pb; OR = 1,10; 95%CI = 0,72-1,69; p = 0,655). Por fim, foram realizadas meta-análises para cada situação clínica/doença que tivesse três ou mais artigos publicados (Câncer de Mama, Câncer de Colo de Útero, Carcinoma Hepatocelular, Esclerose Múltipla, Artrite Reumatoide, Lúpus Eritematoso Sistêmico, Diabetes Tipo 1, Transmissão Vertical de HIV (os alelos das mães foram avaliados), Citomegalovírus Humano, Papilomavírus Humano, Transplante Cardíaco, Transplante Renal, Pré-eclâmpsia (quando os alelos maternos foram avaliados), Pré-eclâmpsia (quando os alelos dos filhos foram avaliados), Falha de implantação e Aborto).

Considerando a presença do polimorfismo de 14-pb da 3'NT do gene *HLA-G* como fator de risco, foram verificados resultados significativos para o grupo chamado Doenças Reprodutivas, que inclui Falha de Implantação, Pré-eclâmpsia, Aborto Recorrente e Aborto Espontâneo (Figura 4.1.2) e para as seguintes doenças analisadas individualmente: Pré-eclâmpsia – quando os alelos dos filhos foram avaliados (Figura 4.1.3A) e Lúpus Eritematoso Sistêmico (Figura 4.1.3B). Por outro lado, considerando a ausência dos 14-pb da 3'NT do gene *HLA-G* como fator de risco, foram observados resultados significativos apenas para duas condições: Câncer de Mama (Figura 4.1.4A) e infecção por Citomegalovírus Humano (Figura 4.1.4B).

O objetivo desta meta-análise foi realizar análise global dos dados, apesar disso as diferenças genéticas entre populações ao redor do mundo foram levadas em consideração. Todos os dados foram reanalisados, excluindo grupos populacionais específicos que poderiam interferir nos resultados obtidos (por exemplo: população Chinesa Han). Entretanto, não foram observadas alterações nos resultados obtidos quando essas populações foram excluídas (dados não apresentados). A única exceção foi observada na meta-análise acerca de infecções por Citomegalovírus, na qual duas das três publicações incluídas eram a respeito da população Chinesa Han, não sendo, portanto, possível fazer a exclusão do grupo populacional nessa avaliação. Por fim, sobre viés de publicação, foi verificado resultado significativo apenas na meta-análise realizada para o grupo de Doenças Autoimunes (p = 0,0009).

4.1.4 Associações entre os alelos do polimorfismo de 14-pb com os haplótipos das regiões promotora-proximal, codificadora e 3'NT do gene *HLA-G* e proteínas

Para compreender as relações entre o polimorfismo de 14-pb e outros sítios polimórficos do *HLA-G*, foi organizada a Tabela 4.1.4, que contém as principais associações entre os alelos do polimorfismo de 14-pb e os haplótipos mais frequentes para as regiões promotora, codificadora e 3'NT e para a proteína *HLA-G*. Na primeira coluna, estão indicados os dois alelos do polimorfismo de 14-pb (ausência/deleção e presença/inserção). Na segunda coluna estão classificadas as regiões NT mais frequentes para cada alelo. Na terceira coluna estão descritos os haplótipos mais frequentes da região promotora, na qual é possível destacar que os haplótipos pertencentes ao mesmo grupo de haplótipos aparecem associados com os dois alelos do 14-pb (por exemplo, o promotor 0104 com ausência dos 14-pb e o promotor 0102 com presença dos 14-pb). Na quarta coluna, estão descritos os haplótipos mais frequentes da região codificadora do gene para cada alelo do polimorfismo de 14-pb. Na última coluna, estão listadas as proteínas mais frequentes, na qual pode-se observar que a proteína G*01:01 é a mais frequente para ambos alelos do polimorfismo de 14-pb. Além disso, embora diferentes proteínas estejam associadas aos alelos de 14-pb, as suas frequências são geralmente baixas e estudos mostram que estas variantes mantêm as funções do *HLA-G*.

Tabela 4.1.3 - Resultados de todas as meta-análises realizadas no presente trabalho.

Grupos de situações clínicas/doenças	Alelos 14-pb	Números				Modelo Fixo			Modelo Randômico			Heterogeneidade (I ² , τ ₂ , p)	Viés de Publicação	
		Estudos	Situações		Casos	Controles	OR	IC	p	OR	IC		P	Teste de Egger
			clínicas/ Doenças											
Doenças associadas com hipotética alta expressão de <i>HLA-G</i>	Del	32	20	6091	8687				1,09	0,98-1,21	0,097	74%, 0,072, <0,01	0,3109	
	Ins								0,92	0,83-1,02	0,097			
Doenças associadas com hipotética alta expressão de <i>HLA-G</i>	Del	107	32	16641	20721				0,98	0,92-1,05	0,617	75%, 0,091, <0,01	0,2346	
	Ins								1,02	0,95-1,09	0,620			
Câncer	Del	21	11	3815	5802				0,94	0,84-1,05	0,312	68%, 0,046, <0,01	0,5915	
	Ins								1,06	0,95-1,19	0,312			

Doenças Autoimunes	Del	27	8	6008	6866	0,99	0,92-1,08	0,089	55%, 0,024, <0,01	0,0009
	Ins					1,01	0,93-1,09	0,089		
Doenças Inflamatórias	Del	15	9	2942	4889	1,11	0,94-1,32	0,224	82%, 0,087, <0,01	0,1447
	Ins					0,90	0,76-1,07	0,225		
Doenças Autoimunes e Inflamatórias	Del	42	17	8950	11755	1,04	0,96-1,12	0,385	10%, 0,045, <0,01	0,0014
	Ins					0,97	0,89-1,05	0,386		
Doenças Infecciosas e Transmissão Vertical de HIV	Del	17	9	2276	2885	1,14	0,93-1,39	0,196	79%, 0,133, <0,01	0,5016
	Ins					0,88	0,72-1,07	0,196		
Rejeição a Transplantes	Del	15	8	1226	1507	0,87	0,61-1,23	0,429	89%, <0,01, 0,415	0,127
	Ins					1,15	0,81-1,64	0,429		
Transplantes Sólidos	Del	11		908	1115	0,81	0,51-1,30	0,385	91%, 0,565, <0,01	*
	Ins					1,23	0,77-1,98	0,385		
Transplantes Celulares	Del	4		318	392	0,87	0,57-1,30	0,716	0%, 0, 0,91	*
	Ins					1,04	0,84-1,29	0,717		

Pacientes <i>versus</i>	Del								1,46	1,28-1,68	0,514		
Controles	Ins	7		857	804				1,20	0,69-2,11	0,515	93%, 0,515,>0,01	*
Transplantes	Del								0,91	0,75-1,10	0,653		
com rejeição		8		369	703							79%, 0,302, <0.01	*
<i>versus</i> sem	Ins								1,10	0,72-1,69	0,655		
rejeição													
Doenças	Del								0,89	0,81-0,99	0,034		
Reprodutivas	Ins	45	7	4189	4574				1,12	1,01-1,24	0,034	58%, 0,065, <0,01	0,701
Câncer de	Del					1,23	1,06-1,42	0,006	1,23	1,06-1,42	0,006		
Mama	Ins	5	-	726	811	0,81	0,70-0,94	0,006	0,81	0,70-0,94	0,006	0%, 0, 0,45	*
Câncer de Colo	Del								1,04	0,78-1,38	0,787		
de Útero	Ins	4	-	592	1383				0,96	0,72-1,28	0,787	65%, 0,051, 0,04	*
Carcinoma	Del								1,14	0,76-1,69	0,529		
Hepatocelular	Ins	3	-	518	1071				0,88	0,59-1,31	0,529	79%, 0,096, <0,01	*
Esclerose	Del					0,93	0,80-1,08	0,322	0,93	0,80-1,08	0,322		
Múltipla	Ins	4	-	799	705	1,08	0,93-1,26	0,322	1,08	0,93-1,26	0,322	0%, 0, 0,91	*

Artrite Reumatoide	Del	6	-	1501	1486	1,04	0,94-1,15	0,462	1,04	0,94-1,15	0,462	0%, 0, 0,82	*
	Ins					0,96	0,87-1,07	0,463	0,96	0,87-1,07	0,463		
Lupus Eritematoso Sistêmico	Del								0,88	0,79-0,99	0,028		
	Ins	10	-	2906	3412				1,13	1,01-1,27	0,028	51%, 0,016, 0,03	0,314
Diabetes Tipo 1	Del	3	-	361	412	1,24	1,01-1,52	0,036	1,24	1,01-1,52	0,049	12%, 0,005, 0,32	*
	Ins					0,80	0,66-1,00	0,036	0,80	0,66-1,00	0,049		
Transmissão Vertical de HIV	Del	5	-	468	702				1,08	0,60-1,95	0,798	88%, 0,386, <0,01	*
	Ins								0,93	0,51-1,67	0,798		
Citomegalovírus Humano	Del	3	-	198	642	2,06	1,60-2,64	<0,0001	2,06	1,60-2,64	<0,0001	0%, 0, 0,76	*
	Ins					0,49	0,38-0,63	<0,0001	0,49	0,38-0,63	<0,0001		
Papilomavírus Humano	Del	4	-	593	702				1,17	0,89-1,56	0,270	60%, 0,049, 0,06	*
	Ins								0,85	0,64-1,13	0,270		
Transplante Cardíaco	Del	3	-	337	261	0,92	0,73-1,17	0,504	0,92	0,73-1,17	0,507	0%, 0, 0,43	*
	Ins					1,08	0,86-1,37	0,504	1,08	0,86-1,37	0,507		
Transplante Renal	Del	3	-	165	166	1,19	0,87-1,64	0,281	1,19	0,87-1,64	0,281	0%, 0, 0,70	*
	Ins					0,84	0,61-1,15	0,281	0,84	0,61-1,15	0,281		

Pré-eclâmpsia (mães)	Del								0,86	0,72-1,04	0,116	54%, 0,052, <0,01	0,227
	Ins	12	-	1059	1457				1,16	0,96-1,39	0,116		
Pré-eclâmpsia (filhos)	Del					0,71	0,52-0,96	0,026	0,71	0,52-0,96	0,026	0%, 0, 0,70	*
	Ins	3	-	122	299	1,42	1,04-1,93	0,026	1,42	1,04-1,93	0,026		
Falha de Implantação	Del								1,38	0,60-3,18	0,447	83%, 0,597, <0,01	*
	Ins	4	-	127	214				0,72	0,31-1,67	0,447		
Aborto	Del								0,89	0,79-1,01	0,075	54%, 0,048, <0,01	0,786
	Ins	24	-	2773	2475				1,12	0,99-1,27	0,077		

As possíveis associações entre o polimorfismo de 14-pb e as situações clínicas/doenças foram mensuradas pelo cálculo do *Odds Ratio* (OR) e Intervalo de Confiança de 95% (CI), para os modelos de efeito randômico e fixo. Os valores de p resultantes do Teste Q e os valores de τ^2 e I^2 foram usados para mensurar a heterogeneidade entre os estudos avaliados em cada meta-análise. Se $I^2 > 50\%$ ou $p < 0.01$ no Teste Q test, somente o modelo de efeito randômico foi aplicado nos dados. Nos outros casos, tanto o modelo fixo quanto o modelo randômico foram calculados. Os resultados das meta-análises que mostraram associação entre o polimorfismo de 14-pb da região 3'NT do gene *HLA-G* com as situações clínicas/doenças avaliadas estão destacados em negrito.

*Devido aos parâmetros definidos para este teste estatístico, o Viés de Publicação não foi avaliado em grupos com menos de 10 estudos.

Figura 4.1.2 - Resultados da meta-análise do grupo Doenças Reprodutivas.

Nesta figura, *Experimental* refere-se aos dados dos pacientes, *Controls* aos dados dos controles, *Events* ao número de alelos de risco e *Total* ao número total de alelos. Estão descritos os valores de *Odds Ratio* (OR) e o Intervalo de Confiança de 95% (CI) para os modelos de efeito fixo e randômico. Nos casos em que os valores de p resultantes do Teste Q foram <0.01 e o valor de I^2 foi $> 50\%$, os valores do modelo de efeito randômico foram aplicados para esses dados.

Figura 4.1.3 - Resultados das meta-análises em Eclâmpsia – quando os alelos dos filhos foram avaliados (A) e Lúpus Eritematoso Sistêmico (B).

Nesta figura, *Experimental* refere-se aos dados dos pacientes, *Controls* aos dados dos controles, *Events* ao número de alelos de risco e *Total* ao número total de alelos. Estão descritos os valores de *Odds Ratio* (OR) e o Intervalo de Confiança de 95% (CI) para os modelos de efeito fixo e randômico. Nos casos em que os valores de p resultantes do Teste Q foram <0.01 e o valor de I^2 foi $> 50\%$, os valores do modelo de efeito randômico foram aplicados para esses dados.

Figura 4.1.4 - Resultados das meta-análises de Câncer de Mama (A) e Infecção por Citomegalovírus Humano (B).

A **B**

Nesta figura, *Experimental* refere-se aos dados dos pacientes, *Controls* aos dados dos controles, *Events* ao número de alelos de risco e *Total* ao número total de alelos. Estão descritos os valores de *Odds Ratio* (OR) e o Intervalo de Confiança de 95% (CI) para os modelos de efeito fixo e randômico. Nos casos em que os valores de p resultantes do Teste Q foram <0.01 e o valor de I^2 foi $> 50\%$, os valores do modelo de efeito randômico foram aplicados para esses dados.

Tabela 4.1.4 – Principais associações entre os alelos do polimorfismo de 14-pb e os haplótipos mais frequentes das regiões promotora-proximal, codificadora e 3'NT do gene *HLA-G* e proteínas.

14-pb	<i>HLA-G</i>	<i>HLA-G</i>	<i>HLA-G</i>	<i>HLA-G</i>
	3'NT	Promotora- proximal	Região Codificadora	Proteína
		010101a		
	UTR-1	010101b	G*01:01:01:01	
14-bp	UTR-3	010101c	G*01:01:01:04	G*01:01
ausência	UTR-4	010101d	G*01:01:01:05	G*01:04
(deleção)	UTR-6	010101f	G*01:04:01	
	UTR-18	0104a	G*01:04:04	
		0104b		
			G*01:01:02:01	G*01:01
14-bp	UTR-2	010102a	G*01:03:01:01	G*01:03
presença	UTR-5	0103a	G*01:01:03:01	G*01:05N
(inserção)	UTR-7	0103d	G*01:05N	G*01:06
			G*01:06	

4.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene *HLA-G*, em um estudo caso-controle para a Doença de Chagas

4.2.1 Expressão de sHLA-G no plasma

De acordo com os dados obtidos, existem diferenças significativas de expressão de sHLA-G entre pacientes com Doença de Chagas (Casos, $n = 80$) e indivíduos saudáveis (Controles, $n = 161$). Essas diferenças foram observadas tanto nas comparações entre os Grupos Casos e Controles (Figura 4.4.1), quanto nas comparações dos subgrupos de formas da doença (Cardíaca – $n = 64$; Digestiva – $n = 6$; e Mista – $n = 10$) separados (Figura 4.4.2), sendo a única exceção o subgrupo da Forma Digestiva.

Figura 4.2.1 – Comparação dos níveis de expressão de sHLA-G nos grupos Casos e Controles.

Nesta figura estão apresentados os níveis de sHLA-G encontrados nos grupos Casos ($n = 80$; $\bar{x} = 9,87$; $s = 9,03$) e Controles ($n = 161$; $\bar{x} = 4,54$; $s = 5,63$); $p \leq 0,001$.

Figura 4.2.2 – Comparação dos níveis de expressão de sHLA-G nos grupos Controles e Casos, subdivididos por formas da doença de Chagas.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos grupos Controles ($n = 161$) e Casos Forma Cardíaca ($n = 64$), Casos Forma Digestiva ($n = 6$) e Casos Forma Mista ($n = 10$). (A) Controles x Forma Cardíaca X Forma Digestiva X Casos Forma Mista – $p \leq 0,001$; (B) Controles x Forma Cardíaca – $p \leq 0,001$; (C) Controles x Forma Digestiva – $p = 0,492$ e (D) Controles x Forma Mista – $p = 0,01$. $\bar{x}$ s e desvios-padrões dos grupos: Controles - $\bar{x} = 4,54$; $s = 5,63$; Forma Cardíaca – $\bar{x} = 9,75$; $s = 9,11$; Forma Digestiva – $\bar{x} = 6,85$; $s = 4,77$; Forma Mista – $\bar{x} = 12,51$; $s = 10,35$.

4.2.2 Expressão de HLA-G em tecidos

Para efeitos de comparação com os dados obtidos, foi verificado no Atlas de Proteínas Humanas (*The Human Protein Atlas*; <https://www.proteinatlas.org/ENSG00000204632-HLA-G/tissue>) que nenhum dos tecidos dos órgãos avaliados neste trabalho (Cólon, Coração e Esôfago) possuem expressão de HLA-G em situações normais. As amostras teciduais de pacientes chagásicos foram avaliadas quanto à intensidade de marcação e quanto ao número de células marcadas. A Figura 4.2.3 ilustra os três tipos de tecidos avaliados.

Figura 4.2.3 - Representação dos três tipos de tecidos avaliados quanto à expressão de HLA-G.

Nesta figura, estão representados os três tipos de tecidos de amostras de pacientes chagásicos que foram avaliados quanto à expressão de HLA-G: (A) Cólon, (B) Coração e (C) Esôfago.

As análises imunoistoquímicas mostraram que o Cólon possui maiores intensidades de marcação de HLA-G quando comparadas com a marcação dos outros dois tipos de tecidos. As diferenças significativas, verificadas quando os Grupos Cólon e Coração foram comparados, estão apresentadas na Figura 4.2.4. Nesse gráfico estão representadas também as variações de concentração, médias e desvios-padrões dos grupos amostrais.

Figura 4.2.4 – Comparação dos níveis de expressão tecidual de HLA-G em amostras de tecidos de pacientes acometidos pela Doença de Chagas.

Neste gráfico está representada a expressão de HLA-G obtida pelas médias de intensidade de marcação (em unidades arbitrárias) encontradas nas amostras de tecidos de Cólón ($n = 17$; $\bar{x} = 80068$; $s = 17862$), Coração ($n = 69$; $\bar{x} = 65433$; $s = 15952$) e Esôfago ($n = 7$; $\bar{x} = 70237$; $s = 10117$) de pacientes com Doença de Chagas; $p = 0,005$.

Conforme descrito, as amostras teciduais foram, adicionalmente, classificadas quanto ao número de células marcadas. De acordo com os dados apresentados na Tabela 4.2.1, é possível observar que, nos resultados obtidos o número de células marcadas varia bastante dentro de cada tipo de tecido.

Tabela 4.2.1 – Classificação das amostras teciduais quanto ao número de células expressando HLA-G.

Amostra (n)	Classificação	Número de Amostras (%)
Cólón (17)	0	1 (6%)
	1	1 (6%)
	2	2 (12%)
	3	3 (18%)

	4	10 (58%)
Coração (69)	0	2 (3%)
	1	11 (16%)
	2	18 (26%)
	3	16 (23%)
	4	22 (32%)
Esôfago (7)	0	0 (0%)
	1	2 (29%)
	2	0 (0%)
	3	3 (42%)
	4	2 (29%)

Nesta tabela estão descritos os subtipos de amostras, a classificação numérica de acordo com o número de células marcadas, o número de amostras e os seus valores percentuais correspondentes, para cada uma das classificações (0 - casos em que não houve identificação de expressão de HLA-G; 1 – quando 1-25% das células expressavam a proteína; 2 - 26-50% das células estavam marcadas; 3 - de 51-75% das células positivas; e 4 - casos com 76-100% das células marcadas para HLA-G.

4.2.3 Variabilidade genética do *HLA-G*

Os dados obtidos após o sequenciamento das amostras de sangue de pacientes chagásicos foram avaliados para verificar possíveis associações entre os polimorfismos do gene *HLA-G* e a doença, inicialmente, em um único grupo que abrangeu todas as amostras. Entretanto, nessas análises realizadas não foram encontradas associações entre os haplótipos e/ou pontos de variação das regiões reguladoras avaliadas (promotora-proximal e 3'NT) e a doença. Posteriormente, os dados foram avaliados considerando as diferentes formas da doença (cardíaca, digestiva e indeterminada). Da mesma forma, não foram observadas associações entre os polimorfismos e/ou haplótipos do gene *HLA-G* e os subtipos da doença. Sendo assim, optou-se por apresentar somente os resultados referentes ao grupo total de amostras, os quais estão descritos nas tabelas 4.2.2, 4.2.3, 4.2.4 e 4.2.5. Nas tabelas, estão descritos os resultados das frequências e dos valores de OR e IC obtidos após as análises de associação para os haplótipos e para os principais pontos de variação encontrados nas duas regiões regulatórias. Para possibilitar as comparações com os dados do grupo controle, os haplótipos referentes à região promotora-proximal foram agrupados de acordo com os principais grupos conhecidos (010101, 010102, 0103e 0104).

Tabela 4.2.2 – Frequências e valores de OR e IC dos grupos de haplótipos da região promotora-proximal do gene *HLA-G*.

Haplótipos	010101	010102	0103	0104	Outros	2n
Controles	0,3956	0,3655	0,0836	0,1538	0,0015	670
Chagas	0,4114	0,3200	0,0829	0,1600	0,0257	350
OR	1,05	0,83	0,99	1,06	8,82	
IC 95%	0,81 - 1,36	0,63 - 1,10	0,62 - 1,58	0,74 - 1,51	1,89 - 41,03	

Nesta tabela estão apresentados os dados de frequência, valor de OR e IC 95% de cada um dos grupos de haplótipos da região promotora-proximal para os grupos de casos e controles.

Tabela 4.2.3 – Frequências e valores de OR e IC de pontos de variação da região promotora-proximal do gene *HLA-G*.

Posição hg19 SNPid Alelo Referência	29826540 rs1736936 G	29826666 rs1736935 A	29826690 rs3823321 G	29826705 rs1736934 A	29826707 rs17875389 A	29826724 rs3115630 T	29826747 rs146374870 G	29826881 rs1632947 G	29827083 rs1632946 C
Casos	0,5030	0,4140	0,8290	0,6740	0,9110	0,0340	0,9910	0,5030	0,5030
Controles	0,4962	0,4148	0,8422	0,6501	0,9224	0,0445	0,9898	0,4924	0,4924
OR	1,03	1,00	0,91	1,11	0,87	0,76	1,22	1,04	1,04
IC 95%	0,80 - 1,33	0,77 - 1,30	0,64 - 1,29	0,84 - 1,46	0,54 - 1,38	0,38 - 1,50	0,31 - 4,75	0,80 - 1,35	0,80 - 1,35
Posição hg19 SNPid Alelo Referência	29827120 rs1233334 G	29827129 rs2249863 T	29827156 rs2735022 A	29827179 rs35674592 G	29827199 rs17875391 A	29827212 rs1632944 G	29827644 rs1233333 G	29827789 rs17875397 C	2n
Casos	0,0890	0,5030	0,5030	0,5030	0,9660	0,5030	0,5090	0,9110	350
Controles	0,1081	0,4911	0,4911	0,4911	0,9669	0,4911	0,4924	0,9224	670
OR	0,81	1,05	1,05	1,05	0,96	1,05	1,07	0,87	
IC 95%	0,52 - 1,26	0,81 - 1,36	0,81 - 1,36	0,81 - 1,36	0,47 - 1,96	0,81 - 1,36	0,82 - 1,38	0,54 - 1,38	

Nesta tabela estão apresentados os dados de frequência, valor de OR e IC 95% de alguns pontos de variação da região promotora-proximal para os grupos de casos e controles.

Tabela 4.2.4 – Frequências e valores de OR e IC dos haplótipos da região 3'NT do gene *HLA-G*.

Haplótipos	UTR-01	UTR-02	UTR-03	UTR-04	UTR-05	UTR-06 +UTR-18	2n
Chagas	0,2614	0,2844	0,1637	0,0805	0,0891	0,0374	348
Controles	0,2507	0,3118	0,1477	0,1135	0,0746	0,0329	670
OR	1,06	0,88	1,13	0,68	1,21	1,14	
IC 95%	0,79 - 1,42	0,66 - 1,17	0,79 - 1,61	0,43 - 1,08	0,76 - 1,94	0,57 - 2,30	
Haplótipos	UTR-07	UTR-13	UTR-15	UTR-17	UTR-23	UTR-44	2n
Chagas	0,0460	0,0029	0,0029	0,0259	0,0029	0,0029	348
Controles	0,0493	0,0030	0,0015	0,0090	0,0030	0,0030	670
OR	0,93	0,96	1,93	2,94	0,96	0,96	
IC 95%	0,50 - 1,71	0,09 - 10,65	0,12 - 30,92	1,04 - 8,32	0,09 - 10,65	0,09 - 10,65	

Nesta tabela estão apresentados os dados de frequência, valor de OR e IC 95% de cada um haplótipos da região 3'NT para os grupos de casos e controles.

Tabela 4.2.5 – Frequências e valores de OR e IC de pontos de variação da região 3'NT do gene *HLA-G*.

hg19 SNPid Alelo Referência	29798581 rs371194629 G	29798608 rs567747015 C	29798610 rs1707 C	29798617 rs1710 G	29798642 rs17179108 C	29798794 rs9380142 A	2n
Casos	0,5776	0,9971	0,0805	0,4080	0,8563	0,7356	348
Controles	0,5776	0,9898	0,1196	0,4148	0,8715	0,7379	670
OR	1,00	3,66	0,65	0,97	0,88	0,99	
IC 95%	0,77 - 1,30	0,45 - 29,90	0,41 - 1,01	0,75 - 1,26	0,60 - 1,28	0,74 - 1,33	

Nesta tabela estão apresentados os dados de frequência, valor de OR e IC 95% de alguns pontos de variação da região 3'NT para os grupos de casos e controles.

4.3 Caracterização das regiões regulatórias do gene *HLA-G* na população mundial

4.3.1 Região Promotora-Distal

Todos os haplótipos identificados nas análises dos dados populacionais disponíveis no Projeto *1000 Genomes* e descritos em publicações prévias estão listados na Tabela 4.3.1. Um total de 48 variações SNV que compõem 48 diferentes haplótipos foram identificados para a região promotora-distal. Desses, somente 8 apresentaram frequências maiores do que 1% na população mundial avaliada e 29 foram identificados pela primeira vez neste trabalho. As frequências dos haplótipos em cada uma das populações avaliadas estão descritas na Tabela 4.3.1 e os haplótipos com frequências maiores que 1% estão representados na Figura 4.3.1. Adicionalmente, foi criada, a partir dos dados de frequência e as variações nucleotídicas dos haplótipos, uma rede de haplótipos *Median-Joining* que está apresentada na Figura 4.3.2. Nessa figura é possível verificar, graficamente, os haplótipos mais frequentes (010101a, 010102a, 0104a e 010102b) e as suas variações dentro dos grupos populacionais.

Nas análises populacionais, foi verificado que a diversidade haplotípica da região promotora-distal variou entre 0,7992 (ASW) a 0,5506 (FIN), quando todas as populações foram analisadas conjuntamente. Em cada um dos cinco grupos, os valores máximos e mínimos desse parâmetro foram os seguintes: populações americanas (AMR) 0,7787 (MXL) a 0,7271 (PEL); populações europeias (EUR) 0,7202 (TSI) a 0,5506 (FIN); populações africanas (AFR) 0,7992 (ASW) a 0,7753 (YRI); populações sul-asiáticas (SAS) 0,7868 (ITU) a 0,7385 (GIH); e populações leste-asiáticas (EAS) 0,7664 (KHV) a 0,6324 (CDX).

No teste de neutralidade genético-populacional de Ewens-Watterson, foram avaliados os valores *F* esperados e observados e a significância de Slatkin. Os valores de *F* observados para essa região estão dentro do intervalo de confiança, sendo que alguns dos valores observados são mais altos e outros mais baixos do que os valores esperados, e nenhuma das populações avaliadas foi verificado diferença significativa entre os valores.

Os valores de distância genética (F_{ST} de Reynolds) foram calculados nas populações aos pares e os dados foram utilizados para realizar a análise da Dispersão em Escala Multidimensional (MDS). O resultado obtido desta análise está apresentado na Figura 4.3.2. O primeiro gráfico (A) apresenta os dados de todas as populações avaliadas, enquanto que, o segundo gráfico (B) apresenta somente os dados das populações do Projeto *1000 Genomes*.

Em ambos, é possível observar que, tendo por base a região promotora-distal, as populações FIN, CHS, CDX e JPT são as mais distantes geneticamente.

No teste de diferenciação entre todos os pares de populações foram observadas diferenças significativas em 320 das 351 comparações realizadas. Por fim, na análise de variância molecular (AMOVA), quando todas as populações avaliadas e considerando os cinco grupos populacionais, foi observado que 95,29% da variabilidade genética está dentro das populações, 3,12% entre grupos, 1,59% entre populações dentro dos grupos e valor de F_{ST} de 4,7%.

As tabelas contendo os dados completos das análises populacionais para a região promotora-distal estão apresentadas no Apêndice A.

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 1A).

Posição chr6 hg19	29792987	29792989	29792990	29792991	29792996	29793021	29793062	29793100	29793105	29793160	29793162	29793176	29793203	29793224	29793243	29793276	29793309	29793310	29793324	29793361	29793362	29793365	29793385	29793404	29793453
Identificação do SNV	rs542428384	.	.	.	rs1632950	rs1736942	rs558600852	rs571891839	rs185842304	rs190164389	rs150426937	rs543301808	rs200585067	rs114251653	rs545198729	rs565008020	rs9278332	rs1736941	rs543960498	rs560849637	rs1736940	rs72838218	rs1736939	rs142798055	rs532106939
SNV	-2635	2633	-2632	-2631	-2626	-2601	-2560	-2522	-2517	-2462	-2460	-2446	-2419	-2398	-2379	-2346	-2313	-2312	-2298	-2261	-2260	-2257	-2237	-2218	-2169
Alelo Referência	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Alelo Alternativo	C	T	C	G	T	G	T	T	T	C	G	T	T	T	T	T	T	G,*	T	C	A	A	A	C	A
Distal-010101a	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010101b	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010101c	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010101d	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010101e	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010101f	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Distal-010102a	T	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Distal-010102b	T	C	T	A	T	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Distal-010102c	T	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	T	*	A	G	A	G	A	CTCT	C
Distal-010102d	T	C	T	A	T	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	A	G	A	CTCT	C
Distal-010102e	T	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	T	*	A	G	A	G	A	CTCT	C
Distal-010102f	C	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Distal-0103a	T	C	T	A	C	A	G	C	C	T	G	A	C	T	G	C	TA	A	A	G	A	G	G	CTCT	C
Distal-0103b	T	C	T	A	C	A	G	C	C	T	G	A	C	T	G	C	TA	A	A	G	A	G	G	CTCT	C
Distal-0104a	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	A	G	C	C
Distal-0104b	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	G	CTCT	C
Distal-0104c	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	A	G	C	C
Distal-0104d	T	T	C	G	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	A	G	C	C
Distal-0104e	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	G	C	C

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 1B).

Posição chr6 hg19	29793460	29793473	29793503	29793542	29793577	29793581	29793590	29793608	29793610	29793630	29793650	29793724	29793761	29793769	29793785	29793876	29793942	29793947	29793982	29794008	29794049	29794097	29794146	Frequência dos haplótipos na População Mundial
Identificação do SNV	rs9461563	rs373926705	rs146027593	rs557634203	.	rs1233335	rs1736938	rs138910665	rs72492280	rs9258492	rs182479069	rs1632949	rs544293020	rs150381639	rs529879422	rs1736937	rs138915780	rs192306068	rs114712708	rs142134285	rs1632948	.	rs114252012	
SNV	-2162	-2149	-2119	-2080	-2045	-2041	-2032	-2014	-2012	-1992	-1972	-1898	-1861	-1853	-1837	-1746	-1680	-1675	-1640	-1614	-1573	-1525	-1476	
Alelo Referência	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	2n = 5070
Alelo Alternativo	C	G	C	A	C	G	G	A	A	G	T	C	T	G	T	A	G	T	C	T	C	C	A	
Distal-010101a	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,3491
Distal-010101b	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0254
Distal-010101c	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	G	C	T	C	T	T	C	0,0032
Distal-010101d	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	T	T	T	C	0,0150
Distal-010101e	T	C	T	T	C	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0000
Distal-010101f	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	C	C	0,0000
Distal-010102a	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,2287
Distal-010102b	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0878
Distal-010102c	T	C	T	T	T	C	G	G	G	G	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0168
Distal-010102d	T	C	C	T	T	C	A	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0014
Distal-010102e	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0000
Distal-010102f	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0004
Distal-0103a	C	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0446
Distal-0103b	C	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	C	T	A	0,0051
Distal-0104a	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,2003
Distal-0104b	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0000
Distal-0104c	T	C	T	T	T	C	G	G	A	T	T	T	C	C	A	C	T	C	T	C	C	T	C	0,0010
Distal-0104d	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0000
Distal-0104e	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0000

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 2A).

Posição chr6 hg19	29792987	29792989	29792990	29792991	29792996	29793021	29793062	29793100	29793105	29793160	29793162	29793176	29793203	29793224	29793243	29793276	29793309	29793310	29793324	29793361	29793362	29793365	29793385	29793404	29793453
Identificação do SNV	rs542428384	.	.	.	rs1632950	rs1736942	rs558600852	rs571891839	rs185842304	rs190164389	rs150426937	rs543301808	rs200585067	rs114251653	rs545198729	rs565008020	rs9278332	rs1736941	rs543960498	rs560849637	rs1736940	rs72838218	rs1736939	rs142798055	rs532106939
SNV	-2635	-2633	-2632	-2631	-2626	-2601	-2560	-2522	-2517	-2462	-2460	-2446	-2419	-2398	-2379	-2346	-2313	-2312	-2298	-2261	-2260	-2257	-2237	-2218	-2169
Alelo Referência	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Alelo Alternativo	C	T	C	G	T	G	T	T	T	C	G	T	T	T	T	T	T	G,*	T	C	A	A	A	C	A
Novo-1	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	C	A	A	G	C	C
Novo-2	T	C	T	A	T	G	G	C	C	T	A	A	T	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-3	T	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-4	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-5	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-6	T	C	T	A	T	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-7	T	C	T	A	C	A	G	C	C	T	G	A	T	T	G	C	TA	A	A	G	A	G	G	CTCT	C
Novo-8	T	C	T	A	T	A	G	C	C	T	A	A	T	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-9	T	C	T	A	C	A	G	C	C	T	A	A	T	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-10	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-11	T	C	T	A	C	A	G	C	C	C	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-12	T	C	T	A	T	G	G	C	C	T	A	T	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-13	T	C	T	A	C	A	G	C	C	T	A	A	T	C	G	C	TA	A	A	G	A	G	G	CTCT	C
Novo-14	T	C	T	A	T	G	G	T	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-15	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	A
Novo-16	T	C	T	A	C	A	G	C	C	T	A	A	T	C	G	C	TA	G	A	G	A	A	G	C	C
Novo-17	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	A	G	C	C
Novo-18	T	C	T	A	T	G	G	C	C	T	A	A	T	C	G	C	T	A	A	G	A	G	A	CTCT	C
Novo-19	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	G	G	G	CTCT	C

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 2B).

Posição chr6 hg19	29793460	29793473	29793503	29793542	29793577	29793581	29793590	29793608	29793610	29793630	29793650	29793724	29793761	29793769	29793785	29793876	29793942	29793947	29793982	29794008	29794049	29794097	29794146	Frequência dos haplótipos na População Mundial
Identificação do SNV	rs9461563	rs373926705	rs146027593	rs557634203	.	rs1233335	rs1736938	rs138910665	rs72492280	rs9258492	rs182479069	rs1632949	rs544293020	rs150381639	rs529879422	rs1736937	rs138915780	rs192306068	rs114712708	rs142134285	rs1632948	.	rs114252012	
SNV	-2162	-2149	-2119	-2080	-2045	-2041	-2032	-2014	-2012	-1992	-1972	-1898	-1861	-1853	-1837	-1746	-1680	-1675	-1640	-1614	-1573	-1525	-1476	
Alelo Referência	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	2n = 5070
Alelo Alternativo	C	G	C	A	C	G	G	A	A	G	T	C	T	G	T	A	G	T	C	T	C	C	A	
Novo-1	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0036
Novo-2	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0034
Novo-3	T	C	T	T	T	C	G	G	G	T	C	C	C	C	T	A	T	C	T	C	C	T	C	0,0022
Novo-4	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	C	C	T	T	A	0,0016
Novo-5	T	C	T	T	T	G	A	G	G	T	C	T	C	G	A	C	T	C	T	C	T	T	C	0,0016
Novo-6	T	C	C	T	T	C	G	G	G	T	C	C	C	C	A	A	T	T	T	C	C	T	C	0,0012
Novo-7	C	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0012
Novo-8	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0008
Novo-9	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0006
Novo-10	T	G	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0004
Novo-11	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0004
Novo-12	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0004
Novo-13	C	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0004
Novo-14	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0004
Novo-15	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0002
Novo-16	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0002
Novo-17	T	C	T	A	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0002
Novo-18	T	C	T	T	T	C	G	G	G	G	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0002
Novo-19	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0002

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 3A).

Posição chr6 hg19	29792987	29792989	29792990	29792991	29792996	29793021	29793062	29793100	29793105	29793160	29793162	29793176	29793203	29793224	29793243	29793276	29793309	29793310	29793324	29793361	29793362	29793365	29793385	29793404	29793453
Identificação do SNV	rs542428384	.	.	.	rs1632950	rs1736942	rs558600852	rs571891839	rs185842304	rs190164389	rs150426937	rs543301808	rs200585067	rs114251653	rs545198729	rs565008020	rs9278332	rs1736941	rs543960498	rs560849637	rs1736940	rs72838218	rs1736939	rs142798055	rs532106939
SNV	-2635	-2633	-2632	-2631	-2626	-2601	-2560	-2522	-2517	-2462	-2460	-2446	-2419	-2398	-2379	-2346	-2313	-2312	-2298	-2261	-2260	-2257	-2237	-2218	-2169
Alelo Referência	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Alelo Alternativo	C	T	C	G	T	G	T	T	T	C	G	T	T	T	T	T	T	G,*	T	C	A	A	A	C	A
Novo-20	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	T	TA	G	A	G	A	A	G	C	C
Novo-21	T	C	T	A	C	A	G	C	C	T	A	A	T	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-22	T	C	T	A	C	A	G	C	T	T	A	A	C	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-23	T	C	T	A	C	A	G	C	C	T	A	A	T	C	G	C	TA	A	A	G	G	G	G	CTCT	C
Novo-24	T	C	T	A	C	A	G	C	C	T	G	A	C	T	G	C	TA	A	T	G	A	G	G	CTCT	C
Novo-25	T	C	T	A	T	A	T	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-26	T	C	T	A	T	G	G	C	C	T	G	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-27	T	C	T	A	C	A	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	A	G	C	C
Novo-28	T	C	T	A	T	G	G	C	C	T	A	A	C	C	G	C	TA	G	A	G	A	G	A	CTCT	C
Novo-29	T	C	T	A	C	A	G	C	C	T	A	A	C	C	T	C	TA	G	A	G	A	A	G	C	C

Tabela 4.3.1 – Haplótipos da região Promotora-Distal do *HLA-G* e suas frequências na população mundial (parte 3B).

Posição chr6 hg19	29793460	29793473	29793503	29793542	29793577	29793581	29793590	29793608	29793610	29793630	29793650	29793724	29793761	29793769	29793785	29793876	29793942	29793947	29793982	29794008	29794049	29794097	29794146	Frequência dos haplótipos na População Mundial
Identificação do SNV	rs9461563	rs373926705	rs146027593	rs557634203	.	rs1233335	rs1736938	rs138910665	rs72492280	rs9258492	rs182479069	rs1632949	rs544293020	rs150381639	rs529879422	rs1736937	rs138915780	rs192306068	rs114712708	rs142134285	rs1632948	.	rs114252012	
SNV	-2162	-2149	-2119	-2080	-2045	-2041	-2032	-2014	-2012	-1992	-1972	-1898	-1861	-1853	-1837	-1746	-1680	-1675	-1640	-1614	-1573	-1525	-1476	
Alelo Referência	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	2n = 5070
Alelo Alternativo	C	G	C	A	C	G	G	A	A	G	T	C	T	G	T	A	G	T	C	T	C	C	A	
Novo-20	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0002
Novo-21	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0002
Novo-22	T	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	T	T	C	0,0002
Novo-23	T	C	T	T	T	G	A	G	G	T	C	T	C	C	A	C	T	C	T	T	T	T	C	0,0002
Novo-24	C	C	T	T	T	C	A	G	G	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0002
Novo-25	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0002
Novo-26	T	C	T	T	T	C	G	G	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0002
Novo-27	T	C	T	T	T	C	G	G	A	T	C	T	T	C	A	C	T	C	T	C	C	T	C	0,0002
Novo-28	T	C	T	T	T	C	G	A	G	T	C	C	C	C	A	A	T	C	T	C	C	T	C	0,0002
Novo-29	T	C	T	T	T	C	G	G	A	T	C	T	C	C	A	C	T	C	T	C	C	T	C	0,0002

Tabela 4.3.2 - Frequências dos haplótipos da região Promotora-Distal do *HLA-G* em diferentes populações do mundo (parte 1).

	Haplótipos 3'NTR <i>HLA-G</i> /Populações	Distal- 010101 ^a	Distal- 010101 ^b	Distal- 010101 ^c	Distal- 010101 ^d	Distal- 010101 ^e	Distal- 010101 ^f	Distal- 010102 ^a	Distal- 010102 ^b	Distal- 010102 ^c	Distal- 010102 ^d	Distal- 010102 ^e	Distal- 010102 ^f	Distal- 0103 ^a	Distal- 0103 ^b
Americanas	BSP	0,3522	0,0239	0,0104	0,0075	0,0015	0,0015	0,2866	0,0493	0,0239	0,0030	0,0015	0,0015	0,0761	0,0075
	CLM	0,3937	0,0479		0,0106			0,2181	0,0106	0,0106			0,0053	0,0532	0,0160
	MXL	0,3358	0,0448	0,0149	0,0075			0,2761	0,0522					0,0746	
	PEL	0,3838	0,0465					0,3314	0,0349	0,0174				0,0349	0,0116
	PUR	0,3904	0,0762		0,0048			0,1618	0,0524	0,0143	0,0048			0,0905	0,0143
Europeias	CEU	0,5151	0,0202	0,0152				0,2322	0,0707	0,0202				0,0202	0,0051
	FIN	0,6363	0,0101	0,0152				0,2020	0,0404	0,0253				0,0101	
	GBR	0,4403	0,0543	0,0054				0,2773	0,0979	0,0217	0,0054			0,0163	0,0054
	IBS	0,4393	0,0748					0,2804	0,0514	0,0093	0,0093			0,0234	0,0093
	TSI	0,4307	0,0231	0,0046				0,2732	0,0417	0,0278	0,0046		0,0046	0,0231	0,0046
Africanas	ACB	0,2605	0,0052		0,0729			0,2240		0,0104				0,1250	
	ASW	0,2726	0,0303	0,0076	0,0303			0,3029	0,0076	0,0076				0,0985	
	ESN	0,2928			0,0960			0,2171		0,0152				0,0808	
	GWD	0,2258	0,0354		0,0265			0,2655		0,0133	0,0044			0,1460	0,0044
	MSL	0,3353	0,0118		0,0529			0,1882						0,1353	
	YRI	0,2568	0,0092		0,0550			0,2523		0,0046				0,0780	
	LWK	0,3316	0,0544		0,0396			0,2722		0,0248				0,0792	0,0050
Sul-Asiáticas	BEB	0,2792	0,0116	0,0116				0,2093	0,2326	0,0058				0,0058	0,0058
	GIH	0,2123	0,0189					0,4197	0,1274	0,0330				0,0142	0,0094
	ITU	0,2573	0,0097	0,0049				0,3154	0,1165	0,0777				0,0291	0,0097
	PJL	0,2396		0,0052				0,3281	0,1615	0,0573				0,0104	0,0156
	STU	0,3009	0,0049					0,2135	0,1553	0,0146				0,0097	0,0049
Leste-Asiáticas	CHB	0,3349		0,0049				0,1601	0,1747	0,0146	0,0049			0,0146	0,0097
	CHS	0,4445						0,0509	0,2269						
	CDX	0,4646						0,0152	0,3586						
	KHV	0,3464	0,0792					0,1039	0,2030						
	JPT	0,2500						0,2019	0,0337					0,0144	
	Global	0,3492	0,0254	0,0032	0,0150			0,2287	0,0877	0,0168	0,0014		0,0004	0,0446	0,0051

Tabela 4.3.2 - Frequências dos haplótipos da região Promotora-Distal do *HLA-G* em diferentes populações do mundo (parte 2).

	Haplótipos 3'NTR <i>HLA-G</i> /Populações	Distal- 0104a	Distal- 0104b	Distal- 0104c	Distal- 0104d	Distal- 0104e	Novo- 1	Novo- 2	Novo- 3	Novo- 4	Novo- 5	Novo- 6	Novo- 7	Novo- 8	Novo- 9	Novo- 10	Novo- 11
Americanas	BSP	0,1433	0,0030	0,0030	0,0030	0,0015											
	CLM	0,2181					0,0106										
	MXL	0,1642					0,0224		0,0075								
	PEL	0,1221							0,0116	0,0058							
	PUR	0,1762					0,0143										
Europeias	CEU	0,0909					0,0051								0,0051		
	FIN	0,0606															
	GBR	0,0652					0,0054			0,0054							
	IBS	0,0981															
	TSI	0,1389					0,0046		0,0046						0,0093		
Africanas	ACB	0,2656						0,0052	0,0104	0,0052		0,0052					
	ASW	0,1666		0,0152			0,0152		0,0152	0,0152		0,0152					
	ESN	0,2474		0,0101				0,0253						0,0051			0,0051
	GWD	0,2655					0,0044				0,0044						
	MSL	0,2059					0,0176	0,0118	0,0059		0,0118	0,0176					
	YRI	0,2982		0,0046				0,0183	0,0046		0,0046			0,0046			
	LWK	0,1336						0,0248	0,0050	0,0149							
Sul-Asiáticas	BEB	0,2209									0,0058		0,0058			0,0058	
	GIH	0,1604															
	ITU	0,1748					0,0049										
	PJL	0,1667											0,0052				
	STU	0,2767											0,0097			0,0049	
Leste-Asiáticas	CHB	0,2767									0,0049						
	CHS	0,2685											0,0046				
	CDX	0,1616															
	KHV	0,2426									0,0099		0,0050				0,0050
	JPT	0,4808												0,0096			
	Global	0,2003		0,0010			0,0036	0,0034	0,0022	0,0016	0,0016	0,0012	0,0012	0,0008	0,0006	0,0004	0,0004

Tabela 4.3.2 - Frequências dos haplótipos da região Promotora-Distal do *HLA-G* em diferentes populações do mundo (parte 4).

	Haplótipos 3'NTR <i>HLA-G</i> /Populações	Novo- 27	Novo- 28	Novo- 29	2n
Americanas	BSP				670
	CLM				188
	MXL				134
	PEL				172
	PUR				210
Europeias	CEU				198
	FIN				198
	GBR				184
	IBS				214
	TSI		0,0046		216
Africanas	ACB				192
	ASW				132
	ESN				198
	GWD				226
	MSL				170
	YRI			0,0046	218
	LWK				202
Sul-Asiáticas	BEB				172
	GIH				212
	ITU				206
	PJL				192
	STU	0,0049			206
Leste-Asiáticas	CHB				206
	CHS				216
	CDX				198
	KHV				202
	JPT				208
	Global	0,0002	0,0002	0,0002	5070

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJL (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.1 - Distribuição dos haplótipos da Região Promotora-Distal com frequências maiores que 1% nas populações avaliadas.

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJI (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.2 - Rede *Median-Joining* dos haplótipos da região Promotora-Distal do *HLA-G*.

As áreas dos círculos são proporcionais à frequência global dos haplótipos e os comprimentos dos ramos ao número de mutações que os separaram. Os círculos são coloridos de acordo com as frequências das regiões das populações (marrom – americanas, cinza escuro – africanas, verde médio – europeias, verde claro – sul-asiáticas e verde-escuro – leste-asiáticas).

Figura 4.3.3 - Dispersão em Escala Multidimensional considerando a Região Promotora-Distal do *HLA-G*.

Nesta figura, estão apresentados os gráficos obtidos após a análise da Dispersão em Escala Multidimensional (MDS) considerando os valores das distâncias genéticas (F_{st} de Reynolds) obtidos a partir dos dados da Região Promotora-Distal do *HLA-G* considerando todas as populações com dados disponíveis para essa região (A) e somente nas populações do Projeto *1000 Genomes* (B). As populações avaliadas foram: BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJL (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã) e JPT (Japoneses residentes em Tokyo no Japão).

4.3.2 Região Promotora-Proximal

Todos os haplótipos identificados nas análises dos dados populacionais disponíveis no Projeto *1000 Genomes* e descritos em publicações prévias estão listados na Tabela 4.3.3. Um total de 58 SNVs que compõem 55 diferentes haplótipos foram identificados para a região promotora-proximal. Desses, somente 11 apresentaram frequências maiores que 1% na população mundial avaliada e 30 foram identificados pela primeira vez neste trabalho. As frequências dos haplótipos em cada uma das populações avaliadas estão descritas na Tabela 4.3.4 e os haplótipos com frequências maiores que 1% estão representados na Figura 4.3.4. A partir dos dados de frequência e as variações nucleotídicas dos haplótipos, foi também criada uma rede de haplótipos *Median-Joining* que está representada na Figura 4.3.5. Nessa figura é possível avaliar, graficamente, os haplótipos mais frequentes (010101a, 010102a e 0104a) e as suas variações de frequência nos grupos populacionais de diferentes origens.

Nas análises populacionais, foi verificado que a diversidade haplotípica da região promotora-proximal variou entre 0,8633 (MSL) a 0,6187 (GIH) quando todas as populações foram analisadas conjuntamente. Em cada um dos cinco grupos, os valores máximos e mínimos desse parâmetro foram os seguintes: populações americanas (AMR) 0,8621 (PUR) a 0,7485 (PEL); populações europeias (EUR) 0,7813 (TSI) a 0,7341 (CEU); populações africanas (AFR) 0,8633 (MSL) a 0,8189 (YRI); populações sul-asiáticas (SAS) 0,7274 (STU) a 0,6187 (GIH); e populações leste-asiáticas (EAS) 0,7477 (CHB) a 0,6294 (CDX).

No teste de neutralidade genético-populacional de Ewens-Watterson foram avaliados os valores F esperados e observados e a significância de Slatkin. Os valores de F observados para esta região estão dentro do intervalo de confiança, sendo que alguns dos valores observados são mais altos e outros mais baixos do que os valores esperados, e somente em uma das populações avaliadas (BSP) foi verificado diferença significativa entre os valores ($p = 0,019$).

Os valores de distância genética (F_{ST} de Reynolds) foram calculados nas populações aos pares e os dados foram utilizados para realizar a análise de MDS. O resultado obtido desta análise está apresentado na Figura 4.3.6, através de dois gráficos, o primeiro (A) contém as informações de todas as populações avaliadas tendo por base a região Promotora e o segundo (B) somente os dados das populações do Projeto *1000 Genomes*. Os dados mostram que todas

as populações apresentam proximidade genética, com exceção da população BSP que aparece um pouco mais distanciada das demais.

No teste de diferenciação entre todos os pares de populações, foram observadas diferenças significativas em 332 das 351 comparações realizadas. Por fim, na análise de AMOVA, quando todas as populações avaliadas e considerando os cinco grupos populacionais, foi observado que 92,62% da variabilidade genética está dentro das populações, 4,55% entre grupos, 2,83% entre populações dentro dos grupos e valor de F_{ST} de 7,3%. Quando apenas as populações dos cinco grupos do Projeto *1000 Genomes* foram avaliadas, os valores de variabilidade observados foram: 95,56% da variabilidade genética está dentro das populações, 3,15% entre grupos, 1,29% entre populações dentro dos grupos e valor de F_{ST} de 4,4%.

As tabelas contendo os dados completos das análises populacionais para a região promotora-proximal estão apresentadas no Apêndice B.

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 1A).

Posição chr6 hg19	29794216	29794245	29794312	29794317	29794336	29794443	29794467	29794482	29794484	29794488	29794501	29794524	29794549	29794632	29794658	29794700	29794756	29794812	29794843	29794860	29794863	29794867	29794897	29794906	29794933	29794956	29794976	29794989	29795076	29795081	29795082
Identificação do SNV	rs527969449	rs113076482	rs112835540	rs1736936	rs558572524	rs1736935	rs3823321	rs1736934	rs17875389	rs540383368	rs3115630	rs146374870	rs192232623	rs562500258	rs16332947	rs370338057	rs561785661	.	rs547123392	rs16332946	rs538982928	rs115528044	rs1233334	rs2249863	rs2735022	rs35674592	rs17875391	rs16332944	rs201221694	rs368205133	.
SNV	-1406	-1377	-1310	-1305	-1286	-1179	-1155	-1140	-1138	-1125	-1121	-1098	-1073	-990	-964	-922	-866	-810	-779	-762	-759	-755	-725	-716	-689	-666	-646	-633	-546	-541	-540
Alelo Referência	AGT	T	C	G	G	A	G	A	A	C	T	G	G	G	G	C	A	C	C	C	T	G	G	T	A	G	A	G	A	GA	A
Alelo Alternativo	A	G	T	A	T	G	A	T	G	T	C	A	T	A	A	A	T	T	T	T	C	A	C,T	G	G	T	G	A	AG	G	G,*
010101a	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101b	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	G	T	A	G	A	G	A	GA	A
010101c	AGT	T	C	G	G	A	G	A	A	C	T	G	G	G	G	C	A	C	C	C	T	G	G	T	A	G	A	G	A	GA	A
010101d	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101f	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	G	*
010101g	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101h	AGT	G	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101i	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101j	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
010101k	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	T	C	T	G	C	T	A	G	A	G	A	GA	A
010102a	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
010102b	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	T	C	T	T	G	C	G	G	T	A	A	A	GA	A
010102c	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	A	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
010102d	AGT	T	C	A	G	G	G	T	A	C	C	A	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
010102e	AGT	T	C	G	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
0103a	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	A	G	AG	GA	A
0103c	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	G	A	GA	A
0103d	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	A	G	AG	GA	A
0103e	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	G	A	GA	A
0103f	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	G	A	GA	A
0103g	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	A	G	AG	GA	G
0104a	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 1B).

Posição chr6 hg19	29795083	29795101	29795113	29795136	29795139	29795145	29795179	29795210	29795222	29795223	29795231	29795253	29795267	29795325	29795338	29795366	29795367	29795369	29795403	29795412	29795421	29795438	29795472	29795477	29795549	29795566	29795621	Frequência dos haplótipos na População Mundial
Identificação do SNV	rs546145666	rs138987412	rs17875393	rs1736933	rs149890776	rs1736932	rs17875394	rs560767298	rs17875395	.	rs17875396	rs1632943	rs191630481	.	rs548611139	rs568354681	rs551170211	rs534338234	rs553952142	rs570647051	rs1233333	rs556503999	rs576355035	rs542140276	rs555381090	rs17875397	rs148951074	
SNV	-539	-521	-509	-486	-483	-477	-443	-412	-400	-399	-391	-369	-355	-297	-284	-256	-255	-253	-219	-210	-201	-184	-150	-145	-73	-56	-1	
Alelo Referência	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	
Alelo Alternativo	G	A	G	C	G	G	A	C	A	A	A	A	A	A	A	A	*	T	T	C	A	C	T	G	G	T	T	2n = 5070
010101a	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,2511
010101b	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0468
010101c	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0258
010101d	A	C	C	A	G	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0270
010101f	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0000
010101g	A	C	C	A	A	C	G	G	G	G	G	C	G	A	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0000
010101h	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0049
010101i	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0000
010101j	A	C	C	A	A	C	G	G	G	A	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0000
010101k	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0000
010102a	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,3304
010102b	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002
010102c	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0014
010102d	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0028
010102e	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0014
0103a	A	C	G	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0140
0103c	A	C	G	A	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0055
0103d	A	C	C	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0101
0103e	A	C	G	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0213
0103f	A	C	G	A	A	G	G	G	A	G	A	A	A	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0020
0103g	A	C	C	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0000
0104a	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,1707

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 2A).

Posição chr6 hg19	29794216	29794245	29794312	29794317	29794336	29794443	29794467	29794482	29794484	29794488	29794501	29794524	29794549	29794632	29794658	29794700	29794756	29794812	29794843	29794860	29794863	29794867	29794897	29794906	29794933	29794956	29794976	29794989	29795076	29795081	29795082
Identificação do SNV	rs527969449	rs113076482	rs112835540	rs1736936	rs558572524	rs1736935	rs3823321	rs1736934	rs17875389	rs540383368	rs3115630	rs146374870	rs192232623	rs562500258	rs1632947	rs370338057	rs561785661	.	rs547123392	rs1632946	rs538982928	rs115528044	rs1233334	rs2249863	rs2735022	rs35674592	rs17875391	rs1632944	rs201221694	rs368205133	.
SNV	-1406	-1377	-1310	-1305	-1286	-1179	-1155	-1140	-1138	-1125	-1121	-1098	-1073	-990	-964	-922	-866	-810	-779	-762	-759	-755	-725	-716	-689	-666	-646	-633	-546	-541	-540
Alelo Referência	AGT	T	C	G	G	A	G	A	A	C	T	G	G	G	G	C	A	C	C	C	T	G	G	T	A	G	A	G	A	GA	A
Alelo Alternativo	A	G	T	A	T	G	A	T	G	T	C	A	T	A	A	A	T	T	T	T	C	A	C,T	G	G	T	G	A	AG	G	G,*
0104b	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
0104c	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
0104d	AGT	T	C	A	G	G	A	A	A	C	C	G	T	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-1	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	G	A
Novo-2	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-3	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-4	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-5	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	G	G	T	A	A	A	GA	A
Novo-6	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-7	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	A	C	G	G	T	A	A	A	GA	A
Novo-8	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-9	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	C	G	C	G	G	T	A	A	A	GA	A
Novo-10	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	G	A	GA	A
Novo-11	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-12	A	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-13	AGT	T	C	A	G	G	G	T	A	C	T	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-14	AGT	T	T	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-15	AGT	T	C	A	G	A	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-16	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-17	AGT	T	C	G	G	G	G	A	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	G	A	GA	A
Novo-18	AGT	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-19	AGT	T	C	A	G	G	G	T	A	C	C	G	T	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 2B).

Posição chr6 hg19	rs29795083 rs29795101 rs29795113 rs29795136 rs29795139 rs29795145 rs29795179 rs29795210 rs29795222 rs29795223 rs29795231 rs29795253 rs29795267 rs29795325 rs29795338 rs29795366 rs29795367 rs29795369 rs29795403 rs29795412 rs29795421 rs29795438 rs29795472 rs29795477 rs29795549 rs29795566 rs29795621																												
Identificação do SNV	rs546145666 rs138987412 rs17875393 rs1736933 rs149890776 rs1736932 rs17875394 rs560767298 rs17875395 . rs17875396 rs1632943 rs191630481 . rs548611139 rs568354681 rs551170211 rs534338234 rs553952142 rs570647051 rs1233333 rs556503999 rs576355035 rs542140276 rs555381090 rs17875397 rs148951074																												Frequência dos haplótipos na População Mundial
SNV	-539 -521 -509 -486 -483 -477 -443 -412 -400 -399 -391 -369 -355 -297 -284 -256 -255 -253 -219 -210 -201 -184 -150 -145 -73 -56 -1																												
Alelo Referência	A C C A A C G G G G G C G G G T C C C G G T C C A C G																												2n = 5070
Alelo Alternativo	G A G C G G A C A A A A A * T T C A C T G G T T																												
0104b	A	C	C	C	A	G	A	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0268	
0104c	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	*	C	C	G	A	T	C	C	A	C	G	0,0000	
0104d	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	*	C	C	G	A	T	C	C	A	C	G	0,0000	
Novo-1	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0432	
Novo-2	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	A	*	C	C	G	A	T	C	C	A	C	G	0,0032	
Novo-3	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	G	A	C	G	0,0012	
Novo-4	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	T	C	G	A	T	C	C	A	C	G	0,0002	
Novo-5	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0012	
Novo-6	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	C	C	G	T	C	C	A	C	G	0,0002
Novo-7	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	
Novo-8	A	A	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0006	
Novo-9	A	C	C	C	A	G	A	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	
Novo-10	A	C	G	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	T	0,0002	
Novo-11	A	C	C	C	A	G	G	G	G	G	G	A	G	G	A	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	
Novo-12	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0010	
Novo-13	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	
Novo-14	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0008	
Novo-15	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	
Novo-16	G	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0014	
Novo-17	A	C	G	A	A	G	G	C	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0002	
Novo-18	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	G	C	G	0,0002	
Novo-19	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002	

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 3A).

Posição chr6 hg19	29794216	29794245	29794312	29794317	29794336	29794443	29794467	29794482	29794484	29794488	29794501	29794524	29794549	29794632	29794658	29794700	29794756	29794812	29794843	29794860	29794863	29794867	29794897	29794906	29794933	29794956	29794976	29794989	29795076	29795081	29795082
Identificação do SNV	rs527969449	rs113076482	rs112835540	rs1736936	rs558572524	rs1736935	rs3823321	rs1736934	rs17875389	rs540383368	rs3115630	rs146374870	rs192232623	rs562500258	rs1632947	rs370338057	rs561785661	.	rs547123392	rs1632946	rs538982928	rs115528044	rs1233334	rs2249863	rs2735022	rs35674592	rs17875391	rs1632944	rs201221694	rs368205133	.
SNV	-1406	-1377	-1310	-1305	-1286	-1179	-1155	-1140	-1138	-1125	-1121	-1098	-1073	-990	-964	-922	-866	-810	-779	-762	-759	-755	-725	-716	-689	-666	-646	-633	-546	-541	-540
Alelo Referência	AGT	T	C	G	G	A	G	A	A	C	T	G	G	G	G	C	A	C	C	C	T	G	G	T	A	G	A	G	A	GA	A
Alelo Alternativo	A	G	T	A	T	G	A	T	G	T	C	A	T	A	A	A	T	T	T	T	C	A	C,T	G	G	T	G	A	AG	G	G,*
Novo-20	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	T	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-21	AGT	T	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-22	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-23	AGT	T	C	A	G	G	G	T	A	C	C	G	G	A	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-24	AGT	G	C	G	G	A	G	A	A	C	C	G	G	G	G	C	A	C	C	C	T	G	C	T	A	G	A	G	A	GA	A
Novo-25	AGT	T	C	G	G	G	G	A	G	T	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	A	G	AG	GA	A
Novo-26	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-27	AGT	T	C	A	T	G	A	A	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-28	AGT	T	C	G	G	G	G	T	G	C	C	G	G	G	G	C	A	C	C	C	T	G	T	T	A	G	G	A	A	GA	A
Novo-29	AGT	T	C	A	G	G	A	A	A	C	C	G	G	G	A	C	T	C	C	T	T	G	C	G	G	T	A	A	A	GA	A
Novo-30	A	T	C	A	G	G	G	T	A	C	C	G	G	G	A	C	A	C	C	T	T	G	C	G	G	T	A	A	A	GA	A

Tabela 4.3.3 – Haplótipos da região Promotora-Proximal do *HLA-G* e suas frequências na população mundial (parte 3B).

Posição chr6 hg19	29795083	29795101	29795113	29795136	29795139	29795145	29795179	29795210	29795222	29795223	29795231	29795253	29795267	29795325	29795338	29795366	29795367	29795369	29795403	29795412	29795421	29795438	29795472	29795477	29795549	29795566	29795621	
Identificação do SNV	rs546145666	rs138987412	rs17875393	rs1736933	rs149890776	rs1736932	rs17875394	rs560767298	rs17875395	.	rs17875396	rs1632943	rs191630481	.	rs548611139	rs568354681	rs551170211	rs534338234	rs553952142	rs570647051	rs1233333	rs556503999	rs576355035	rs542140276	rs555381090	rs17875397	rs148951074	Frequência dos haplótipos na População Mundial
SNV	-539	-521	-509	-486	-483	-477	-443	-412	-400	-399	-391	-369	-355	-297	-284	-256	-255	-253	-219	-210	-201	-184	-150	-145	-73	-56	-1	
Alelo Referência	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	2n = 5070
Alelo Alternativo	G	A	G	C	G	G	A	C	A	A	A	A	A	A	A	A	*	T	T	C	A	C	T	G	G	T	T	
Novo-20	A	C	C	A	G	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0002
Novo-21	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	G	T	C	C	A	C	G	0,0004
Novo-22	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	C	C	C	A	C	G	0,0002
Novo-23	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002
Novo-24	A	C	C	A	A	C	G	G	G	G	G	C	G	G	G	T	C	C	C	G	G	T	T	C	A	C	G	0,0004
Novo-25	A	C	C	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0002
Novo-26	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	T	G	A	T	C	C	A	C	G	0,0002
Novo-27	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0004
Novo-28	A	C	G	A	A	G	G	G	A	G	A	A	G	G	G	T	C	C	C	G	G	T	C	C	A	T	G	0,0002
Novo-29	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0002
Novo-30	A	C	C	C	A	G	G	G	G	G	G	A	G	G	G	T	C	C	C	G	A	T	C	C	A	C	G	0,0004

Frequência dos
haplótipos na
População
Mundial

2n = 5070

	Haplótipos Promotora <i>HLA-G</i> /Populações	010101a	010101b	010101c	010101d	010101h	010102a	010102b	010102c	010102d	010102e	0103a	0103c	0103d	0103e
Americanas	BSP	0,1980	0,0640	0,0570	0,0160	0,0540				0,3280			0,0030	0,0410	0,0060
	CLM	0,2181	0,0479	0,0851	0,0372	0,0053	0,2394	0,0053				0,0319	0,0053	0,0160	0,0320
	MXL	0,2239	0,0672	0,0149	0,0373	0,0075	0,2985			0,0373		0,0821		0,0075	0,0075
	PEL	0,3450	0,0175	0,0058	0,0117		0,3509			0,0351		0,0234		0,0117	0,0117
	PUR	0,2333	0,0477	0,0667	0,0429		0,2095			0,0143	0,0048	0,0429		0,0286	0,0429
Europeias	CEU	0,3838	0,0404	0,1010	0,0051		0,3182					0,0202			0,0051
	FIN	0,3586	0,2071	0,0758			0,2576					0,0101			
	GBR	0,3098	0,0272	0,0978	0,0109		0,3859				0,0054	0,0109		0,0054	0,0054
	IBS	0,2944	0,0234	0,0607	0,0467		0,3364				0,0093	0,0140		0,0093	0,0093
	TSI	0,2824	0,0463	0,0926	0,0139		0,3333		0,0139		0,0046	0,0139		0,0139	0,0046
Africanas	ACB	0,1198	0,0469	0,0052	0,0573	0,0365	0,2448		0,0052			0,0104	0,0156	0,0417	0,0521
	ASW	0,1515	0,0379	0,0076	0,0758	0,0076	0,3182		0,0152			0,0303	0,0455	0,0076	0,0303
	ESN	0,1617	0,0808		0,0556	0,0202	0,2374						0,0404	0,0051	0,0303
	GWD	0,1327	0,0310		0,0664		0,2788				0,0044		0,0088	0,0531	0,0840
	MSL	0,1706	0,0647		0,1294		0,1941						0,0118	0,0471	0,0882
	YRI	0,1238	0,0872		0,0459	0,0229	0,2615						0,0275	0,0138	0,0367
	LWK	0,1931	0,0396	0,0050	0,0891	0,0297	0,3020								0,0595

Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do *HLA-G* em diferentes populações do mundo (parte 1B).

	Haplótipos Promotora <i>HLA-G</i> / Populações	010101a	010101b	010101c	010101d	010101h	010102a	010102b	010102c	010102d	010102e	0103a	0103c	0103d	0103e
Sul-Asiáticas	BEB	0,2558	0,0407				0,4535					0,0058			0,0058
	GIH	0,1462	0,0613	0,0047			0,5755					0,0142			0,0094
	ITU	0,2282	0,0194	0,0097	0,0049		0,5097					0,0340			0,0097
	PJL	0,1990	0,0366	0,0105			0,5497		0,0052			0,0105			0,0157
	STU	0,2039	0,0825	0,0097			0,3883					0,0097			0,0049
Leste-Asiáticas	CHB	0,3010	0,0340	0,0049			0,3495				0,0049	0,0146			0,0097
	CHS	0,4213	0,0093	0,0046			0,2824								
	CDX	0,4646					0,3737								
	KHV	0,3416	0,0198				0,3020								
	JPT	0,2404	0,0048				0,2356					0,0144			
	Global	0,2512	0,0468	0,0258	0,0270	0,0049	0,3305	0,0002	0,0014	0,0028	0,0014	0,0140	0,0055	0,0101	0,0213

Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do *HLA-G* em diferentes populações do mundo (parte 2B).

	Haplótipos Promotora <i>HLA-G</i> /Populações	0103f	0104a	0104b	Novo- 01	Novo- 02	Novo- 03	Novo- 04	Novo- 05	Novo- 06	Novo- 07	Novo- 08	Novo- 09	Novo- 10	Novo- 11
Sul-Asiáticas	BEB	0,2035	0,0116	0,0116				0,0058	0,0058						
	GIH	0,1651		0,0189											
	ITU	0,1699		0,0097											
	PJL	0,1518							0,0052						
	STU	0,2767		0,0097					0,0097						
Leste-Asiáticas	CHB	0,1796	0,0971									0,0049			
	CHS	0,1667	0,0972	0,0046					0,0046			0,0046	0,0046		
	CDX	0,1313	0,0303												
	KHV	0,2129	0,0297	0,0842					0,0050						
	JPT	0,3654	0,1346									0,0048			
	Global	0,0020	0,1707	0,0268	0,0432	0,0032	0,0012	0,0002	0,0012	0,0002	0,0002	0,0006	0,0002	0,0002	0,0002

	Haplótipos Promotora HLA-G/Populações	Novo-12	Novo-13	Novo-14	Novo-15	Novo-16	Novo-17	Novo-18	Novo-19	Novo-20	Novo-21	Novo-22	Novo-23	Novo-24	Novo-25
Americanas	BSP														
	CLM														
	MXL														
	PEL	0,0058		0,0058											
	PUR			0,0048											
Europeias	CEU														
	FIN	0,0101	0,0051												
	GBR			0,0109											
	IBS	0,0047						0,0047	0,0047	0,0047					
	TSI	0,0046									0,0046				
Africanas	ACB														
	ASW														
	ESN														
	GWD					0,0044	0,0044	0,0044							
	MSL													0,0059	
	YRI					0,0046									
	LWK					0,0248									

Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do *HLA-G* em diferentes populações do mundo (parte 3B).

	Haplótipos Promotora <i>HLA-G</i> / Populações	Novo- 12	Novo- 13	Novo- 14	Novo- 15	Novo- 16	Novo- 17	Novo- 18	Novo- 19	Novo- 20	Novo- 21	Novo- 22	Novo- 23	Novo- 24	Novo- 25
Sul-Asiáticas	BEB														
	GIH	0,0047													
	ITU	0,0049													
	PJL														
	STU														
Leste-Asiáticas	CHB														
	CHS														
	CDX														
	KHV	0,0050													
	JPT														
	Global	0,0010	0,0002	0,0008	0,0002	0,0014	0,0002	0,0002	0,0002	0,0002	0,0004	0,0002	0,0002	0,0004	0,0002

Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do *HLA-G* em diferentes populações do mundo (parte 4A).

	Haplótipos Promotora <i>HLA-G</i> / Populações	Novo- 26	Novo- 27	Novo- 28	Novo- 29	Novo- 30	n
Americanas	BSP						628
	CLM						188
	MXL						134
	PEL						171
	PUR			0,0048			210
Europeias	CEU						198
	FIN						198
	GBR						184
	IBS						214
	TSI					0,0093	216
Africanas	ACB						192
	ASW						132
	ESN						198
	GWD						226
	MSL						170
	YRI						218
	LWK						202

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJI (residentes de Lahore em Punjab no Paquistão), STU (srilanqueses Tamil residentes no

Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Tabela 4.3.4 - Frequências dos haplótipos da região Promotora-Proximal do *HLA-G* em diferentes populações do mundo (parte 4B).

	Haplótipos Promotora <i>HLA-G</i> / Populações	Novo- 26	Novo- 27	Novo- 28	Novo- 29	Novo- 30	n
Sul-Asiáticas	BEB						172
	GIH						212
	ITU						206
	PJL	0,0052	0,0105				191
	STU				0,0049		206
Leste-Asiáticas	CHB						206
	CHS						216
	CDX						198
	KHV						202
	JPT						208
	Global	0,0002	0,0004	0,0002	0,0002	0,0004	5068

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJL (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.4 - Distribuição dos haplótipos da Região Promotora-Proximal com frequências maiores que 1% nas populações avaliadas.

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJI (residentes de Lahore em Punjabi no Paquistão), STU (srilankeses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.5 - Rede *Median-Joining* dos haplótipos da região Promotora-Proximal do *HLA-G*.

As áreas dos círculos são proporcionais à frequência global dos haplótipos e os comprimentos dos ramos ao número de mutações que os separaram. Os círculos são coloridos de acordo com as frequências das regiões das populações (marrom – americanas, cinza escuro – africanas, verde médio – europeias, verde claro – sul-asiáticas e verde-escuro – leste-asiáticas).

Figura 4.3.6 - Dispersão em Escala Multidimensional considerando a Região Promotora-Proximal do *HLA-G*.

Nesta figura, estão apresentados os gráficos obtidos após a análise da Dispersão em Escala Multidimensional (MDS) considerando os valores das distâncias genéticas (F_{st} de Reynolds) obtidos a partir dos dados da Região Promotora-Proximal do *HLA-G* considerando todas as populações com dados disponíveis para essa região (A) e somente nas populações do Projeto *1000 Genomes* (B).

4.3.3 Região *Enhancer L*

Os haplótipos identificados nas análises dos dados populacionais disponíveis no Projeto *1000 Genomes* estão listados na Tabela 4.3.5. Cinco SNV foram identificados, compondo 6 diferentes haplótipos para a região *Enhancer L*. Todos os haplótipos foram identificados pela primeira vez no presente trabalho e somente 2 apresentaram frequências maiores que 1% na população mundial avaliada. As frequências dos haplótipos em cada uma das populações avaliadas estão descritas na Tabela 4.3.6 e as frequências dos haplótipos nas populações avaliadas estão representadas na Figura 4.3.7. As frequências e as variações nucleotídicas dos haplótipos foram ainda usadas para criar uma rede de haplótipos *Median-Joining* que está na Figura 4.3.8. Na figura é possível observar os haplótipos mais frequentes da região (*EnhancerL-01* e *EnhancerL-02*) e as suas variações dentro das populações.

Nas análises populacionais, foi verificado que a diversidade haplotípica da região 3'NT variou entre 0,4894 (GIH) a 0,0300 (CDX) quando todas as populações foram analisadas conjuntamente. Em cada um dos cinco grupos, os valores máximos e mínimos desse parâmetro foram os seguintes: populações americanas (AMR) 0,4534 (PEL) a 0,2727 (PUR); populações europeias (EUR) 0,4112 (TSI) a 0,3241 (FIN); populações africanas (AFR) 0,4372 (ASW) a 0,3147 (MSL); populações sul-asiáticas (SAS) 0,4894 (GIH) a 0,3329 (BEB); e populações leste-asiáticas (EAS) 0,3391 (JPT) a 0,0300 (CDX).

No teste de neutralidade genético-populacional de Ewens-Watterson foram avaliados os valores *F* esperados e observados e a significância de Slatkin. Todos os valores de *F* observados são menores do que os valores esperados, com exceção das populações CHS e CDX, nas quais os valores observados foram maiores do que os esperados. Valores observados menores que valores esperados foram verificados na maioria dos casos. O cálculo da significância de Slatkin não apresentou valores significativos para nenhuma das populações.

Os valores de distância genética (F_{ST} de Reynolds) foram calculados nas populações aos pares e os dados foram utilizados para realizar a análise de MDS. O resultado obtido desta análise está apresentado na Figura 4.3.9. No gráfico estão representadas as populações do Projeto *1000 Genomes*, as populações CDX, CHS, KHV e GIH apresentaram as maiores distâncias genéticas.

No teste de diferenciação entre todos os pares de populações foram observadas diferenças significativas em 147 das 325 comparações realizadas. Por fim, na AMOVA, quando as populações dos cinco grupos do Projeto *1000 Genomes* avaliadas e considerando os cinco grupos populacionais, foi observado que 95,6% da variabilidade genética está dentro das populações, 2,58% entre grupos, 1,82% entre populações dentro dos grupos e valor médio de F_{ST} de 4,4%.

As tabelas contendo os dados completos das análises populacionais para a região *Enhancer L* estão apresentadas no Apêndice C.

Tabela 4.3.5 – Haplótipos da região *Enhancer L* e suas frequências na população mundial.

Posição chr6 hg19	29783064	29783078	29783083	29783125	29783142	Frequência dos haplótipos na População Mundial
Identificação do SNP	rs1077433	rs550851704	rs566038094	rs117884929	rs551878526	
Alelo Referência	G	T	G	G	T	
Alelo Alternativo	A	A	T	A	C	2n = 5070
EnhancerL-01	G	T	G	G	T	0,7665
EnhancerL-02	A	T	G	G	T	0,2323
EnhancerL-03	G	T	T	G	T	0,0002
EnhancerL-04	A	T	G	G	C	0,0002
EnhancerL-05	G	A	G	G	T	0,0004
EnhancerL-06	G	T	G	A	T	0,0004

Tabela 4.3.6 - Frequências dos haplótipos da região *Enhancer L* em diferentes populações do mundo.

Haplótipos <i>Enhancer L</i> /Populações		EnhancerL-01	EnhancerL-02	EnhancerL-03	EnhancerL-04	EnhancerL-05	EnhancerL-06	2n
Americanas	CLM	0,7770	0,2230	0,0000	0,0000	0,0000	0,0000	188
	MXL	0,7160	0,2840	0,0000	0,0000	0,0000	0,0000	134
	PEL	0,6570	0,3430	0,0000	0,0000	0,0000	0,0000	172
	PUR	0,8380	0,1620	0,0000	0,0000	0,0000	0,0000	210
Europeias	CEU	0,7680	0,2270	0,0000	0,0050	0,0000	0,0000	198
	FIN	0,7980	0,2020	0,0000	0,0000	0,0000	0,0000	198
	GBR	0,7230	0,2770	0,0000	0,0000	0,0000	0,0000	184
	IBS	0,7200	0,2800	0,0000	0,0000	0,0000	0,0000	214
	TSI	0,7130	0,2870	0,0000	0,0000	0,0000	0,0000	216
Africanas	ACB	0,7550	0,2400	0,0050	0,0000	0,0000	0,0000	192
	ASW	0,6820	0,3180	0,0000	0,0000	0,0000	0,0000	132
	ESN	0,7780	0,2220	0,0000	0,0000	0,0000	0,0000	198
	GWD	0,7350	0,2650	0,0000	0,0000	0,0000	0,0000	226
	MSL	0,8060	0,1940	0,0000	0,0000	0,0000	0,0000	170
	YRI	0,7430	0,2570	0,0000	0,0000	0,0000	0,0000	218
	LWK	0,7230	0,2770	0,0000	0,0000	0,0000	0,0000	202
Sul-Asiáticas	BEB	0,7910	0,2090	0,0000	0,0000	0,0000	0,0000	172
	GIH	0,5800	0,4200	0,0000	0,0000	0,0000	0,0000	212
	ITU	0,6750	0,3160	0,0000	0,0000	0,0090	0,0000	206
	PJL	0,6670	0,3330	0,0000	0,0000	0,0000	0,0000	192
	STU	0,7820	0,2180	0,0000	0,0000	0,0000	0,0000	206
Les	CHB	0,8400	0,1600	0,0000	0,0000	0,0000	0,0000	206

CHS	0,9440	0,0560	0,0000	0,0000	0,0000	0,0000	216
CDX	0,9848	0,0152	0,0000	0,0000	0,0000	0,0000	198
KHV	0,8960	0,1040	0,0000	0,0000	0,0000	0,0000	202
JPT	0,7880	0,2020	0,0000	0,0000	0,0000	0,0100	208
Global	0,7665	0,2323	0,0002	0,0002	0,0004	0,0004	5070

Populações Americanas (AMR): BSP (residentes no estado de São Paulo no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJI (residentes de Lahore em Punjab no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.7 - Distribuição dos haplótipos da Região *Enhancer L* nas populações avaliadas.

Populações Americanas (AMR): CLM (colombianos residentes em Medellín na Colômbia), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJL (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.8 - Rede *Median-Joining* dos haplótipos da região *Enhancer L*.

As áreas dos círculos são proporcionais à frequência global dos haplótipos e os comprimentos dos ramos ao número de mutações que os separaram. Os círculos são coloridos de acordo com as frequências das regiões das populações (marrom – americanas, cinza escuro – africanas, verde médio – europeias, verde claro – sul-asiáticas e verde-escuro – leste-asiáticas).

Figura 4.3.9 - Dispersão em Escala Multidimensional considerando a Região *Enhancer L*.

Nesta figura, está apresentado o gráfico obtido após a análise da Dispersão em Escala Multidimensional (MDS) considerando os valores das distâncias genéticas (Fst de Reynolds) obtidos a partir dos dados da Região *Enhancer L* considerando as populações do Projeto *1000 Genomes*.

4.3.4 Região 3'NT

Todos os haplótipos identificados nas análises dos dados populacionais disponíveis no Projeto *1000 Genomes* e descritos em publicações prévias estão listados na Tabela 4.3.7. Um total de 21 SNVs que compõem 52 diferentes haplótipos foram identificados para a região 3'NT. Desses, somente 8 apresentaram frequências maiores que 1% na população mundial avaliada e 8 foram identificados pela primeira vez neste trabalho. As frequências dos haplótipos em cada uma das populações avaliadas estão descritas na Tabela 4.3.8 e os haplótipos com frequências maiores que 1% estão representados na Figura 4.3.10. Adicionalmente, foi criada uma rede de haplótipos *Median-Joining* a partir dos dados de frequência e as variações nucleotídicas dos haplótipos, a qual está apresentada na Figura 4.3.11. Nesse gráfico é possível verificar os haplótipos mais frequentes da região 3'NT (UTR-01, UTR-02 e UTR-03) e as suas variações dentro dos grupos populacionais.

Nas análises populacionais, foi verificado que a diversidade haplotípica da região 3'NT variou entre 0,8421 (PUR) a 0,6324 (CDX) quando todas as populações foram analisadas conjuntamente. Em cada um dos cinco grupos, os valores máximos e mínimos desse parâmetro foram os seguintes: populações americanas (AMR) 0,8421 (PUR) a 0,7279 (PEL); populações europeias (EUR) 0,8155 (CYP) a 0,7306 (FIN); populações africanas (AFR) 0,8308 (YAN) a 0,7410 (GUI); populações sul-asiáticas (SAS) 0,8060 (BEB) a 0,7336 (GIH); e populações leste-asiáticas (EAS) 0,7899 (KHV) a 0,6324 (CDX).

No teste de neutralidade genético-populacional de Ewens-Watterson, foram avaliados os valores F esperados e observados e a significância de Slatkin. Todos os valores de F observados são menores do que os valores esperados. O cálculo da significância de Slatkin apresentou valores significativos em 15 das 42 populações avaliadas (BSC, BRS, BPR, BPE, FRM, FRP, POR, MAL, YAN, SER, MSL, LWK, MDE, GIH e KHV).

Os valores de distância genética (F_{ST} de Reynolds) foram calculados nas populações aos pares e os dados foram utilizados para realizar a análise de MDS. O resultado obtido desta análise está apresentado na Figura 4.3.12. No primeiro gráfico (A), o qual mostra todo o grupo populacional avaliado, as populações JPT, CDX e CHS apresentam maior distanciamento genético; no segundo gráfico (B), quando somente as populações do Projeto *1000 Genomes* foram avaliadas, as mesmas populações (JPT, CDX e CHS) apresentaram as maiores distâncias

genéticas; e no terceiro gráfico (C), quando somente as populações da revisão bibliográfica foram avaliadas, as populações com maiores distanciamentos genéticos foram GUI, POR, YAN, SER e MAL.

No teste de diferenciação entre todos os pares de populações foram observadas diferenças significativas 50 em das 861 comparações realizadas. Por fim, na AMOVA, quando todas as populações avaliadas e considerando os cinco grupos populacionais, foi observado que 95,7% da variabilidade genética está dentro das populações, 2,56% entre grupos, 1,74% entre populações dentro dos grupos e valor de F_{ST} de 4,3%. Quando apenas as populações dos cinco grupos do Projeto *1000 Genomes* foram avaliadas, os valores de variabilidade observados foram: 94,33% da variabilidade genética está dentro das populações, 3,65% entre grupos, 2,02% entre populações dentro dos grupos valor de F_{ST} de 5,6%. Quando somente os dados populacionais da revisão de literatura foram utilizados, quatro grupos foram identificados (Americanas, Europeias, Africanas e Leste-asiáticas), e os valores de variância obtidos foram: 97,54% da variabilidade genética está dentro das populações, 1,18% entre grupos, 1,28% entre populações dentro dos grupos e valor de F_{ST} de 2,4%.

As tabelas contendo os dados completos das análises populacionais para a região 3'NT estão apresentadas no Apêndice D.

Tabela 4.3.7 – Haplótipos da região 3'NT do *HLA-G* e suas frequências na população mundial.

Posição chr6 hg19	29798563	29798581	29798608	29798610	29798617	29798634	29798639	29798642	29798651	29798659	29798699	29798714	29798718	29798728	29798749	29798784	29798790	29798794	29798803	29798834	29798896	Frequência dos haplótipos na População Mundial
Identificação do SNV	rs551200694	rs371194629	rs567747015	rs1707	rs1710	rs17179101	rs146339774	rs17179108	.	rs569057854	rs180827037	.	rs554784083	rs138249160	rs1063320	rs554076817	rs187320344	rs9380142	rs1610696	rs1233331	rs541542414	
SNV	2961	14 bp	3001	3003	3010	3027	3032	3035	3044	3052	3092	3107	3111	3121	3142	3177	3183	3187	3196	3227	3289	
Alelo Referência	T	G	C	C	G	C	G	C	A	C	G	C	G	T	C	G	G	A	C	G	C	
Alelo Alternativo	C	GATTTGTTTCATGCCT	T	T	C	A	C	T	T	T	T	G	A	C	G	T	A	G	G	A	T	2n = 5070
UTR-01	T	G	C	T	G	C	G	C	A	C	G	C	G	T	C	G	G	G	C	G	C	0,2577
UTR-02	T	GATTTGTTTCATGCCT	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,2496
UTR-03	T	G	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,2002
UTR-04	T	G	C	C	G	C	G	C	A	C	G	C	G	T	C	G	G	A	C	G	C	0,0923
UTR-05	T	GATTTGTTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0456
UTR-06	T	G	C	T	G	C	G	C	A	C	G	C	G	T	C	G	G	A	C	G	C	0,0304
UTR-07	T	GATTTGTTTCATGCCT	C	T	C	A	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0882
UTR-08	T	GATTTGTTTCATGCCT	C	T	G	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0002
UTR-09	T	GATTTGTTTCATGCCT	C	T	G	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-10	T	G	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0004
UTR-11	T	G	C	C	C	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-13	T	G	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0010
UTR-14	T	G	C	T	G	C	G	C	A	C	G	C	G	T	G	G	G	G	C	G	C	0,0002
UTR-15	T	GATTTGTTTCATGCCT	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0002
UTR-16	T	GATTTGTTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0000
UTR-17	T	GATTTGTTTCATGCCT	T	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0057
UTR-18	T	G	C	T	G	C	G	C	A	C	G	C	G	T	C	G	G	A	C	A	C	0,0187
UTR-19	T	GATTTGTTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	G	C	G	C	0,0000
UTR-20	T	G	C	T	G	C	C	C	A	C	G	C	G	T	C	G	G	A	C	G	C	0,0016
UTR-21	T	G	C	T	G	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000

UTR-22	T	GATTGTTCATGCCT	C	C	G	C	G	C	A	C	G	C	G	T	C	G	G	A	G	G	C	0,0000
UTR-23	T	G	C	T	C	C	G	C	A	T	G	C	G	T	G	G	G	A	C	G	C	0,0010
UTR-24	T	G	C	T	C	C	G	C	A	C	G	C	A	T	C	G	G	G	C	G	C	0,0000
UTR-25	T	GATTGTTCATGCCT	C	T	G	C	G	T	A	C	G	C	G	C	G	G	G	A	C	G	C	0,0000
UTR-26	T	GATTGTTCATGCCT	T	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0000
UTR-27	T	G	C	C	G	C	G	C	A	C	T	C	G	T	C	G	G	A	C	G	C	0,0004
UTR-29	T	G	C	T	G	C	G	C	A	C	G	C	A	T	C	G	G	G	C	G	C	0,0000
UTR-30	T	GATTGTTCATGCCT	C	T	G	C	G	C	A	C	G	C	G	T	C	G	G	G	C	G	C	0,0000
UTR-31	T	G	C	T	C	A	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-32	T	GATTGTTCATGCCT	C	T	C	C	G	C	A	C	G	G	G	T	G	G	G	A	G	G	C	0,0000
UTR-33	T	GATTGTTCATGCCT	C	T	C	C	G	T	A	C	G	G	G	T	G	G	G	A	C	G	C	0,0000
UTR-34	T	G	C	T	G	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0000
UTR-35	T	G	C	T	G	C	G	C	A	C	G	G	G	T	C	G	G	A	C	G	C	0,0000
UTR-36	T	G	C	T	G	C	G	C	A	C	G	G	G	T	C	G	G	G	C	G	C	0,0000
UTR-37	T	G	C	T	C	C	G	C	A	C	G	G	G	T	G	G	G	A	C	G	C	0,0000
UTR-38	T	G	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	G	C	G	C	0,0002
UTR-39	T	G	T	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-40	T	G	C	T	G	A	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-41	T	G	C	C	G	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-42	T	GATTGTTCATGCCT	C	T	G	A	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
UTR-43	T	G	C	C	G	C	G	C	A	C	G	C	G	T	G	G	G	A	G	G	C	0,0000
UTR-44	T	GATTGTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	C	G	G	G	A	C	G	C	0,0014
UTR-45	T	GATTGTTCATGCCT	C	C	G	C	G	T	A	C	G	C	G	C	G	G	G	A	C	G	C	0,0000
UTR-46	T	G	C	T	C	C	G	C	T	C	G	C	G	T	G	G	G	A	C	G	C	0,0000
NOVO-1	C	GATTGTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	C	0,0026
NOVO-2	T	GATTGTTCATGCCT	C	T	C	C	G	T	A	C	G	C	G	T	G	G	G	A	C	G	T	0,0010
NOVO-3	T	G	C	T	C	C	G	C	A	C	G	C	G	T	G	G	A	A	C	G	C	0,0002
NOVO-4	T	G	C	C	G	C	G	C	A	C	G	C	G	T	C	G	G	G	C	G	C	0,0002
NOVO-5	T	G	C	T	G	C	G	C	A	C	G	C	G	T	C	T	G	G	C	G	C	0,0002
NOVO-6	T	GATTGTTCATGCCT	C	T	C	C	G	C	A	C	G	C	G	T	C	G	G	A	G	G	C	0,0002
NOVO-7	T	G	C	T	C	C	G	C	A	C	G	C	G	T	G	G	G	A	C	G	T	0,0004
NOVO-8	T	G	C	T	C	C	G	C	A	C	G	C	A	T	G	G	G	A	C	G	C	0,0002

Tabela 4.3.8 - Frequências dos haplótipos da região 3'NT do *HLA-G* em diferentes populações do mundo (parte 1).

	Haplótipos 3'NTR <i>HLA-G</i> /Populações	UTR-01	UTR-02	UTR-03	UTR-04	UTR-05	UTR-06	UTR-07	UTR-08	UTR-10	UTR-13	UTR-14	UTR-15	UTR-17	UTR-18	UTR-20	UTR-23	UTR-27
Americanas	BSC	0,3050	0,2500	0,1300	0,0900	0,0650	0,0850	0,0600	0,0150									
	BRS	0,2805	0,2716	0,1131	0,0936	0,0930	0,0226	0,0138			0,0005							
	BPR	0,2790	0,2760	0,1740	0,0610	0,0840	0,0710	0,0420										
	BSP	0,2607	0,2695	0,1411	0,1190	0,0850	0,0509	0,0556	0,0047		0,0019		0,0005	0,0042		0,0018	0,0009	
	BPE	0,2910	0,2280	0,1580	0,0670	0,0870	0,0880	0,0350										
	CLM	0,2501	0,2340	0,2181	0,1436	0,0798	0,0106	0,0106							0,0479			
	MXL	0,2388	0,2760	0,1642	0,0896	0,0970	0,0149	0,0522	0,0075						0,0448			
	PEL	0,3546	0,3605	0,1221	0,0233	0,0465	0,0116	0,0349							0,0407	0,0058		
	PUR	0,2666	0,1762	0,1619	0,1238	0,1048	0,0190	0,0524			0,0143				0,0619			
Europeias	FRM	0,3350	0,2060	0,1230	0,1560	0,0280	0,0740	0,0740	0,0040									
	FRP	0,2510	0,2890	0,1310	0,1200	0,0560	0,0860	0,0590	0,0050									
	POR	0,2500	0,3420	0,0920		0,0420	0,0750											
	CYP	0,2486	0,2892	0,1405	0,0892	0,0622	0,0567	0,0784	0,0027	0,0027	0,0135							0,0027
	ITA	0,3480	0,3140	0,1080	0,1220	0,0290		0,0740										
	CEU	0,3889	0,2525	0,0909	0,1414	0,0303	0,0051	0,0707							0,0202			
	FIN	0,3687	0,2273	0,0606	0,2828	0,0101		0,0404							0,0101			
	GBR	0,3208	0,2989	0,0598	0,1250	0,0272		0,0978			0,0054				0,0543	0,0054		
	IBS	0,3459	0,2897	0,0981	0,0888	0,0327	0,0047	0,0514							0,0794			
	TSI	0,3009	0,3102	0,1343	0,1343	0,0324	0,0185	0,0417		0,0046	0,0046				0,0139			
Africanas	MAL	0,1340	0,2990	0,2390	0,0500	0,1430	0,1210											
	YAN	0,1460	0,1800	0,2800	0,1060	0,0770	0,1230	0,0030										
	SER	0,1150	0,3430	0,2720	0,0540	0,0940	0,0610											
	GUI	0,1920	0,3830	0,0670		0,0170	0,0750											
	TOR	0,2500	0,2670	0,2000	0,0830	0,0830	0,1000											
	ACB	0,1406	0,2500	0,2709	0,1094	0,0990	0,0885							0,0156	0,0052	0,0052		
	ASW	0,2120	0,3257	0,1667	0,0606	0,0985	0,0379	0,0076						0,0076	0,0303	0,0152	0,0152	
	ESN	0,1768	0,2373	0,2524	0,0960	0,0556	0,1414							0,0152	0,0051		0,0101	
	GWD	0,1814	0,2833	0,2567	0,0442	0,0929	0,0310						0,0044	0,0487	0,0354			

	MSL	0,2705	0,1941	0,2059	0,0824	0,1294	0,0588							0,0471	0,0118			
	YRI	0,1330	0,2615	0,3072	0,1284	0,0505	0,0780							0,0138	0,0092		0,0046	
	LWK	0,2524	0,3019	0,1337	0,0842	0,0842	0,0891								0,0297	0,0149		0,0099
Sul-Asiáticas	MDE	0,2360	0,3730	0,1540	0,0950	0,0710		0,0710										
	BEB	0,1512	0,2209	0,2209	0,1453	0,0116	0,0116	0,2385										
	GIH	0,1038	0,4528	0,1651	0,0943	0,0236	0,0330	0,1274										
	ITU	0,1844	0,3931	0,1748	0,0583	0,0437	0,0243	0,1165					0,0049					
	PJL	0,1875	0,3906	0,1667	0,0521	0,0260	0,0052	0,1667										
	STU	0,1165	0,2330	0,2717	0,1748	0,0146	0,0146	0,1602		0,0049								
Leste-Asiáticas	CHB	0,3057	0,1748	0,2718	0,0388	0,0243		0,1748										
	CHS	0,4306	0,0602	0,2639	0,0185			0,2222										
	CDX	0,4646	0,0152	0,1616				0,3586										
	KHV	0,3019	0,1089	0,2426	0,0545		0,0891	0,2030										
	JPT	0,2452	0,2019	0,5000	0,0048	0,0144		0,0337										
	Global	0,2577	0,2496	0,2002	0,0923	0,0456	0,0304	0,0882	0,0002	0,0004	0,0010	0,0002	0,0002	0,0057	0,0187	0,0016	0,0010	0,0004

Populações Americanas (AMR): BRS (residentes no estado do Rio Grande do Sul no Brasil), BPR (residentes no estado do Paraná no Brasil), BSP (residentes no estado de São Paulo no Brasil), BPE (residentes no estado de Pernambuco no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): FRM (residentes em Marselha na França), FRP (residentes em Paris na França), POR (residentes em Coimbra em Portugal), CYP (residentes no Chipre), ITA (residentes em Milão na Itália), CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): MAL (pessoas dos grupos étnicos Dogon, Peulh, Tamashek e Bambara, de Mali), YAN (residentes na República do Congo), SER (residentes no Senegal), GUI (residentes nativos de Guiné-Bissau), TOR (residentes em Tori-Bossito no Benin), ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): MDE (pessoas residentes em países do Oriente Médio), BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJL (residentes de Lahore em Punjab no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Tabela 4.3.8 - Frequências dos haplótipos da região 3'NT do *HLA-G* em diferentes populações do mundo (parte 2).

	Haplótipos 3'NTR <i>HLA-G</i> /Populações	UTR- 38	UTR- 39	UTR- 44	UTR- 46	NOVO- 1	NOVO- 2	NOVO- 3	NOVO- 4	NOVO- 5	NOVO- 6	NOVO- 7	NOVO- 8	Indeterminado	2n
Americanas	BSC														200
	BRS													0,1113	1276
	BPR													0,0130	380
	BSP			0,0023										0,0019	310
	BPE													0,0460	256
	CLM					0,0053									188
	MXL								0,0075	0,0075					134
	PEL														172
	PUR		0,0048			0,0048	0,0095								210
Europeias	FRM														246
	FRP													0,0030	374
	POR													0,1990	120
	CYP			0,0108	0,0027									0,0001	370
	ITA													0,0050	204
	CEU														198
	FIN														198
	GBR		0,0054												184
	IBS		0,0093												214
	TSI		0,0046												216
Africanas	MAL													0,0140	458
	YAN													0,0850	350
	SER													0,0610	478
	GUI													0,2660	120
	TOR													0,0170	60
	ACB					0,0104	0,0052								192
	ASW					0,0227									132
	ESN					0,0101									198
	GWD		0,0044			0,0044	0,0088						0,0044		226

Figura 4.3.10 - Distribuição dos haplótipos da Região 3'NT com frequências maiores que 1% nas populações avaliadas.

Populações Americanas (AMR): BRS (residentes no estado do Rio Grande do Sul no Brasil), BPR (residentes no estado do Paraná no Brasil), BSP (residentes no estado de São Paulo no Brasil), BPE (residentes no estado de Pernambuco no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (porto-riquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); Populações Europeias (EUR): FRM (residentes em Marselha na França), FRP (residentes em Paris na França), POR (residentes em Coimbra em Portugal), CYP (residentes no Chipre), ITA (residentes em Milão na Itália), CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); Populações Africanas (AFR): MAL (pessoas dos grupos étnicos Dogon, Peulh, Tamashek e Bambara, de Mali), YAN (residentes na República do Congo), SER (residentes no Senegal), GUI (residentes nativos de Guiné-Bissau), TOR (residentes em Tori-Bossito no Benin), ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); Populações Sul-Asiáticas (SAS): MDE (pessoas residentes em países do Oriente Médio), BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PIL (residentes de Lahore em Punjabi no Paquistão), STU (srilanqueses Tamil residentes no Reino Unido); Populações Leste-Asiáticas (EAS): CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã), JPT (Japoneses residentes em Tokyo no Japão); e população Global (todas as populações do Projeto 1000 Genomes).

Figura 4.3.11 - Rede *Median-Joining* dos haplótipos da região 3'NT do *HLA-G*.

As áreas dos círculos são proporcionais à frequência global dos haplótipos e os comprimentos dos ramos ao número de mutações que os separaram. Os círculos são coloridos de acordo com as frequências das regiões das populações (marrom – americanas, cinza escuro – africanas, verde médio – europeias, verde claro – sul-asiáticas e verde-escuro – leste-asiáticas).

Figura 4.3.12 - Dispersão em Escala Multidimensional considerando a Região 3'NT do *HLA-G*.

Nesta figura, estão apresentados os gráficos obtidos após a análise da Dispersão em Escala Multidimensional (MDS) considerando os valores das distâncias genéticas (F_{st} de Reynolds) obtidos a partir dos dados da Região 3'NT do *HLA-G* considerando todas as populações com dados disponíveis para essa região (A), somente nas populações do Projeto *1000 Genomes* (B) e somente com os dados populacionais encontrados por meio da revisão bibliográfica (C). As populações avaliadas foram: BRS (residentes no estado do Rio Grande do Sul no Brasil), BPR (residentes no estado do Paraná no Brasil), BSP (residentes no estado de São Paulo no Brasil), BPE (residentes no estado de Pernambuco no Brasil), MXL (pessoas com ancestralidade mexicana residentes em Los Angeles nos EUA), PUR (portorriquenhos residentes em Porto Rico), CLM (colombianos residentes em Medellín na Colômbia) e PEL (peruanos residentes em Lima no Peru); FRM (residentes em Marselha na França), FRP (residentes em Paris na França), POR (residentes em Coimbra em Portugal), CYP (residentes no Chipre), ITA (residentes em Milão na Itália), CEU (residentes de Utah com ancestralidade do norte e oeste europeu), TSI (residentes da Toscana na Itália), GBR (britânicos na Inglaterra e Escócia), FIN (finlandeses na Finlândia), IBS (ibéricos na Espanha); MAL (pessoas dos grupos étnicos Dogon, Peulh, Tamashek e Bambara, de Mali), YAN (residentes na República do Congo), SER (residentes no Senegal), GUI (residentes nativos de Guiné-Bissau), TOR (residentes em Tori-Bossito no Benin), ACB (africanos caribenhos em Barbados), MSL (tribo Mende em Serra Leoa), ESN (tribo Esan na Nigéria), GWD (gambianos residentes nas divisas da região oeste de Gâmbia), ASW (americanos com ancestralidade africana na região sudoeste dos EUA), YRI (iorubás residentes em Ibadan na Nigéria), LWK (luhyas residentes em Webuye no Quênia); MDE (pessoas residentes em países do Oriente Médio), BEB (residentes de Bengala em Bangladesh), GIH (indianos Gujaratis residentes em Houston, Texas), ITU (indianos Telugo residentes no Reino Unido), PJI (residentes de Lahore em Punjab no Paquistão), STU (srilankeses Tamil residentes no Reino Unido); CHB (chineses Han de Pequim na China), CHS (chineses Han do sul da China), CDX (chineses Dai de Xishuangbanna na China), KHV (Kinh residentes em Ho Chi Minh City no Vietnã) e JPT (Japoneses residentes em Tóquio no Japão).

4.3.5 Desequilíbrio de Ligação

O Desequilíbrio de Ligação (LD) entre os pares de SNP presentes em um segmento que compreende as regiões do *Enhancer L*, promotora-distal, promotora-proximal e 3'NT do gene *HLA-G*. A Figura 4.3.13, apresentada a seguir, foi gerada usando sítios variáveis com frequência alélica mínima $\geq 40\%$ (MAF $\geq 40\%$) e ponto de corte 0,05 para HWE. As áreas em vermelho ou cinza escuro indicam LD forte, áreas em vermelho claro ou cinza claro indicam LD moderado, áreas em azul ou quase branco indicam LD fraca e áreas brancas indicam que não há LD. As variantes do tipo inserção/deleção foram ocultadas desta análise. Na mesma figura está apresentado o desequilíbrio de ligação do gene *HLA-G* como um todo. Nas duas representações gráficas, é possível observar que um único bloco é formado, o que indica intenso desequilíbrio de ligação entre os polimorfismos analisados.

Figura 4.3.13 - Desequilíbrio de Ligação entre as regiões *Enhancer L*, promotora-distal, promotora-proximal e 3'NT do gene *HLA-G*.

Desequilíbrio de ligação entre os pares de SNP presentes em um segmento que compreende as regiões do *Enhancer L*, promotora-distal, promotora-proximal e 3'NT do gene *HLA-G*. A imagem foi gerada usando sítios variáveis com MAF $\geq 40\%$ e ponto de corte 0,05 para HW. Áreas em vermelho/cinza escuro indicam LD forte, vermelho claro/cinza claro indicam LD moderado, azul/quase branco indicam LD fraca e brancas indicam que não há LD. Variantes do tipo inserção/deleção foram ocultadas desta análise. No quadrante inferior, está apresentado o desequilíbrio de ligação do gene *HLA-G* como um todo.

4.4 Relação entre as concentrações de HLA-G solúvel e os haplótipos das regiões regulatórias do gene *HLA-G*, em amostras de indivíduos saudáveis

Os dados apresentados na Tabela 4.4.1 correspondem aos haplótipos das regiões reguladoras Promotora-Distal, Promotora-Proximal e 3'NT do gene *HLA-G*, que foram identificados após a análise dos dados genéticos das 80 amostras de indivíduos saudáveis (doadores de sangue) pertencentes ao banco de dados e amostras do nosso laboratório. A relação da expressão de sHLA-G nessas amostras e os haplótipos/grupos de haplótipos identificados estão demonstradas nas figuras apresentadas a seguir, e são referentes às regiões reguladoras Promotora-Distal (Figuras 4.4.1 e 4.4.2), Promotora-Proximal (Figuras 4.4.3 e 4.4.4) e 3'NT (Figuras 4.4.5 e 4.4.6). Para a realização dos testes estatísticos, foram incluídos somente os haplótipos que foram identificados em três ou mais indivíduos. Nos gráficos, estão representados os resultados das médias e dos desvios-padrão por região reguladora.

Tabela 4.4.1 – Haplótipos identificados nas regiões reguladoras Promotora-Distal, Promotora-Proximal e 3'NT do gene *HLA-G* em uma amostra de indivíduos saudáveis.

Nome	Identificação
Região Promotora-Distal	
DISTAL-010101a	T C T A C A A C T A A G G G C T C T T T T C A G T C T C T C T T C
DISTAL-010101b	T C T A C A A C T A A G G G C T C T T T T G A G T C T C T C T T C
DISTAL-010102a	T C T A T G A C T A G A G A C T C T T T T C G G T C C A T C C T C
DISTAL-010102b	T C T A T A A C T A G A G A C T C T T T T C G G T C C A T C C T C
DISTAL-0103a	T C T A C A G T T A A A G G C T C T C T T C A G T C T C T C C T C
DISTAL-0104a	T C T A C A A C T A G A A G C T T T C G A T C T C T C C T C
Região Promotora-Proximal	
PROMO-G010101a/f/d/j	A G T T G A G A A C G G G C C C C T A G A G G C
PROMO-G010101b	A G T T G A G A A C G G G C C C G T A G A G G C
PROMO-G010101c	A G T T G A G A A T G G G C C C G T A G A G G C
PROMO-G010102a	A G T T A G G T A C G G A C C T C G G T A A A C
PROMO-G0103a	A G T T G G G A G C G G G C C C T T A G A G G T
PROMO-G0104a/b	A G T T A G A A A C G G A C C T C G G T A A A C

Região 3'NT

UTR-1	T G C T G C G C A C G C G T C G G G C G C
UTR-2	T G A T T T G T T C A T G C C T C T C C G C A C G C G T G G G A G G C
UTR-3	T G C T C C G C A C G C G T G G G A C G C
UTR-4	T G C C G C G C A C G C G T C G G A C G C
UTR-5	T G A T T T G T T C A T G C C T C T C C G T A C G C G T G G G A C G C
UTR-18	T G C T G C G C A C G C G T C G G A C A C

Figura 4.4.1 – Expressão de sHLA-G de acordo com os haplótipos da região Promotora-Distal.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos haplótipos da região promotora distal. Somente os haplótipos com três ou mais ocorrências estão apresentados neste gráfico ($n = 62$; $p = 0,920$). O número de amostras avaliadas, as médias e os desvios-padrões encontrados para cada um dos haplótipos analisados foram, respectivamente: 010101a_010101a ($n = 9$; $\bar{x} = 5,92$, $s = 5,42$), 010101a_010102a ($n = 14$; $\bar{x} = 7,61$, $s = 6,24$), 010101b_010102a ($n = 6$; $\bar{x} = 5,56$, $s = 8,68$), 010101a_0104a ($n = 12$; $\bar{x} = 7,59$, $s = 5,42$), 010102a_010102a ($n = 6$; $\bar{x} = 6,06$, $s = 6,89$), 0104a_010102a ($n = 3$; $\bar{x} = 5,81$, $s = 7,94$), 0104a_0103a ($n = 4$; $\bar{x} = 12,1$, $s = 8,24$), 010101a_0103a ($n = 5$; $\bar{x} = 5,40$, $s = 6,24$) e 010101a_010102b ($n = 3$; $\bar{x} = 7,37$, $s = 9,81$).

Figura 4.4.2 – Expressão de sHLA-G de acordo com os grupos haplótipos da região Promotora-Distal.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos grupos de haplótipos da região promotora-distal. Somente os grupos de haplótipos com três ou mais ocorrências estão apresentados neste gráfico ($n = 62$; $p = 0,842$). O número de amostras avaliadas, as médias e os desvios-padrões encontrados para cada um dos haplótipos analisados foram, respectivamente: 010101_010101 ($n = 9$; $\bar{x} = 5,92$, $s = 5,42$), 010101_010102 ($n = 23$; $\bar{x} = 6,77$, $s = 7,03$), 010101_0103 ($n = 5$; $\bar{x} = 5,40$, $s = 6,24$), 010101_0104 ($n = 12$; $\bar{x} = 7,60$, $s = 5,42$), 010102_010102 ($n = 6$; $\bar{x} = 6,06$, $s = 6,89$), 010102_0104 ($n = 3$; $\bar{x} = 5,81$, $s = 7,94$) e 0103_0104 ($n = 4$; $\bar{x} = 12,10$, $s = 8,24$).

Figura 4.4.3 – Expressão de sHLA-G de acordo com os haplótipos da região Promotora-Proximal.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos haplótipos da região promotora-proximal. Somente os haplótipos com três ou mais ocorrências estão apresentados neste gráfico ($n = 56$; $p = 0,277$). O número de amostras avaliadas, as médias e os desvios-padrões encontrados para cada um dos haplótipos analisados foram, respectivamente: 010101afdj_010101afdj ($n = 6$; $\bar{x} = 3,71$, $s = 6,02$), 010101afdj_010101c ($n = 3$; $\bar{x} = 7,79$, $s = 2,74$), 010101afdj_010102a ($n = 14$; $\bar{x} = 5,41$, $s = 6,18$), 010101afdj_0104ab ($n = 7$; $\bar{x} = 8,53$, $s = 6,23$), 010101b_010102a ($n = 5$; $\bar{x} = 11,08$, $s = 8,36$), 010101b_0104ab ($n = 3$; $\bar{x} = 8,59$, $s = 1,67$), 010101c_010102a ($n = 4$; $\bar{x} = 9,67$, $s = 6,74$), 010102a_010102a ($n = 8$; $\bar{x} = 6,37$, $s = 6,54$), 0104ab_010102a ($n = 3$; $\bar{x} = 5,81$, $s = 7,94$) e 0104ab_0103a ($n = 3$; $\bar{x} = 7,46$, $s = 6,41$).

Figura 4.4.4 – Expressão de sHLA-G de acordo com os grupos de haplótipos da região Promotora-Proximal.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos grupos de haplótipos da região promotora-proximal. Somente os grupos de haplótipos com três ou mais ocorrências estão apresentados neste gráfico ($n = 56$; $p = 0,232$). O número de amostras avaliadas, as médias e os desvios-padrões encontrados para cada um dos haplótipos analisados foram, respectivamente: 010101_010101 ($n = 9$; $\bar{x} = 5,07$, $s = 5,32$), 010101_010102 ($n = 23$; $\bar{x} = 7,38$, $s = 6,93$), 010101_0104 ($n = 10$; $\bar{x} = 8,55$, $s = 5,15$), 010102_010102 ($n = 8$; $\bar{x} = 6,37$, $s = 6,53$), 010102_0104 ($n = 3$; $\bar{x} = 5,81$, $s = 7,94$) e 0103_0104 ($n = 6$; $\bar{x} = 16,14$, $s = 2,03$).

Figura 4.4.5 – Expressão de sHLA-G de acordo com os haplótipos da região 3'NT.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nos haplótipos da região 3'NT. Somente os haplótipos com três ou mais ocorrências estão apresentados neste gráfico (n = 60; p = 0,891). O número de amostras avaliadas, as médias e os desvios-padrões encontrados para cada um dos haplótipos analisados foram, respectivamente: UTR1_UTR1 (n = 3; $\bar{x}$ = 6,73, s = 7,85), UTR1_UTR2 (n = 6; $\bar{x}$ = 5,03, s = 4,61), UTR1_UTR3 (n = 6; $\bar{x}$ = 8,69, s = 6,81), UTR1_UTR4 (n = 6; $\bar{x}$ = 5,52, s = 4,67), UTR1_UTR5 (n = 4; $\bar{x}$ = 5,97, s = 6,58), UTR18_UTR2 (n = 6; $\bar{x}$ = 6,98, s = 8,27), UTR2_UTR2 (n = 7; $\bar{x}$ = 5,46, s = 6,49), UTR3_UTR2 (n = 3; $\bar{x}$ = 5,81, s = 7,94), UTR3_UTR5 (n = 4; $\bar{x}$ = 12,11, s = 8,24), UTR4_UTR2 (n = 9; $\bar{x}$ = 9,05, s = 6,70) e UTR4_UTR3 (n = 6; $\bar{x}$ = 6,49, s = 3,92).

Figura 4.4.6 – Expressão de sHLA-G de acordo com a presença dos dois haplótipos mais frequentes da região 3'NT.

Nesta figura, estão as comparações dos níveis de sHLA-G encontrados nas amostras que apresentavam os dois haplótipos mais frequentes da região 3'NT (UTR1 e UTR2). Os grupos foram divididos da seguinte forma: UTR1_UTR* (indivíduos que possuíam a UTR1 e qualquer outra UTR), UTR1_UTR2 (indivíduos com a UTR1 e UTR2) e UTR2_UTR* (indivíduos com UTR2 e qualquer outra UTR). Foram avaliadas 50 amostras ($p = 0,785$). O número de amostras avaliadas em cada grupo, as médias e os desvios-padrões encontrados foram, respectivamente: UTR1_UTR* ($n = 19$, $\bar{x} = 6,81$, $s = 5,91$), UTR1_UTR2 ($n = 6$; $\bar{x} = 5,03$, $s = 4,61$), UTR2_UTR* ($n = 25$; $\bar{x} = 7,15$, $s = 6,88$).

5. DISCUSSÃO

5.1 Avaliação da associação dos alelos do polimorfismo de 14-pb do gene *HLA-G* em doenças de diversas etiologias

Por intermédio da revisão sistemática seguida de meta-análise realizada neste trabalho, um número considerável de artigos publicados foi avaliado (125), nos quais, o principal objetivo, foi avaliar possíveis associações entre o polimorfismo de 14-pb (presença ou ausência) da 3'NT do gene *HLA-G* com diversas doenças. Os trabalhos publicados e analisados nesta seção do presente estudo foram realizados em vários países, sendo a maioria deles realizados no Brasil, China e Itália (Tabela 4.1.1). Dentre as principais razões para esse grande número de estudos é a suposição de que esse polimorfismo pode ser diretamente relacionado com os níveis de expressão de *HLA-G*. Uma vez que o *splicing* alternativo que ocorre nesta região pode gerar sequências completas, bem como, sequências com deleções de 14-pb e 92-pb, níveis de expressão distintos de *HLA-G* podem ser observados (HARRISON et al., 1993; HVIID et al., 2003, 2006; CASTELLI et al., 2014). É também importante destacar que, o polimorfismo de 14-pb é facilmente detectado por técnicas simples e baratas, o que possivelmente justifica a escolha desse marcador para estudos genéticos. Nos artigos avaliados, a grande maioria dos dados foi obtida utilizando a técnica de PCR seguida de eletroforese de gel de agarose ou poliacrilamida (Tabela 4.1.2).

O polimorfismo de 14-pb é o sítio de variação mais estudado do *HLA-G*. Além disso, esta meta-análise revelou inconsistências entre alguns estudos de associação previamente publicados. Essa consideração é baseada nos seguintes resultados, que foram apresentados na Tabela 4.1.3: i) para muitas condições (por exemplo, Câncer de Colo de Útero, Carcinoma Hepatocelular, Esclerose Múltipla, Artrite Reumatoide, Transmissão Vertical de HIV, infecção por Papilomavírus Humano, Transplante Cardíaco, Transplante Renal, Pré-eclâmpsia - quando os alelos maternos foram analisados, Falhas de Implantação, Aborto), alguns estudos encontraram associação do alelo de inserção de 14-pb com a susceptibilidade para o desenvolvimento da doença (BRIEN et al., 2001; MOREAU et al., 2008; LARSEN et al., 2010; ZHANG et al., 2012; LASHLEY et al., 2014; ARJMAND; SAMADI, 2015; AMODIO et al., 2016), outros revelaram a associação do alelo de deleção de 14-pb com susceptibilidade (GAZIT et al., 2004; GLAS et al., 2007; VEIT et al., 2008; ZHENG et al., 2009; MISRA et al., 2013) e outros ainda não encontraram associação de susceptibilidade com nenhum dos dois alelos do polimorfismo de 14-pb (KRONER et al., 2007; CRISPIM et al., 2008; CHIUSOLO et al., 2011; TWITO et al., 2011; ZIDI et al., 2015; BEN FREDJ et al., 2016; LAZARTE et al., 2016); ii) quando

as doenças foram agrupadas de acordo com uma hipotética alta (1) ou baixa (2) expressão de *HLA-G*, não foram verificadas associações entre os grupos e os alelos de 14-pb. Nessas comparações, cânceres e células infectadas cronicamente por vírus podem expressar *HLA-G* para reduzir ou impedir a ação de células citotóxicas do organismo hospedeiro e, nessas condições, a expressão de *HLA-G* poderia hipoteticamente aumentar, independentemente do genótipo do indivíduo. Por outro lado, em doenças autoimunes, doenças inflamatórias crônicas e transplantes uma baixa expressão de *HLA-G* pode ser hipoteticamente associada com um quadro clínico mais brando. Assim, o esperado seria que o alelo de deleção dos 14-pb fosse associado com pior desfecho no caso de tumores e infecções virais crônicas e com um melhor prognóstico nos casos de doenças autoimunes e transplantes. De acordo com a presente meta-análise, essas associações não foram confirmadas; iii) quando as doenças foram agrupadas de acordo com as principais características, foi verificada a associação de risco entre o grupo Doenças Reprodutivas (composto por estudos sobre Falha de Implantação, Pré-eclâmpsia, Aborto Recorrente, Falha Reprodutiva e Aborto Espontâneo) e o alelo de presença dos 14-pb (Figura 4.1.2). Outras meta-análises realizadas por Fan et al. (2017) e Wang et al. (2013) também verificaram o alelo de presença de 14-pb como risco para Falhas de Implantação e Aborto, respectivamente.

A presença da inserção dos 14-pb foi associada com susceptibilidade: i) à Pré-eclâmpsia, quando os alelos dos filhos foram avaliados e ii) ao Lúpus Eritematoso Sistêmico (Figura 4.1.3). Esse último dado está de acordo com as meta-análises realizadas por Catamo et al. (2015) e Zhang et al. (2014) que também descrevem o alelo de inserção de 14-pb como risco para Lúpus Eritematoso Sistêmico. Enquanto que o alelo de deleção dos 14-pb foi associado com o risco para infecção por Citomegalovírus e Câncer de Mama (Figura 4.1.4). A meta-análise realizada por Ge et al. (2014), similarmente, descreve o alelo de ausência de 14-pb como fator de risco para Câncer de Mama em populações asiáticas.

Analisados conjuntamente, esses achados indicam que a variação de 14-pb na 3'NT do *HLA-G* não é um marcador genético eficiente para estudos de associação de doenças. É necessário considerar que essas são doenças complexas e, embora a presença de *HLA-G* tenha sido previamente descrita na maioria dessas doenças, provavelmente, não é uma relação direta de causa-efeito. Além disso, esse polimorfismo não é o único elemento responsável pelo controle pós-transcricional da expressão de *HLA-G*. Como previamente mencionado, além do polimorfismo de 14-pb (que teve o alelo de inserção associado com baixa expressão de RNAm), pelo menos outros dois sítios de variação

dessa região (+3142C/G e +3187A/G) (TAN et al., 2007; VEIT; CHIES, 2009) têm sido funcionalmente estudados. Os alelos +3142G (associado com baixa expressão devido a ação de determinados miRNA) e +3187A (associado com baixa estabilidade do RNAm devido à presença de um motivo rico em AU) têm sido associados com baixa expressão de *HLA-G*. Além dessas, outros sítios de variação (+3003 C/G, +3010 G/C, +3027 C/A, +3035 C/T, +3196 C/G), que ainda não foram funcionalmente estudados, têm potencial para modular a expressão pós-transcricional de *HLA-G* conforme evidências demonstradas em estudos *in silico* (CASTELLI et al., 2009, 2014; PORTO et al., 2015).

Outra evidência é baseada na associação de sítios de variação da 3'NT e níveis de *HLA-G* solúvel, a qual indica associação dos polimorfismos de 14-pb, +3010 G/C, +3027 C/A e +3035 C/T com a expressão de *HLA-G* (MARTELLI-PALOMINO et al., 2013). Sendo assim, pelo menos quatro dos oito polimorfismos comumente observados na população mundial podem exercer, sozinhos ou em combinação, efeitos na regulação da expressão pós-transcricional de *HLA-G*. Portanto, para estudar a influência dos polimorfismos desse segmento e suas associações com doenças, seria mais interessante realizar a tipificação de toda a região 3'NT, ao invés de fazer unicamente a genotipagem do polimorfismo de 14-pb. Além disso, para compreender melhor a expressão de *HLA-G* também é necessário levar em consideração a regulação transcricional do gene, a qual é determinada principalmente pela região promotora. Uma vez que não adianta os mecanismos de regulação pós-transcricional serem conhecidos se não acontecer a transcrição do gene. Estudos adicionais, como ensaios *in vitro*, podem ajudar a compreender o papel das regiões reguladoras do *HLA-G*, o que poderia afetar as interações do gene com fatores de transcrição e miRNA e, consequentemente, a expressão do gene. Nesse sentido, é também importante considerar o intenso desequilíbrio de ligação do *HLA-G*, que poderia, eventualmente, justificar as associações observadas entre os alelos do polimorfismo de 14-pb e algumas doenças. Muitas das associações observadas nos estudos avaliados nesta meta-análise podem ser reflexo da ação de outro polimorfismo do *HLA-G*.

Para compreender um pouco das relações entre o polimorfismo de 14-pb e outros sítios polimórficos do *HLA-G*, foi organizada a Tabela 4.1.4, a qual contém as principais associações entre os alelos do polimorfismo de 14-pb e os haplótipos mais frequentes das regiões promotora, codificadora e 3'NT e as proteínas *HLA-G* formadas. Considerando a tabela como um todo, percebe-se que o polimorfismo de 14-pb não é um bom marcador para compreender os diferentes níveis de expressão de *HLA-G*, nem para avaliar as funções do gene, uma vez que existem grandes similaridades entre os haplótipos mais frequentes para ambos alelos.

Por fim, embora algumas associações tenham sido observadas, a lição aprendida nesta meta-análise é que o polimorfismo de 14-pb da 3'NT do *HLA-G* deve ser analisado no contexto da região 3'NT como um todo ou até mesmo do gene inteiro, uma vez que, como dito anteriormente, esse gene está sob forte desequilíbrio de ligação (CASTELLI et al., 2011). Esta é a primeira meta-análise que inclui avaliação global dos dados de todos os artigos de associação que tiveram como objeto de estudo o polimorfismo de 14-pb do *HLA-G*. Como previamente mencionado, as outras meta-análises que avaliaram esse sítio de variação focaram em somente uma doença (WANG; JIANG; ZHANG, 2013; FAN et al., 2014, 2017; ZHANG et al., 2014; CATAMO et al., 2015b; LEE; BAE; SONG, 2015; PABALAN et al., 2015; COELHO et al., 2016; NIU et al., 2017; SHI et al., 2017) ou em um pequeno grupo de doenças (GE et al., 2014; KIM et al., 2015; LI et al., 2015). Considerando as meta-análises realizadas no presente estudo, pode-se concluir que, embora o polimorfismo de 14-pb da região 3'NT do *HLA-G* é um sítio de variação de fácil detecção e amplamente utilizado em estudos de associação com doenças, esse polimorfismo não pode ser usado como um marcador genético isolado para avaliar susceptibilidade a doenças.

5.2 Avaliação da expressão de HLA-G e dos polimorfismos das regiões regulatórias do gene *HLA-G*, em um estudo caso-controle para a Doença de Chagas

Tendo como base os resultados obtidos na meta-análise apresentada na seção anterior, esta seção do trabalho teve como objetivo avaliar a expressão de HLA-G e os polimorfismos das duas regiões regulatórias mais estudadas até o momento (3'NT e promotora-proximal), por meio de uma metodologia desenvolvida por nosso grupo de pesquisa, em amostras de pacientes com diagnóstico para Doença de Chagas. Para isso, um estudo caso-controle foi conduzido, o qual teve como premissa que, assim como em outras doenças infecciosas, haveria aumento da expressão de HLA-G em pacientes chagásicos, o que facilitaria a infecção e o acometimento de órgãos pelo parasita devido à modulação negativa do sistema imune. Os resultados obtidos confirmam a hipótese inicial, tanto quanto à expressão de HLA-G solúvel (as amostras dos pacientes tinham níveis significativamente mais altos de HLA-G do que as amostras do grupo controle – FIGURA 4.2.1 e FIGURA 4.2.2), como quanto os relativos à expressão tecidual da proteína (nas amostras de imunoistoquímica de coração, esôfago e cólon de pacientes chagásicos, que em situações normais não apresentam expressão de HLA-G, foram detectadas marcações para a proteína – FIGURA 4.2.4 e TABELA 4.2.1). Ambas análises demonstraram que os pacientes com a doença possuem expressão aumentada da proteína. Ainda,

em estudo prévio do nosso grupo, observamos diminuição da expressão de HLA-G nas áreas de desnervação de coração e cólon de pacientes chagásicos (DIAS et al., 2015).

Considerando ainda que a patogênese dessa doença possa estar relacionada à variabilidade genética tanto do parasita quanto do paciente (hospedeiro), as regiões regulatórias do gene *HLA-G* foram avaliadas, incluindo as regiões promotora-proximal e 3'NT. Conforme relatado previamente, não foram observadas associações entre os haplótipos e/ou SNVs do gene. Os resultados obtidos neste trabalho não confirmaram as associações verificadas em um estudo anterior de Dias e colaboradores (2015), no qual os genótipos +3003 T/T, +3187 G/G e +3196 G/C foram associados com o risco para a Doença de Chagas. Cabe salientar que, neste estudo os SNVs foram detectados por sequenciamento de nova geração, ao passo que o realizado por Dias et al (2015) utilizou sequenciamento capilar (sequenciamento de Sanger).

Como demonstrado na seção anterior relativa à meta-análise, essas inconsistências nas associações encontradas também foram observadas em estudos caso-controle para outras doenças em que o *HLA-G* foi avaliado. As discordâncias observadas no presente trabalho, em relação ao estudo anterior de Dias e colaboradores (2015), reafirmam a necessidade de ampliar as análises das regiões regulatórias do gene que estão sendo avaliadas, para compreender se os polimorfismos do *HLA-G* estão de fato envolvidos na susceptibilidade à doença e em variações da resposta imune do paciente, ou se o aumento da expressão da proteína, verificado no plasma e nos tecidos, é uma consequência da resposta imune à infecção crônica causada pelo parasita *Trypanosoma cruzi*. Além disso, seria interessante reproduzir essa metodologia em outras populações, para verificar se são encontradas ou não associações. Pois, havendo somente um estudo para comparação, torna-se difícil concluir sobre os efeitos dos polimorfismos do gene *HLA-G* e da expressão da proteína em pacientes com diagnóstico para Doença de Chagas.

5.3 Caracterização das regiões regulatórias do gene *HLA-G* na população mundial

Esta seção do trabalho forneceu uma análise abrangente das variações genéticas de regiões regulatórias do gene *HLA-G* na população mundial. Nela foram incluídas as duas regiões amplamente estudadas (e avaliadas na seção anterior): promotora-proximal e 3'NT; e duas regiões recentemente descritas como possíveis reguladoras da expressão de *HLA-G*: promotora-distal e *Enhancer* L. As regiões foram analisadas para verificar polimorfismos conhecidos, identificar novos polimorfismos e conhecer os haplótipos formados por suas combinações. Os dados das populações avaliadas foram

divididos, de acordo com a definição feita pelo Projeto *1000 Genomes*, considerando cinco regiões continentais, em: americanas, europeias, africanas, sul-asiáticas e leste-asiáticas.

Para as duas regiões promotoras (promotora-distal e promotora-proximal) foram utilizados os dados coletados na Fase 3 do Projeto *1000 Genomes* e os dados de uma amostra populacional da região sudeste brasileira, mais especificamente do estado de São Paulo. Apesar de serem representativas da mesma região, as amostras avaliadas não foram as mesmas, nem conduzidas no mesmo trabalho. A região promotora-distal foi caracterizada em diferentes populações mundiais, pela primeira vez, no presente trabalho. A região promotora-proximal havia sido avaliada, de maneira semelhante, por Gineau e colaboradores (2014). No trabalho em questão, foram utilizados os dados da Fase 2 do Projeto *1000 Genomes*. Dentre as diferenças das fases do Projeto *1000 Genomes*, incluem-se: i) a ampliação do número de populações (de 21 para 26), ii) aumento do número de indivíduos (de 1936 para 2535), iii) aperfeiçoamento da qualidade das metodologias utilizadas para o sequenciamento das amostras e iv) melhor avaliação dos dados. Considerando as características dos HLA não-clássicos, ambas regiões apresentaram alto número de SNVs e haplótipos. Tendo a região promotora-distal 48 SNVs e 48 haplótipos (TABELA 4.3.1), enquanto que a região promotora-proximal teve 58 SNV e 55 haplótipos (TABELA 4.3.3), muitos deles identificados pela primeira vez no presente trabalho. Apesar do número considerável de diferentes haplótipos encontrados nas duas regiões, somente cerca de 20% deles possuem frequências maiores que 1% na população mundial (FIGURA 4.3.1 e FIGURA 4.3.4). Para as avaliações da região 3'NT, foram usados os dados coletados das 26 populações, da Fase 3 do Projeto *1000 Genomes*, e os dados de 16 amostras populacionais obtidos através de revisão de literatura. Esta região havia sido avaliada, anteriormente, em um trabalho publicado por Sabbagh e colaboradores (2014), no qual foram avaliados os dados da Fase 2 do Projeto *1000 Genomes*. Assim como as regiões promotoras supracitadas, a região 3'NT apresenta alto número de SNVs e haplótipos, quando consideradas as características dos HLA não-clássicos. Nela foram observados 21 SNVs que compõem 52 diferentes haplótipos, oito dos quais foram identificados no presente trabalho (TABELA 4.3.7). Da mesma forma que as regiões promotoras, apesar do número de haplótipos encontrados, apenas 20% deles têm frequências maiores que 1% na população mundial e apenas três deles (UTR-1, UTR-2 e UTR-3) correspondem a mais de 70% dos haplótipos observados (FIGURA 4.3.10). Além disso, uma característica interessante é que os dois haplótipos mais frequentes (UTR-1 e UTR-2) são divergentes em grande parte dos SNPs observados na região (TABELA 4.3.7). As diferenças de frequências dos haplótipos das regiões nas populações

mundiais estão graficamente representadas pelas redes *Median-Joining* (FIGURA 4.3.2, FIGURA 4.3.5 e FIGURA 4.3.11). Nelas é possível observar que as maiores diferenças estão nos grupos populacionais asiáticos.

Analisando os resultados obtidos da AMOVA, as duas regiões promotora-distal e promotora-proximal e a região 3'NT mostraram alta diversidade haplotípica intrapopulacional. Cerca de 95% da variabilidade genética dessas regiões está dentro das populações. Além disso, os valores de F_{ST} encontrados para as três regiões foram em torno de 4%. Esses valores estão muito abaixo dos valores tipicamente observados para *loci* do genoma nuclear (10 a 15%) e são contrastantes, pois o esperado é que houvesse uma perda gradual da diversidade com o aumento da distância da África, que seria resultante da colonização e da dispersão dos humanos pelo planeta. Os dados sugerem, portanto, padrões de variação que não podem ser explicados somente como resultado da história populacional humana. A avaliação da diversidade haplotípica mostrou que os valores médios encontrados para essas regiões (promotora-distal: 0,74; promotora-proximal: 0,77; e 3'NT: 0,78) são mais altos que os valores da maioria dos *loci* de variantes nucleares humanas.

Os gráficos obtidos nas análises de MDS (FIGURA 4.3.3, FIGURA 4.3.6 e FIGURA 4.3.12), os quais avaliaram dados de distância genética entre as populações, mostraram que a maioria das populações estão próximas e que três das populações asiáticas, CDX, CHS e JPT, e a população FIN são as que mais se distanciam do restante do grupo. Durante as análises da região promotora-proximal, foi possível observar um desvio dos resultados ocasionado pela inclusão da população BSP (como pode ser observado na FIGURA 4.3.6). Dentre os possíveis motivos estão as limitações da técnica utilizada, as quais poderiam ocasionar discrepâncias na interpretação dos dados. Por isso, optou-se por avaliar e discutir apenas os dados referentes ao grupo amostral que inclui apenas os dados provenientes do Projeto *1000 Genomes*, excluindo, portanto, essa população das análises.

Considerando os estudos realizados para avaliação das regiões do *HLA-G* conjuntamente com os dados obtidos no presente trabalho, verifica-se nessas regiões: (i) alto número de polimorfismos e haplótipos; (ii) elevada diversidade nucleotídica; (iii) baixa diferenciação populacional; e (iv) presença de dois haplótipos de alta frequência divergentes, no caso da região 3'NT. Essas características observadas são compatíveis com seleção balanceadora. Evidências de seleção balanceadora na região promotora do gene *HLA-G* haviam sido sugeridas em outros três estudos prévios (GINEAU et al., 2014; TAN et al., 2005; CASTELLI et al., 2011). Tendo em consideração, como discutido nas outras seções desta tese, que as variações de expressão da proteína *HLA-G* podem ser

vantajosas ou deletérias, a seleção balanceadora, com provável manutenção de um haplótipo de alta e um de baixa expressão da proteína seria benéfico para o indivíduo, uma vez que o mesmo teria capacidade de modular as suas respostas a diferentes agentes estressores, como mudanças ambientais e exposição a patógenos.

A região *Enhancer L*, por sua vez, é uma pequena região (121 pb) que apresenta 5 SNVs e 6 haplótipos, sendo que somente dois deles possuem frequências polimórficas na população mundial (TABELA 4.3.5). Esses dois haplótipos são caracterizados por possuírem um único SNV (rs1077433) como diferença. O haplótipo nominado neste trabalho *Enhancer-01*, possui uma guanina; enquanto que o haplótipo *Enhancer-02* possui uma adenina nessa posição, sendo que o haplótipo *Enhancer-02* possui uma frequência quatro vezes menor na população mundial. Essa proporção é mantida em todos os grupos populacionais, com exceção do grupo das populações leste-asiáticas que difere das demais, pois nele o haplótipo *Enhancer-01* continua sendo o mais frequente, mas em proporções mais elevadas, chegando a corresponder a 98% dos haplótipos para esta região na população CDX (FIGURA 4.3.7). Os outros haplótipos aparecem pontualmente em algumas populações, não sendo verificados na maioria das populações avaliadas (TABELA 4.3.8). Ainda sobre a diversidade haplotípica da região, a população GIH é a que tem a maior diversidade, tendo os dois haplótipos frequências próximas de 50% e a população CDX, a menor diversidade (98% e 2%). E, portanto, é também essa última população que apresenta a maior distância genética das demais populações, seguida por outras populações asiáticas (CHS, KHV e GIH) (FIGURA 4.3.8 e FIGURA 4.3.9). Como os dados provenientes dessa região se limitam, praticamente, a dois haplótipos, é difícil afirmar se há algum tipo de seleção atuando na região.

Analisando as regiões reguladoras estudadas nesta tese como um todo, observou-se que há um alto DL entre os pares de SNP que compreendem essa região do gene *HLA-G*, corroborando dados de estudos anteriores que haviam estudado algumas regiões especificamente (SABBAGH et al., 2014; GINEAU et al., 2014; CASTELLI et al., 2011). Conforme demonstrado na Figura 4.3.13, nessa região um único bloco é formado, o que indica intenso desequilíbrio de ligação entre os polimorfismos analisados e que, provavelmente, esse segmento é herdado em conjunto. Portanto, os SNVs observados em uma região podem ser diretamente relacionados com os de outra região do segmento gênico.

5.4 Relação entre as concentrações de HLA-G solúvel e os haplótipos das regiões regulatórias do gene *HLA-G*

Nesta seção, avaliamos as concentrações de sHLA-G, considerando os haplótipos das regiões regulatórias do *HLA-G* em uma amostra de indivíduos saudáveis, no intuito de compreender como e quais variações polimórficas estão envolvidas na expressão de HLA-G. Os dados foram avaliados para três das regiões abordadas nesta tese: promotora-distal, promotora-proximal e 3'NT. A região *Enhancer L* não foi avaliada, pois ela não havia sido contemplada pelo sequenciamento que gerou o banco de dados que foi utilizado para estas análises. Para quantificar os níveis de sHLA-G diretamente nas amostras de plasma, conforme proposto, as análises tiveram que ser realizadas considerando os dois haplótipos de cada indivíduo. Na Tabela 4.4.1 estão apresentados os haplótipos que foram identificados para as regiões reguladoras na amostra avaliada. Os valores de sHLA-G observados foram, portanto, resultantes da combinação dos dois haplótipos. Para que a ação de cada haplótipo na expressão de HLA-G fosse avaliada separadamente, seria necessário realizar a clonagem dos haplótipos de cada uma das regiões reguladoras, o que se torna inviável economicamente quando uma amostra populacional está sendo analisada, como é o caso da avaliada neste trabalho. Esse tipo de abordagem foi realizado em outros tipos de estudos, os quais também se propuseram a avaliar as diferenças de expressão da proteína considerando os haplótipos de regiões reguladoras de *HLA-G*, como é o caso dos artigos publicados por Martelli-Palomino e colaboradores (2014), Poras e colaboradores (2017) e Dias e colaboradores (2017).

No primeiro dos trabalhos acima citados, de Martelli-Palomino e colaboradores (2014), os haplótipos mais frequentes da região reguladora 3'NT foram analisados. Nele, foi verificada associação da UTR-1 com expressão de sHLA-G mais alta, enquanto que as UTR-5 e UTR-7 foram associadas com níveis mais baixos de expressão e as demais UTR avaliadas (UTR-2, UTR-3, UTR-4 e UTR-6) com níveis intermediários de sHLA-G. Dados esses corroborados por Poras e colaboradores (2017), que descreveram em seu artigo experimentos em que os haplótipos da região 3'NT foram clonados e inseridos em linhagens celulares que expressam HLA-G (JEG-3 - coriocarcinoma e FON - melanoma) e linhagens que não expressam HLA-G (M8 - melanoma e U251MG - glioblastoma) para que tivessem seus padrões de expressão avaliados através da atividade da luciferase. Da mesma forma, as UTR-5 e UTR-7 são associadas a baixos níveis de HLA-G, a UTR-1 com expressão mais alta e as UTR-2, UTR-3, UTR-4 e UTR-18 a níveis intermediários da proteína. Apesar dos artigos supracitados terem identificado esse padrão de expressão entre as diferentes UTR, com os dados obtidos nesta tese não foi possível verificar essas variações descritas pelos dois autores (FIGURA 4.4.5 e FIGURA

4.4.6). Não tendo sido, portanto, observadas diferenças significativas entre as diferentes combinações de haplótipos.

O terceiro trabalho citado, de Dias e colaboradores (2018), avaliou os principais haplótipos da região promotora-convencional. Nele, os haplótipos foram clonados e inseridos em duas linhagens de celulares: JEG-3 (coriocarcinoma) e FON+ (melanoma), ambas produtoras de HLA-G. Os haplótipos G010101a (em JEG-3) e G010101b (em FON+) exibiram maior atividade promotora, e consequente maior expressão de HLA-G; enquanto os haplótipos G010101d (em JEG-3) e G010102a (em FON+) exibiram menor atividade promotora. No presente trabalho, a região promotora-proximal também foi avaliada, mas não foram observadas diferenças significativas entre as diferentes combinações de haplótipos nem quando as avaliações foram feitas considerando somente os grupos de haplótipos (FIGURA 4.4.3 e FIGURA 4.4.4). Da mesma forma que na região 3'NT, percebe-se a necessidade de ampliação do número amostral, para que cada combinação de haplótipos tenha uma maior representatividade nas análises, e assim as diferenças de expressão se tornem mais perceptíveis. Ainda relacionado à região promotora do gene, nesta tese, a região promotora-distal foi avaliada pela primeira vez quanto à variação da expressão da proteína em uma amostra populacional saudável. Assim como as outras duas regiões, não foram verificadas diferenças de expressão de HLA-G entre as combinações de haplótipos (FIGURA 4.4.1 E FIGURA 4.4.2).

Dentre as possíveis causas para que não tenham sido observadas diferenças de expressão entre os haplótipos das regiões reguladoras, está o número amostral, pois segregando as amostras organizadas de acordo com os dois haplótipos presentes, cada grupo amostral (exemplo, UTR-1+UTR-2, UTR-2+UTR-3, ...) ficou com um pequeno número de indivíduos para que as comparações fossem realizadas. Seria importante verificar se ampliando o número amostral tornaria possível perceber diferenças de expressão entre as combinações de haplótipos. Outra possível causa, seria que os diferentes padrões de expressão das UTR não seriam perceptíveis pelo fato de estarem sendo avaliadas aos pares (pois, foram considerados os dois haplótipos de cada indivíduo). Nesse caso, considerando que foi verificada a existência de seleção balanceadora nesses segmentos do gene *HLA-G*, seria esperado, realmente, que não houvesse diferenças significativas entre as combinações de haplótipos em indivíduos saudáveis, como é o caso da população avaliada. Pois, o que se supõe é que a maioria da população seria heterozigota (tendo um haplótipo de alta e um de baixa expressão) e, em uma situação normal (não-patológica), as pessoas teriam uma expressão semelhante, não sendo possível identificar as diferenças na expressão da proteína. Para avaliar melhor essas diferenças em

uma amostra saudável, teríamos que, em uma situação hipotética, ter somente indivíduos homozigotos para realizar as comparações (ou pelo menos, um alto número amostral dessas combinações de haplótipos). De um modo geral, avaliando criticamente os dados obtidos neste trabalho, percebe-se a necessidade de ampliação do número amostral, para que mesmo as pequenas diferenças na produção da proteína HLA-G sejam detectadas. Além disso, foi possível observar que a técnica de ELISA, atualmente utilizada e disponível para a detecção de sHLA-G, precisa ser reavaliada com o intuito de aumentar a sua sensibilidade, para que assim pequenas diferenças de expressão sejam constatadas com maior precisão e acurácia. Rever esses pontos críticos, possivelmente, possibilitará mensurar os reais efeitos das UTR na expressão do gene *HLA-G*.

6. CONCLUSÕES

A revisão sistemática seguida de meta-análise acerca da variabilidade dos 14 pares de bases na região 3'NT do gene *HLA-G* revelou que, embora esse sítio seja o mais estudado em doenças de várias etiologias, ele não é um marcador genético eficiente para estudos de associação de doenças. Os níveis plasmáticos de sHLA-G e a expressão tecidual de HLA-G estavam mais elevados em pacientes com doença de Chagas nas suas diferentes formas de apresentação quando comparados com respectivos controles, entretanto, não observamos associação entre os sítios polimórficos das regiões promotora e 3'NT do gene *HLA-G* com a magnitude de expressão plasmática ou tecidual da proteína. Nas diversas populações mundiais, os alelos mais frequentes observados nas regiões reguladoras do gene *HLA-G* apresentam-se com frequências similares e verificou-se indícios de seleção balanceadora atuando nesses segmentos gênicos. Além disso, foi verificado um alto desequilíbrio de ligação entre os pares de SNP que compreendem toda essa região. Na população brasileira do Nordeste do Estado de São Paulo, não observamos associação entre os níveis de sHLA-G com a variabilidade das regiões reguladoras do gene. Embora não existam estudos clínicos utilizando HLA-G recombinante ou anticorpos anti-HLA-G, a proteína apresenta grande espectro de inibição de diversas células do sistema imune, tendo potencial aplicação clínica e o conhecimento da diversidade das regiões reguladoras pode contribuir para o entendimento acerca da magnitude de expressão do gene.

BIBLIOGRAFIA

- AFKHAMI, F.; KHANIANI, M. S.; FARZADI, L.; PAKNEJAD, Z.; DERAKHSHAN, S. M. The HLA-G 14-bp Insertion / Deletion Polymorphism in Recurrent Spontaneous Abortion among Iranian Women. *Iran J Allergy Asthma Immunol*, v. 13, n. 5, p. 364–369, 2014.
- AGHDAIE, M.; AZARPIRA, N.; KAZEMI, K.; GERAMIZADEH, B.; DARAI, M.; MALEKHOSEINI, S. Frequency of HLA-G exon 8 polymorphisms and kidney allograft outcome in Iranian population. *Mol Biol Rep*, v. 38, n. 5, p. 3593–3597, 2011.
- AGNIHOTRI, V.; GUPTA, A.; KUMAR, R.; DATT, A.; DWIVEDI, S. Promising link of HLA-G polymorphism, tobacco consumption and risk of Head and Neck Squamous Cell Carcinoma (HNSCC) in North Indian population. *Human Immunology*, v. 78, n. 2, p. 172–178, 2017.
- AKHTER, A.; FARIDI, R. M.; DAS, V.; PANDEY, A.; NAIK, S.; AGRAWAL, S. In vitro up-regulation of HLA-G using dexamethasone and hydrocortisone in first-trimester trophoblast cells of women experiencing recurrent miscarriage. *Tissue Antigens*, v. 80, n. 2, p. 126–135, 2012.
- ALBUQUERQUE, R. S. De; MENDES-JUNIOR, C. T.; LUCENA-SILVA, N.; LEAL, C.; MEIRE, D.; VEIGA-CASTELLI, L. C.; FOSS-FREITAS, M. C.; CÉSAR, M.; HASSAN, N.; DEGHAIDE, S.; MOREAU, P.; GREGORI, S.; CASTELLI, E. C.; ANTÔNIO, E. Association of HLA-G 3' untranslated region variants with type 1 diabetes mellitus. *Human Immunology*, v. 77, n. 4, p. 358–364, 2016.
- ALVAREZ M, PIEDADE J, BALSEIRO S, RIBAS G, REGATEIRO F. HLA-G 3'-UTR SNP and 14-bp deletion polymorphisms in Portuguese and Guinea-Bissau populations. *Int J Immunogenet*. Dez;36(6):361-6, 2009.
- AMODIO, G.; CANTI, V.; MAGGIO, L.; ROSA, S.; TERESA, M.; ROVERE-QUERINI, P.; GREGORI, S. Association of genetic variants in the 3' UTR of HLA-G with Recurrent Pregnancy Loss. *Human Immunology*, v. 77, n. 10, p. 886–891, 2016.
- ARJMAND, F.; SAMADI, M. Association of 14-bp insertion/deletion polymorphism of HLA-G gene with idiopathic recurrent miscarriages in infertility center patients in Yazd, Iran. *Journal of Immunotoxicology*, v. 0, n. 0, p. 1–6, 2015.
- ARUNA, M.; SIRISHA, P. V. S.; BHASKAR, S. A.; TARAKESWARI, S.; THANGARAJ, K.; REDDY, B. M. Role of 14-bp insertion/deletion polymorphism in HLA-G among Indian women with recurrent spontaneous abortions. *Tissue Antigens*, v. 77, n. 2, p. 131–135, 2011.
- AZARPIRA, N.; AGHDAIE, M. H.; KAZEMI, K.; GERAMIZADEH, B.; DARAI, M. HLA-G polymorphism (rs16375) and acute rejection in liver transplant recipients. *Dis Markers*, v. 2014, p. 1–5, 2014.
- BANDELT H, FORSTER P, RÖHL A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16(1):37–48, 1999.
- BARRETT JC, FRY B, MALLER J, DALY MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, Jan 15, 2005.
- BEN FREDJ, N.; SAKLY, K.; BORTOLOTTI, D.; AISSI, M.; FRIH-AYED, M.; ROTOLA, A.; CASELLI, E.; CURA, F.; SAKLY, N.; AOUNI, M.; DI, D.; RIZZO, R. The association between functional HLA-G 14bp insertion/deletion and +3142 C>G polymorphisms and susceptibility to multiple sclerosis. *Immunology Letters*, v. 180, p. 24–30, 2016.

- BERGER, D. S.; HOGGE, W. A.; BARMADA, M. M.; FERRELL, R. E. Comprehensive analysis of HLA-G: implications for recurrent spontaneous abortion. *Reprod Sci*, v. 17, n. 4, p. 331–338, 2010.
- BIELSKA, M.; BOJO, M.; KLIMKIEWICZ-WOJCIECHOWSKA, G.; JESIONEK-KUPNICKA, D.; BOROWIEC, M.; KALINKA-WARZOCZA, E.; PROCHOREC-SOBIESZEK, M.; ROBAK, T.; WARZOCZA, K.; MŁYNNARSKI, W.; LECH-MARANDA, E. Human leukocyte antigen-G polymorphisms influence the clinical outcome in diffuse large B-cell lymphoma. *Genes Chromosomes Cancer*, v. 54, n. 3, p. 185–193, 2015.
- BORTOLOTTI, D.; GENTILI, V.; ROTOLA, A.; LUCA, D. Di; RIZZO, R. Implication of HLA-G 3' untranslated region polymorphisms in human papillomavirus infection. *Tissue Antigens*, v. 83, n. 2, p. 113–118, 2014.
- BOUKOUACI, W.; BUSSON, M.; FORTIER, C.; AMOKRANE, K.; DE LATOUR, R.; ROBIN, M.; KRISHNAMOORTHY, R.; TOUBERT, A.; CHARRON, D.; SOCIÉ, G.; TAMOUZA, R. Association of HLA-G low expressor genotype with severe acute graft-versus-host disease after sibling bone marrow transplantation. *Front Immunol*, v. 2, n. 74, p. 1–6, 2011.
- BRIEN, M. O.; MCCARTHY, T.; JENKINS, D.; PAUL, P.; DAUSSET, J.; CAROSELLA, E. D.; MOREAU, P. Altered HLA-G transcription in pre-eclampsia is associated with allele specific inheritance: possible role of the HLA-G gene in susceptibility to the disease. *Cell Mol Life Sci*, v. 58, n. 12–13, p. 1943–1949, 2001.
- CAPITTINI C, BERGAMASCHI P, SACHETTO S, TRUGLIO M, VIOLA M, MARCHESI A, GENOVESE V, ROMANO B, GUARENE M, POMA R, MARTINETTI M, TINELLI C, SALVANESCHI L. The plasma levels of soluble HLA-G molecules correlate directly with CD34+ cell concentration and HLA-G 14bp insertion/deletion polymorphism in cord blood donors. *Blood Transfus. Jan;12 Suppl 1:s361-6*, 2014.
- CARLINI F, TRAORE K, CHEROUAT N, ROUBERTOUX P, BUHLER S, CORTEY M, SIMON S, DOUMBO O, CHIARONI J, PICARD C, DI CRISTOFARO J. HLA-G UTR haplotype conservation in the Malian population: association with soluble HLA-G. *PLoS One*. Dec 23;8(12), 2013.
- CAROSELLA, E. D.; ROUAS-FREISS, N.; ROUX, D. T.; MOREAU, P.; LEMAOUT, J. HLA-G : An Immune Checkpoint Molecule. 1. ed. Elsevier Inc., 2015.
- CASTELLI, E. C.; MOREAU, P.; CHIROMATZO, A. O. E.; MENDES-JUNIOR, C. T.; VEIGA-CASTELLI, L. C.; YAGHI, L.; GIULIATTI, S.; CAROSELLA, E. D.; DONADI, E. A. In silico analysis of microRNAs targeting the HLA-G 3' untranslated region alleles and haplotypes. *Human Immunology*, v. 70, n. 12, p. 1020–1025, 2009.
- CASTELLI, E. C.; MENDES-JUNIOR, C. T.; VEIGA-CASTELLI, L. C.; ROGER, M.; MOREAU, P.; DONADI, E. A. A comprehensive study of polymorphic sites along the HLA-G gene: Implication for gene regulation and evolution. *Molecular Biology and Evolution*, v. 28, n. 11, p. 3069–3086, 2011.
- CASTELLI, E. C.; VEIGA-CASTELLI, L. C.; YAGHI, L.; MOREAU, P.; DONADI, E. a. Transcriptional and posttranscriptional regulations of the HLA-G gene. *Journal of Immunology Research*, v. 2014, p. 1–14, 2014.
- CASTELLI EC, GERASIMOU P, PAZ MA, RAMALHO J, PORTO IOP, LIMA THA, SOUZA AS, VEIGA-CASTELLI LC, COLLARES CVA, DONADI EA, MENDES-JUNIOR CT, COSTEAS P. HLA-G variability and haplotypes detected by massively parallel sequencing procedures in the geographically distinct population samples of Brazil and Cyprus. *Mol Immunol*. Mar; 83:115-126, 2017.

- CASTELLI EC, PAZ MA, SOUZA AS, RAMALHO J, MENDES-JUNIOR CT. Hla-mapper: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. *Hum Immunol*. Setembro;79(9):678-684, 2018.
- CATAMO, E.; ADDOBBATI, C.; SEGAT, L.; FRAGOSO, T. S.; BARBOSA, A. D.; JR, R.; DUARTE, A. L. B. P.; MONASTA, L.; P, S.-G.; S, C. HLA-G gene polymorphisms associated with susceptibility to rheumatoid arthritis disease and its severity in Brazilian patients. *Tissue Antigens*, v. 84, n. 3, p. 308–315, 2014.
- CATAMO, E.; ADDOBBATI, C.; SEGAT, L.; FRAGOSO, T. S.; DANTAS, A. T.; MARIZ, H. D. A. Comprehensive analysis of polymorphisms in the HLA-G 5' upstream regulatory and 3' untranslated regions in Brazilian patients with systemic lupus erythematosus. *Tissue Antigens*, v. 85, n. 6, p. 458–465, 2015a.
- CATAMO, E.; ADDOBBATI, C.; SEGAT, L.; SOTERO FRAGOSO, T.; TAVARES DANTAS, A.; DE ATAIDE MARIZ, H.; FERREIRA DA ROCHA JUNIOR, L.; BRANCO PINTO DUARTE, A.; COELHO, A.; DE MOURA, R.; POLESELLO, V.; CROVELLA, S.; SANDRIN GARCIA, P. Comprehensive analysis of polymorphisms in the HLA-G 5' upstream regulatory and 3' untranslated regions in Brazilian patients with systemic lupus erythematosus. *Tissue Antigens*, v. 85, n. 6, p. 458–465, 2015b.
- CATAMO, E.; ZUPIN, L.; SEGAT, L.; CELSI, F.; CROVELLA, S. HLA-G and susceptibility to develop celiac disease. *Human Immunology*, v. 76, n. 1, p. 36–41, 2015c.
- CHEN, Y.; GAO, X.; DENG, Y.; ZHANG, H. Relationship between HLA-G gene polymorphism and the susceptibility of esophageal cancer in Kazakh and Han nationality in Xinjiang. *Biomarkers*, v. 17, n. May 2011, p. 9–15, 2012.
- CHIUSOLO, P.; BELLES, S.; PICCIRILLO, N.; GIAMMARCO, S.; MARIETTI, S.; RITIS, D. De; METAFUNI, E.; STIGNANI, M.; BARICORDI, O. R.; SICA, S.; LEONE, G.; RIZZO, R. The role of HLA – G 14-bp polymorphism in allo-HSCT after short-term course MTX for GvHD prophylaxis. *Bone Marrow Transplantation*, v. 47, n. 1, p. 120–124, 2011.
- CHRISTIANSEN, O. B.; KOLTE, A. M.; DAHL, M.; LARSEN, E. C.; STEFFENSEN, R.; NIELSEN, H. S.; HVIID, T. V. Maternal homozygosity for a 14 base pair insertion in exon 8 of the HLA-G gene and carriage of HLA class II alleles restricting HY immunity predispose to unexplained secondary recurrent miscarriage and low birth weight in children born to these patients. *Human Immunology*, v. 73, n. 7, p. 699–705, 2012.
- CILIÃO ALVES, D. C.; DE OLIVEIRA CRISPIM, J. C.; CASTELLI, E. C.; MENDES-JUNIOR, C. T.; DEGHAIDE, N. H. S.; BARROS SILVA, G. E.; COSTA, R. S.; SABER, L. T.; MOREAU, P.; DONADI, E. A. Human leukocyte antigen-G 3' untranslated region polymorphisms are associated with better kidney allograft acceptance. *Human immunology*, v. 73, n. 1, p. 52–9, jan. 2012.
- COELHO, A. V. C.; MOURA, R. R.; CROVELLA, S.; CELSI, F. HLA-G genetic variants and hepatocellular carcinoma: a meta-analysis. *Genet Mol Res*, v. 15, n. 3, p. 1–9, 2016.
- CONSIGLIO, C. R.; VEIT, T. D.; MONTICIELO, O. A.; MUCENIC, T.; XAVIER, R. M.; BRENOL, J. C. T.; CHIES, J. A. Association of the HLA-G gene +3142C>G polymorphism with systemic lupus erythematosus. *Tissue Antigens*, v. 77, n. 6, p. 540–545, 2011.

CRISPIM, J.; MENDES-JUNIOR, C.; WASTOWSKI, I.; COSTA, R.; CASTELLI, E.; SABER, L.; DONADI, E. Frequency of insertion/deletion polymorphism in exon 8 of HLA-G and kidney allograft outcome. *Tissue antigens*, v. 71, n. 1, p. 35–41, jan. 2008.

DAHL M, KLITKOU L, CHRISTIANSEN OB, DJURISIC S, PIOSIK ZM, SKOVBO P, MØLLER AM, STEFFENSEN R, HVIID TV. Human leukocyte antigen (HLA)-G during pregnancy part II: associations between maternal and fetal HLA-G genotypes and soluble HLA-G. *Hum Immunol*. Abril;76(4):260-71, 2015.

DARDANO, A.; RIZZO, R.; POLINI, A.; STIGNANI, M.; TOGNINI, S.; PASQUALETTO, G.; URSINO, S.; COLATO, C.; FERDEGHINI, M.; BARICORDI, O. R.; MONZANI, F. Soluble human leukocyte antigen-g and its insertion/deletion polymorphism in papillary thyroid carcinoma: novel potential biomarkers of disease? *J Clin Endocrinol Metab*, v. 97, n. November 2012, p. 4080–4086, 2016.

OLIVEIRA MLG, VEIGA-CASTELLI LC, MARCORIN L, DEBORTOLI G, PEREIRA ALE, FRACASSO NCA, SILVA GDV, SOUZA AS, MASSARO JD, SIMÕES AL, SABBAGH A, DONADI EA, CASTELLI EC, MENDES-JUNIOR CT. Extended HLA-G genetic diversity and ancestry composition in a Brazilian admixed population sample: Implications for HLA-G transcriptional control and for case-control association studies. *Hum Immunol*. Nov;79(11):790-799, 2018.

DIAS FC, CASTELLI EC, COLLARES CV, MOREAU P, DONADI EA. The Role of HLA-G Molecule and HLA-G Gene Polymorphisms in Tumors, Viral Hepatitis, and Parasitic Diseases. *Front Immunol*. Feb 2;6:9, 2015.

DIAS FC, BERTOL BC, PORAS I, SOUTO BM, MENDES-JUNIOR CT, CASTELLI EC, GINEAU L, SABBAGH A, ROUAS-FREISS N, CAROSELLA ED, DONADI EA, MOREAU P. The genetic diversity within the 1.4 kb HLA-G 5' upstream regulatory region moderately impacts on cellular microenvironment responses. *Sci Rep*. Abril 4;8(1):5652, 2018.

DI CRISTOFARO J, EL MOUJALLY D, AGNEL A, MAZIÈRES S, CORTEY M, BASIRE A, CHIARONI J, PICARD C. HLA-G haplotype structure shows good conservation between different populations and good correlation with high, normal and low soluble HLA-G expression. *Hum Immunol*. Feb;74(2):203-6, 2013.

DONADI, E. a; CASTELLI, E. C.; ARNAIZ-VILLENA, A.; ROGER, M.; REY, D.; MOREAU, P. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cellular and molecular life sciences: CMLS*, v. 68, n. 3, p. 369–95, fev. 2011.

ENGHELABIFAR, M.; ALLAFAN, S.; KHAYATZADEH, J.; ABADI, K. S.; NAZARABADI, M. H.; MORADI, F.; MUSAVIFAR, N.; JALALI, M.; MOJARRAD, M. Association of the maternal 14-bp insertion/deletion polymorphism in the histocompatibility leukocyte antigen G gene with recurrent implantation failure. *Iran J Reprod Med*, v. 12, n. 9, p. 641–646, 2014.

ESKANDARI-NASAB, E.; HASHEMI, M.; HASANI, S.; OMRANI, M.; TAHERI, M.; MASHHADI, M. Association between HLA-G 3'NTR 14-bp ins/del polymorphism and susceptibility to breast cancer. *Cancer Biomark*, v. 13, n. 4, p. 253–259, 2013.

EXCOFFIER, L, AND LISCHER, H.E.L. Arlequin Suite ver 3.5, a New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567, 2010.

FABRIS, A.; CATAMO, E.; SEGAT, L.; MORGUTTI, M.; CLAUDIO, L.; CROVELLA, S. Association between HLA-G 3'NTR 14-bp polymorphism and HIV vertical transmission in Brazilian children. *AIDS*, v. 23, n. 2, p. 177–82, 2009.

FABRIS, A.; SEGAT, L.; CATAMO, E.; MORGUTTI, M.; VENDRAMIN, A.; CROVELLA, S. HLA-G 14 bp deletion/insertion polymorphism in celiac disease. *The American journal of gastroenterology*, v. 106, n. 1, p. 139–44, jan. 2011.

FAN, W.; HUANG, Z.; LI, S.; XIAO, Z. The HLA-G 14-bp polymorphism and recurrent implantation failure: a meta-analysis. *J Assist Reprod Genet*, v. 34, n. 11, p. 1559–1565, 2017.

FAN, W.; LI, S.; HUANG, Z.; CHEN, Q. Relationship between HLA-G polymorphism and susceptibility to recurrent miscarriage: a meta-analysis of non-family-based studies. *J Assist Reprod Genet*, v. 31, n. 2, p. 173–184, 2014.

FERGUSON, R.; RAMANAKUMAR, A. V.; KOUSHIK, A.; FRANCO, E.; ROGER, M. Human leukocyte antigen G polymorphism is associated with an increased risk of invasive cancer of the uterine cervix. *Int J Cancer*, v. 131, n. 3, p. 312–319, 2012.

FERREIRA LM, MEISSNER TB, MIKKELSEN TS, MALLARD W, O'DONNELL CW, TILBURGS T, GOMES HA, CAMAHORT R, SHERWOOD RI, GIFFORD DK, RINN JL, COWAN CA, STROMINGER JL. A distant trophoblast-specific enhancer controls HLA-G expression at the maternal-fetal interface. *Proc Natl Acad Sci USA*. Mai 10;113(19):5364-9, 2016.

FERREIRA, C.; GAMA, S.; CHACHÁ, F.; FERNANDES, F.; CORRÊA, A.; CARVALHO, R. De; HASSAN, N.; DEGHAIDE, S.; RODRIGUES, S.; MARANO, L. A.; MENDES-JUNIOR, C. T.; NAIRA, L.; RAMALHO, Z.; ZUCOLOTO, S.; ANTÔNIO, E.; LOURDES, A. De; MARTINELLI, C. The HLA-G 14-base pair deletion allele and the deletion / deletion genotype are associated with persistent HBe antigenemia in chronic hepatitis B infection.

DE FIGUEIREDO-FEITOSA NL, MARTELLI PALOMINO G, CILIÃO ALVES DC, MENDES JUNIOR CT, DONADI EA, MACIEL LM. HLA-G 3' untranslated region polymorphic sites associated with increased HLA-G production are more frequent in patients exhibiting differentiated thyroid tumors. *Clin Endocrinol*. Abril;86(4):597-605, 2017.

GARCÍA-GONZÁLEZ, I. J.; VALLE, Y.; RIVAS, F.; FIGUERA-VILLANUEVA, L. E.; MUÑOZ-VALLE, J. F.; FLORES-SALINAS, H. E.; GUTIÉRREZ-AMAVIZCA, B. E.; DÁVALOS-RODRÍGUEZ, N. O.; PADILLA-GUTIÉRREZ, J. R. The 14 bp Del/Ins HLA-G polymorphism is related with high blood pressure in acute coronary syndrome and type 2 diabetes mellitus. *Biomed Res Int*, v. 2014, p. 1–8, 2014.

GARZIERA, M.; CATAMO, E.; CROVELLA, S.; MONTICO, M.; CECCHIN, E.; LONARDI, S.; MINI, E.; NOBILI, S.; ROMANATO, L.; TOFFOLI, G. Association of the HLA-G 3'NTR polymorphisms with colorectal cancer in Italy: a first insight. *Int J Immunogenet*, v. 43, n. 1, p. 32–39, 2016.

GAZIT, E.; SLOMOV, Y.; GOLDBERG, I.; BRENNER, S.; LOEWENTHAL, R. HLA-G is Associated with Pemphigus Vulgaris in Jewish Patients. *Hum Immunol*, v. 65, n. 1, p. 39–46, 2004.

GE, Y.; GE, Q.; LI, M.; SHI, G.; XU, X.; XU, L.; XU, Z.; LU, T.; WU, R.; ZHOU, L.; WU, J.; LIANG, K.; DOU, Q.; ZHU, J.; LI, W.; JIA, R. Association between human leukocyte antigen - G 14-bp insertion / deletion polymorphism and cancer risk : A meta-analysis and systematic review. *Human Immunology*, v. 75, n. 8, p. 827–832, 2014.

GENRE, J.; REGINALDO, F. P. S.; ANDRADE, J. M. D. L.; LIMA, F. P.; COUTINHO, A. V.; DONADI, E. A.; CRISPIM, J. C. HLA-G 14-bp Ins/Ins Genotype in Patients Harboring *Helicobacter pylori* Infection: A Potential Risk Factor? *Scand J Immunol*, v. 83, n. 1, p. 52–57, 2016.

GERAGHTY, D. E.; KOLLER, B. H.; ORR, H. T. A human major histocompatibility complex class I gene that encodes a protein with a shortened cytoplasmic segment. *Proc Natl Acad Sci USA*, v. 84, n. December, p. 9145–9149, 1987.

GERASIMOU, P.; SKORDIS, N.; PICOLOS, M.; SPYRIDONIDIS, A.; COSTEAS, P. HLA-G 14-bp polymorphism affects the age of onset in Type I Diabetes Mellitus. *Int J Immunogenet*, v. 43, n. 3, p. 135–142, 2016.

GINEAU L, LUISI P, CASTELLI EC, MILET J, COURTIN D, CAGNIN N, PATILLON B, LAAYOUNI H, MOREAU P, DONADI EA, GARCIA A, SABBAGH A. Balancing immunity and tolerance: genetic footprint of natural selection in the transcriptional regulatory region of HLA-G. *Genes Immun*. Jan-Feb;16(1):57-70, 2015.

GLAS, J.; TÖRÖK, H.; TONENCHI, L.; WETZKE, M.; BEYNON, V.; TESHOME, M.; COTOFANA, S.; SCHIEMANN, U.; GRIGA, T.; KLEIN, W.; EPPLEN, J.; FOLWACZNY, C.; FOLWACZNY, M.; MUSSACK, T.; WEISS, E.; WEISS, H. The 14-bp deletion polymorphism in the HLA-G gene displays significant differences between ulcerative colitis and Crohn's disease and is associated with ileocecal resection in Crohn's disease. *Int Immunol*, v. 19, n. 5, p. 621–626, 2007.

GRAEBIN, P.; VEIT, T. D.; ALHO, C. S.; DIAS, F. S.; CHIES, J. A. B. Polymorphic variants in exon 8 at the 3' UTR of the HLA-G gene are associated with septic shock in critically ill patients. *Critical Care*, v. 16, n. 5, p. R211, 2012.

HACHIYA, Y.; KAWASAKI, A.; OKA, S.; KONDO, Y.; ITO, S.; MATSUMOTO, I.; KUSAOI, M.; AMANO, H.; SUDA, A.; SETOGUCHI, K.; NAGAI, T.; SHIMADA, K.; SUGII, S.; OKAMOTO, A.; CHIBA, N.; SUEMATSU, E.; OHNO, S.; KATAYAMA, M.; KONO, H.; HIROHATA, S.; TAKASAKI, Y.; HASHIMOTO, H.; SUMIDA, T.; NAGAOKA, S.; TOHMA, S.; FURUKAWA, H.; TSUCHIYA, N. Association of HLA-G 3' Untranslated Region Polymorphisms with Systemic Lupus Erythematosus in a Japanese Population: A Case-Control Association Study. *PLoS One*, v. 11, n. 6, p. 1–14, 2016.

HADDAD, R.; CILIÃO ALVES, D.; ROCHA-JUNIOR, M.; AZEVEDO, R.; POMBO-DE-OLIVEIRA, M.; TAKAYANAGUI, O.; DONADI, E.; COVAS, D.; KASHIMA, S. HLA-G 14-bp insertion/deletion polymorphism is a risk factor for HTLV-1 infection. *AIDS Res Hum Retroviruses*, v. 27, n. 3, p. 283–288, 2011.

HAGHI, M.; ALI, M.; FEIZI, H.; SADEGHIZADEH, M.; LOTFI, S. 14-bp Insertion/Deletion Polymorphism of the HLA-G gene in Breast Cancer among Women from North Western Iran. *Asian Pac J Cancer Prev*, v. 16, n. 14, p. 6155–6158, 2015.

HAHN EC, ZAMBRA FMB, KAMADA AJ, DELONGUI F, GRION CMC, REICHE EMV, CHIES JAB. Association of HLA-G 3'UTR polymorphisms and haplotypes with severe sepsis in a Brazilian population. *Hum Immunol*. Nov;78(11-12):718-723, 2017.

HARRISON, G. A.; HUMPHREY, K. E.; JAKOBSEN, I. B.; COOPER, D. W. A 14 bp deletion polymorphism in the HLA-G gene. *Human Molecular Genetics*, v. 2, n. 12, p. 2200, 1993.

- HASHEMI, M.; SANDOUGHI, M.; FAZELI, S. A.; BAHARI, G.; REZAEI, M.; ZAKERI, Z. Evaluation of HLA-G 14 bp Ins/Del and +3142G>C Polymorphism with Susceptibility and Early Disease Activity in Rheumatoid Arthritis. *Adv Med*, v. 2016, p. 1–7, 2016.
- HONG, H.; PAXIMADIS, M.; GRAY, G.; KUHN, L.; TIEMESSEN, C. Maternal human leukocyte antigen-G (HLA-G) genetic variants associate with in utero mother-to-child transmission of HIV-1 in Black South Africans. *Infect Genet Evol*, v. 30, p. 147–158, 2015.
- HVIID, T.; HYLENIUS, S.; HOEGH, A. M.; KRUSE, C.; CHRISTIANSEN, O. B. HLA-G polymorphisms in couples with recurrent spontaneous abortions. *Tissue Antigens*, v. 60, n. 2, p. 122–132, 2002.
- HVIID, T.; HYLENIUS, S.; RØRBYE, C.; NIELSEN, L. HLA-G allelic variants are associated with differences in the HLA-G mRNA isoform profile and HLA-G mRNA levels. *Immunogenetics*, v. 55, n. 2, p. 63–79, 2003.
- HVIID, T.; HYLENIUS, S.; LINDHARD, A.; CHRISTIANSEN, O. B. Association between human leukocyte antigen-G genotype and success of in vitro fertilization and pregnancy outcome. *Tissue Antigens*, v. 64, n. 1, p. 66–69, 2004.
- HVIID, T. V. F.; RIZZO, R.; MELCHIORRI, L.; STIGNANI, M.; BARICORDI, O. R. Polymorphism in the 5' upstream regulatory and 3' untranslated regions of the HLA-G gene in relation to soluble HLA-G and IL-10 expression. *Hum Immunol*, v. 67, n. 1–2, p. 53–62, 2006.
- HYLENIUS, S.; ANDERSEN, A. N.; MELBYE, M.; HVIID, T. V. F. Association between HLA-G genotype and risk of pre-eclampsia: a case-control study using family triads. *Mol Hum Reprod*, v. 10, n. 4, p. 237–246, 2004.
- IVERSEN, A.; TOAI, O.; NGUYEN, D.; FØLL, L.; POLIAKOVA, I.; MALM, V.; ACAR, N.; MYHRE, R.; KLUNGLAND, H.; AUSTGULEN, R. The HLA-G 14bp gene polymorphism and decidual HLA-G 14bp gene expression in pre-eclamptic and normal pregnancies. *J Reprod Immunol*, v. 78, n. 2, p. 158–165, 2008.
- JAHAN, P.; DEEPTHI, G.; LATHA, P.; RANI, V. U. Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health A study on the role of HLA-G 14 bp and ACE IN / DEL polymorphisms in pre-eclamptic South Indian women. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, v. 4, n. 2, p. 164–169, 2014.
- JASSEM, R. M.; SHANI, W. S.; LOISEL, D. A.; SHARIEF, M. HLA-G polymorphisms and soluble HLA-G protein levels in women with recurrent pregnancy loss from Basrah province in Iraq. *Hum Immunol*, v. 73, n. 8, p. 811–817, 2012.
- JEONG, K.; KIM, S.; KANG, B.; CHUNG, J.; SHIN, M.; LEE, M. Association between an HLA-G 14 bp insertion/deletion polymorphism and non-segmental vitiligo in the Korean population. *Arch Dermatol Res*, v. 306, n. 6, p. 577–582, 2014a.
- JEONG, S.; PARK, S.; PARK, B.; PARK, Y.; KWON, O.; KIM, H. Human leukocyte antigen-G (HLA-G) polymorphism and expression in breast cancer patients. *Plos One*, v. 9, n. 5, p. e98284., 2014b.
- JIANG, Y.; CHEN, S.; JIA, S.; ZHU, Z.; GAO, X. Association of HLA-G 3' UTR 14-bp insertion/deletion polymorphism with hepatocellular carcinoma susceptibility in a Chinese population. *DNA Cell Biol*, v. 30, n. 12, p. 1027–1032, 2011.

- JIN, Z.; XU, C.; TIAN, P.; XUE, W.; DING, X.; ZHENG, J. Impact of HLA-G 14-bp polymorphism on acute rejection and cytomegalovirus infection in kidney transplant recipients from northwestern China. *Transplant Immunology*, v. 27, n. 2–3, p. 69–74, 2012.
- KIM, S. K.; JEONG, K. H.; KANG, I. J.; CHUNG, J. H.; SHIN, M. K.; LEE, M. H. Relationship between the HLA-G 14 bp insertion/deletion polymorphism and susceptibility to autoimmune disease: a meta-analysis. *Genet Mol Res*, v. 14, n. 4, p. 15839–15847, 2015.
- KRONER, A.; GRIMM, A.; JOHANNSEN, K.; MÄURER, M.; WIENDL, H. The genetic influence of the nonclassical MHC molecule HLA-G on multiple sclerosis. *Human immunology*, v. 68, n. 5, p. 422–5, maio 2007.
- LARSEN, M.; HYLENIUS, S.; ANDERSEN, A.; HVIID, T. The 3'-untranslated region of the HLA-G gene in relation to pre-eclampsia: revisited. *Tissue antigens*, v. 75, n. 3, p. 253–61, mar. 2010.
- LASHLEY, L. E. E. L. O.; WESTERLAKEN, L. A. J. Van Der; HAASNOOT, G. W.; DRABBELS, J. J. M. Maternal HLA-C2 and 14 bp insertion in HLA-G is associated with recurrent implantation failure after in vitro fertilization treatment. *Tissue Antigens*, v. 84, n. 6, p. 536–544, 2014.
- LAU, D. T.; NORRIS, M. D.; MARSHALL, G. M.; HABER, M.; ASHTON, L. J. HLA-G polymorphisms, genetic susceptibility, and clinical outcome in childhood neuroblastoma. *Tissue Antigens*, v. 78, n. 6, p. 421–427, 2011.
- LAZARTE, J.; GOLDRAICH, L.; MANLHIOT, C.; KOZUSZKO, S.; RAO, V.; DELGADO, D. Human Leukocyte Antigen-G Polymorphisms Association with Cancer Post-Heart Transplantation. *Human Immunology*, v. 77, n. 9, p. 805–811, 2016.
- LEE, Y.; BAE, S.; SONG, G. Meta-analysis of associations between functional HLA-G polymorphisms and susceptibility to systemic lupus erythematosus and rheumatoid arthritis. *Rheumatol Int*, v. 35, n. 6, p. 953–961, 2015.
- LI, T.; HUANG, H.; LIAO, D.; LING, H. Genetic polymorphism in HLA-G 3'NTR 14-bp ins/del and risk of cancer: a meta-analysis of case-control study. *Mol Genet Genomics*, v. 290, n. 4, p. 1235–1245, 2015.
- LIN, A.; YAN, W. H.; DAI, M. Z.; CHEN, X. J.; LI, B. L.; CHEN, B. G.; FAN, L. A. Maternal human leukocyte antigen-G polymorphism is not associated with pre-eclampsia in a Chinese Han population. *Tissue Antigens*, v. 68, n. 4, p. 311–316, 2006.
- LIN, A.; YAN, W.; XU, H.; TANG, L.; CHEN, X.; ZHU, M.; ZHOU, M. 14 bp deletion polymorphism in the HLA-G gene is a risk factor for idiopathic dilated cardiomyopathy in a Chinese Han population. *Tissue antigens*, v. 70, n. 5, p. 427–31, nov. 2007.
- LIU, Y.; LAI, M.; LOU, Y.; HAN, Q.; YANG, Q.; CHEN, M.; LI, J.; WANG, H.; YAN, W.; ZHENG, X. Elevation of plasma-soluble HLA-G in childhood nephrotic syndrome is associated with IgE. *Ann Clin Biochem*, v. 54, n. 1, p. 69–75, 2017.
- LUCENA-SILVA N, MONTEIRO AR, DE ALBUQUERQUE RS, GOMES RG, MENDES-JUNIOR CT, CASTELLI EC, DONADI EA. Haplotype frequencies based on eight polymorphic sites at the 3' untranslated region of the HLA-G gene in individuals from two different geographical regions of Brazil. *Tissue Antigens*. Abril; v.79, n.4, p. 272-278, 2012.

LUCENA-SILVA, N.; SALES, V.; SOUZA, B. De; GOMES, R. G.; COSTA, Y.; MUNIZ, N.; ALBUQUERQUE, R. S. De; LUNA, A.; MONTEIRO, R.; TADEU, G.; DINIZ, N.; ROSÂNGELA, M.; DUARTE, C.; MENDES-JUNIOR, C. T.; CASTELLI, C.; DONADI, E. A.; LUCENA-SILVA, N.; SALES, V.; SOUZA, B. De; GOMES, R. G.; FANTINATTI, A.; COSTA, Y.; MUNIZ, N.; ALBUQUERQUE, R. S. De; LUNA, A.; MONTEIRO, R.; TADEU, G.; DINIZ, N.; ROSÂNGELA, M.; DUARTE, C.; MENDES-JUNIOR, C. T.; CASTELLI, C.; DONADI, E. A. HLA-G 3' untranslated region polymorphisms are associated with systemic lupus erythematosus in 2 Brazilian populations. *J Rheumatol*, v. 40, n. 7, p. 1104–1113, 2013a.

LUCENA-SILVA, N.; TEIXEIRA, M. A. G.; RAMOS, A. D. L.; ALBUQUERQUE, R. S. De; DINIZ, G. T. N.; MENDES-JUNIOR, C. T.; CASTELLI, E. C.; DONADI, E. A. The +3187A/G HLA-G polymorphic site is associated with polar forms and reactive reaction in leprosy. *Mol Genet Genomic Med*, v. 1, n. 3, p. 123–130, 2013b.

MARIASELVAM, C. M.; CHAABEN, A. B.; SALAH, S.; CHARRON, D.; KRISHNAMOORTHY, R.; TAMOUZA, R. Human leukocyte antigen-G polymorphism influences the age of onset and autoantibody status in rheumatoid arthritis. *Tissue Antigens*, v. 85, n. 3, p. 182–189, 2015.

MARTELLI-PALOMINO, G.; PANCOTTO, J. A.; MUNIZ, Y. C.; MENDES-JUNIOR, C. T.; CASTELLI, E. C.; MASSARO, J. D.; KRAWICE-RADANNE, I.; PORAS, I.; REBMANN, V.; CAROSELLA, E. D.; ROUAS-FREISS, N.; MOREAU, P.; DONADI, E. A. Polymorphic Sites at the 3' Untranslated Region of the HLA-G Gene Are Associated with Differential HLA-G Soluble Levels in the Brazilian and French Population. *PLoS ONE*, v. 8, n. 10, p. 1–10, 2013.

MICHITA, R. T.; ZAMBRA, F. M. B.; FRAGA, L. R.; SANSEVERINO, M. T. V.; CALLEGARI-JACQUES, S. M.; VIANNA, P.; CHIES, J. A. B. A tug-of-war between tolerance and rejection – New evidence for 3' UTR HLA-G haplotypes influence in recurrent pregnancy loss. *Human Immunology*, v. 77, n. 10, p. 892–897, 2016.

MISRA, M. K.; PRAKASH, S.; KAPOOR, R.; PANDEY, S. K.; SHARMA, R. K.; AGRAWAL, S. Association of HLA-G promoter and 14-bp insertion-deletion variants with acute allograft rejection and end-stage renal disease. *Tissue Antigens*, v. 82, n. 5, p. 317–326, 2013.

MOHAMMADI, N.; ADIB, M.; ALSAHEBFOSOU, F.; KAZEMI, M.; ETEMADIFAR, M. An investigation into the association between HLA-G 14 bp insertion / deletion polymorphism and multiple sclerosis susceptibility. *Journal of Neuroimmunology*, v. 290, p. 115–118, 2016.

MORANDI, F.; RIZZO, R.; FAINARDI, E.; ROUAS-FREISS, N.; PISTOIA, V. Recent Advances in Our Understanding of HLA-G Biology: Lessons from a Wide Spectrum of Human Diseases. *J Immunol Res*, v. 2016, p. 1–14, 2016.

MOREAU, P.; CONTU, L.; ALBA, F.; LAI, S.; SIMOES, R.; ORRÙ, S.; CARCASSI, C.; ROGER, M.; RABREAU, M.; CAROSELLA, E. HLA-G gene polymorphism in human placentas: possible association of G*0106 allele with preeclampsia and miscarriage. *Biology of reproduction*, v. 79, n. 3, p. 459–67, set. 2008.

NARDI, S.; SLOWIK, R.; MICHELON, T.; FELIPE, L.; WAGNER, B.; NEUMANN, J.; HORN, P.; GRAC, M.; REBMANN, V. High Amounts of Total and Extracellular Vesicle-Derived Soluble HLA-G are Associated with HLA-G 14-bp Deletion Variant in Women with Embryo Implantation Failure. *Am J Reprod Immunol*, v. 75, n. 6, p. 661–671, 2016.

- NILSSON, L. L.; DJURISIC, S.; ANDERSEN, A. N.; MELBYE, M.; BJERRE, D. Distribution of HLA-G extended haplotypes and one HLA-E polymorphism in a large-scale study of mother-child dyads with and without severe preeclampsia and eclampsia. *HLA*, v. 88, n. 4, p. 172–186, 2016.
- NIU, Z.; WANG, L.; TK, R.; GUO, Y.; SB, W.; YAO, Y. A meta-analysis of the impact of human leukocyte antigen-G on the outcomes of IVF/ICSI. *Reproductive BioMedicine Online*, v. 34, n. 6, p. 611–618, 2017.
- NOWAK, I.; MALINOWSKI, A.; BARCZ, E.; WILCZYŃSKI, J.; WAGNER, M.; MAJORCZYK, E.; MOTAK-POCHRZĘST, H.; BANASIK, M.; KUŚNIERCZYK, P. Possible Role of HLA-G, LILRB1 and KIR2DL4 Gene Polymorphisms in Spontaneous Miscarriage. *Arch Immunol Ther Exp*, v. 64, n. 6, p. 505–514, 2016.
- OMAR, S. Y. Al; MANSOUR, L.; ALKHURIJI, A. F.; ALWASEL, S. Genetic association between the HLA-G 14-bp insertion/deletion polymorphism and the recurrent spontaneous abortions in Saudi Arabian women. *Genet Mol Res*, v. 14, n. 1, p. 286–293, 2015.
- PABALAN, N.; JARJANAZI, H.; SUN, C.; IVERSEN, A. C. Meta-analysis of the human leukocyte antigen-G (HLA-G) 14 bp insertion/deletion polymorphism as a risk factor for preeclampsia. *Tissue Antigens*, v. 86, n. 3, p. 186–194, 2015.
- PARK, K. S.; PARK, J. S.; NAM, J. H.; BANG, D.; SOHN, S.; LEE, E. S. HLA-E*0101 and HLA-G*010101 reduce the risk of Behcet's disease. *Tissue Antigens*, v. 69, n. 2, p. 139–144, 2007.
- PEDROZA, L.; SAUMA, M.; VASCONCELOS, J. M.; TAKESHITA, L. Y. C.; RIBEIRO-RODRIGUES, E. M.; SASTRE, D. Systemic lupus erythematosus: association with KIR and SLC11A1 polymorphisms, ethnic predisposition and influence in clinical manifestations at onset revealed by ancestry genetic markers in an urban Brazilian population. *Lupus*, v. 20, n. 3, p. 265–273, 2011.
- PORAS I, YAGHI L, MARTELLI-PALOMINO G, MENDES-JUNIOR CT, MUNIZ YCN, et al. Haplotypes of the HLA-G 3' Untranslated Region respond to endogenous factors of HLA-G+ and HLA-G- cell lines differentially. *Plos One* v.12, n.1, e0169032, 2017.
- PORTO, I. O. P.; MENDES-JUNIOR, C. T.; FELÍCIO, L. P.; GEORG, R. C.; MOREAU, P.; DONADI, E. a.; CHIES, J. A. B.; CASTELLI, E. C. microRNAs targeting the immunomodulatory HLA-G gene: A new survey searching for microRNAs with potential to regulate HLA-G. *Molecular Immunology*, v. 65, n. 2, p. 230–241, 2015.
- QUACH, K.; GROVER, S. A.; KENIGSBURG, S.; LIBRACH, C. L. A combination of single nucleotide polymorphisms in the 3' untranslated region of HLA-G is associated with preeclampsia. *Human Immunology*, v. 75, n. 12, p. 1163–1170, 2014.
- RAMOS, C. S.; GONÇALVES, A. S.; MARINHO, L. C.; GOMES, M. A. A.; SADDI, V. A.; LOPES, A. C.; SIMÕES, R. T.; WASTOWSKI, I. J. Analysis of HLA-G gene polymorphism and protein expression in invasive breast ductal carcinoma. *Human Immunology*, v. 75, n. 7, p. 667–672, 2014.
- RASSI JR, A; RASSI, A; MARIN-NETO, JA. Chagas disease. *The Lancet*, vol. 375, no. 9723, pp. 1388–1402, 2010.
- RIZZO, R.; HVIID, T. V. F.; GOVONI, M.; PADOVAN, M.; RUBINI, M.; MELCHIORRI, L.; STIGNANI, M.; CARTURAN, S. HLA-G genotype and HLA-G expression in systemic lupus erythematosus: HLA-G as a putative susceptibility gene in systemic lupus erythematosus. *Tissue Antigens*, v. 71, n. 6, p. 520–529, 2008.

- RIZZO, R.; RUBINI, M.; GOVONI, M.; PADOVAN, M.; MELCHIORRI, L.; STIGNANI, M.; CARTURAN, S.; FERRETTI, S.; TROTTA, F.; BARICORDI, O. R. HLA-G 14-bp polymorphism regulates the methotrexate response in rheumatoid arthritis. *Pharmacogenet Genomics*, v. 16, n. 9, p. 615–623, 2006.
- ROUSSEAU, P.; LE DISCORDE, M.; MOUILLOT, G.; MARCOU, C.; CAROSELLA, E. D.; MOREAU, P. The 14 bp Deletion-Insertion Polymorphism in the 3' UT Region of the HLA-G Gene Influences HLA-G mRNA Stability. *Human Immunology*, v. 64, n. 11, p. 1005–1010, 2003.
- SABBAGH, A.; LUISI, P.; CASTELLI, E. C.; GINEAU, L.; COURTIN, D.; MILET, J.; MASSARO, J. D.; LAAYOUNI, H.; MOREAU, P. Worldwide genetic variation at the 3'untranslated region of the HLA-G gene: balancing selection influencing genetic diversity. *Genes and Immunity*, v. 15, n. 2, p. 95–106, 2014.
- SAKLY, K.; MAATOUK, M.; HAMMAMI, S.; HARZALLAH, O.; SAKLY, W.; FEKI, S.; MIRSHAHI, M.; GHEDIRA, I.; SAKLY, N. HLA-G 14 bp insertion / deletion polymorphism and its association with sHLA-G levels in Behçet's disease Tunisian patients.
- SEGAT, L.; CATAMO, E.; FABRIS, A.; MORGUTTI, M.; D'AGARO, P.; CAMPELLO, C.; CROVELLA, S. HLA-G*0105N allele is associated with augmented risk for HIV infection in white female patients. *AIDS*, v. 31, n. 24, p. 1961–1964, 2010.
- SEGAT, L.; CATAMO, E.; FABRIS, A.; PADOVAN, L.; MORGUTTI, M.; CROVELLA, S. HLA-G 3' UTR haplotypes and HIV vertical transmission. *AIDS*, v. 23, n. 14, p. 1916–1918, 2009.
- SEGAT, L.; ZUPIN, L.; KIM, H.; CATAMO, E.; THEA, D.; KANKASA, C.; ALDROVANDI, G.; KUHN, L.; CROVELLA, S. HLA-G 14 bp deletion/insertion polymorphism and mother-to-child transmission of HIV. *Tissue Antigens*, v. 83, n. 3, p. 161–167, 2014.
- SHANKARKUMAR, U.; SHANKARKUMAR, A.; CHEDDA, Z.; GHOSH, K. Role of 14-bp deletion/insertion polymorphism in exon 8 of the HLA-G gene in recurrent spontaneous abortion patients. *J Hum Reprod Sci*, v. 4, n. 3, p. 143–146, 2011.
- SHI, X.; XIE, X.; JIA, Y.; LI, S. Maternal genetic polymorphisms and unexplained recurrent miscarriage: a systematic review and meta-analysis. *Clin Genet*, v. 91, n. 2, p. 265–284, 2017.
- SILVA, G. K.; VIANNA, P.; DEGANI, T.; CROVELLA, S.; CATAMO, E.; ALICIA, E.; CORDERO, A.; SUÑÉ, V.; KUHMMER, R.; SPRINZ, E.; KUHMMER, R.; ARTUR, J.; CHIES, B. Influence of HLA-G polymorphisms in human immunodeficiency virus infection and hepatitis C virus co-infection in Brazilian and Italian individuals. *Infect Genet Evol*, v. 21, p. 418–423, 2014.
- SILVA, I. D.; MUNIZ, Y. C. N.; SOUSA, M. C. P. S.; SILVA, K. R.; CASTELLI, E. C.; FILHO, J. C. G.; OSTA, A. P.; LIMA, M. I. M.; SIMÕES, R. T. HLA-G 3'NTR polymorphisms in high grade and invasive cervico-vaginal cancer. *Hum Immunol*, v. 74, n. 4, p. 452–458, 2013.
- SILVA, H. P. V.; URURAHY, M. A. G.; SOUZA, K. S. C.; LOUREIRO, M. B.; OLIVEIRA, Y. M. C.; OLIVEIRA, G. H. M.; LUCHESSI, A. D.; CARVALHO, K. T. C.; FREITAS, J. The association between the HLA-G 14-bp insertion / deletion polymorphism and type 1 diabetes. *Genes Immun*, v. 17, n. 1, p. 13–18, 2016.
- SIMÕES, R.; GONÇALVES, M.; CASTELLI, E.; JÚNIOR, C.; BETTINI, J.; DISCORDE, M.; DUARTE, G.; QUINTANA, S.; SIMÕES, A.; MOREAU, P.; CAROSELLA, E.; SOARES, E.; DONADI, E. HLA-G polymorphisms in women with squamous intraepithelial lesions harboring human papillomavirus. *Mod Pathol*, v. 22, n. 8, p. 1075–1082, 2009.

SIPAK-SZMIGIEL, O.; CYBULSKI, C.; WOKOŁORCZYK, D.; LUBIŃSKI, J.; KURZAWA, R.; BACZKOWSKI, T.; RADWAN, M.; RADWAN, P.; RONIN-WALKNOWSKA, E. HLA-G polymorphism and in vitro fertilization failure in a Polish population. *Tissue Antigens*, v. 73, n. 4, p. 348–352, 2009.

SIZZANO F, TESTI M, ZITO L, CROCCHIOLO R, TROIANO M, MAZZI B, TURCHIANO G, TORCHIO M, PULTRONE C, GREGORI S, CHIESA R, GAZIEV J, SODANI P, MARKTEL S, AMOROSO A, RONCAROLO MG, LUCARELLI G, CICERI F, ANDREANI M, FLEISCHHAUER K. Genotypes and haplotypes in the 3' untranslated region of the HLA-G gene and their association with clinical outcome of hematopoietic stem cell transplantation for beta-thalassemia. *Tissue Antigens*. Maio;79(5):326-32, 2012.

SURYANARAYANA, V.; RAO, L.; KANAKAVALLI, M.; PADMALATHA, V.; RASESWARI, T.; DEENADAYAL, M.; SINGH, L. Association between novel HLA-G genotypes and risk of recurrent miscarriages: a case-control study in a South Indian population. *Reprod Sci*, v. 15, n. 8, p. 817–824, 2008.

TAN, Z.; RANDALL, G.; FAN, J.; CAMORETTI-MERCADO, B.; BROCKMAN-SCHNEIDER, R.; PAN, L.; SOLWAY, J.; GERN, J.; LEMANSKE, R.; NICOLAE, D.; OBER, C. Allele-specific targeting of microRNAs to HLA-G and risk of asthma. *American journal of human genetics*, v. 81, n. 4, p. 829–834, out. 2007.

TAN Z, SHON AM, OBER C. Evidence of balancing selection at the HLA-G promoter region. *Hum Mol Genet*. Dez 1;14(23):3619-28, 2005.

TAN Z, RANDALL G, FAN J, CAMORETTI-MERCADO B, BROCKMAN-SCHNEIDER R, PAN L, SOLWAY J, GERN JE, LEMANSKE RF, NICOLAE D, OBER C. Allele-specific targeting of microRNAs to HLA-G and risk of asthma. *Am J Hum Genet*. Out, 81(4):829-34, 2007.

TEIXEIRA, A. C.; SOUZA, F. F.; MARANO, L. A.; DEGHAIDE, N. H. S.; FERREIRA, S. C. The 14bp-deletion allele in the HLA-G gene confers susceptibility to the development of hepatocellular. *Tissue Antigens*, v. 81, n. 6, p. 408–413, 2013.

THUDE, H.; JANSSEN, M.; STERNECK, M.; NASHAN, B.; KOCH, M. 14-bp ins/del polymorphism and +3142C>G SNP of the HLA-G gene have a significant impact on acute rejection after liver transplantation. *Hum Immunol*, v. 77, n. 12, p. 1159–1165, 2016.

TORRES, M.; LUQUE, J.; LORITE, P.; ISLA-TEJERA, B.; PALOMEQUE, T.; AUMENTE, M.; ARIZON, J.; PEÑA, J. 14-Base pair polymorphism of human leukocyte antigen-G as genetic determinant in heart transplantation and cyclosporine therapy monitoring. *Human immunology*, v. 70, n. 10, p. 830–5, out. 2009.

TRIPATHI, P.; ABBAS, A.; NAIK, S.; AGRAWAL, S. Role of 14-bp deletion in the HLA-G gene in the maintenance of pregnancy. *Tissue Antigens*, v. 64, n. 6, p. 706–710, 2004.

TWITO, T.; JOSEPH, J.; MOCIORNITA, A.; RAO, V.; ROSS, H.; DELGADO, D. The 14-bp deletion in the HLA-G gene indicates a low risk for acute cellular rejection in heart transplant recipients. *The Journal of heart and lung transplantation: the official publication of the International Society for Heart Transplantation*, v. 30, n. 7, p. 778–82, jul. 2011.

VARGAS, R.; SARTURI, P.; MATTAR, S.; BOMPEIXE, E.; SILVA, J.; PIRRI, A.; BICALHO, M. Association of HLA-G alleles and 3' UTR 14 bp haplotypes with recurrent miscarriage in Brazilian couples. *Human immunology*, v. 72, n. 6, p. 479–85, jun. 2011.

VEIT, T.; CORDERO, E.; MUCENIC, T.; MONTICIELO, O.; BRENOL, J.; XAVIER, R.; DELGADO-CAÑEDO, A.; CHIES, J. Association of the HLA-G 14 bp polymorphism with systemic lupus erythematosus. *Lupus*, v. 18, n. 5, p. 424–430, 2009.

- VEIT, T. D.; CHIES, J. A. B. Tolerance versus immune response - MicroRNAs as important elements in the regulation of the HLA-G gene expression. *Transplant Immunology*, v. 20, n. 4, p. 229–231, 2009.
- VEIT, T. D.; LIMA, C. P. S. De; CAVALHEIRO, L. C.; BRENOL, C. V.; BRENOL, J. C. T. HLA-G +3142 polymorphism as a susceptibility marker in two rheumatoid arthritis populations in Brazil. *Tissue Antigens*, v. 83, n. 4, p. 260–266, 2014.
- VEIT, T. D.; VIANNA, P.; SCHEIBEL, I.; BRENOL, C. V.; BRENOL, C.; BRENOL, J. C. T.; XAVIER, R. M.; DELGADO-CAÑEDO, A.; GUTIERREZ, J. E.; BRANDALIZE, A. P. C.; SCHULER-FACCINI, L.; CHIES, J. A. B. Association of the HLA-G 14-bp insertion/deletion polymorphism with juvenile idiopathic arthritis and rheumatoid arthritis. *Tissue antigens*, v. 71, n. 5, p. 440–6, maio 2008.
- VIANNA, P.; DALMÁZ, C.; VEIT, T.; TEDOLDI, C.; ROISENBERG, I.; CHIES, J. Immunogenetics of pregnancy: role of a 14-bp deletion in the maternal HLA-G gene in primiparous pre-eclamptic Brazilian women. *Human immunology*, v. 68, n. 8, p. 668–74, ago. 2007.
- WANG, X.; JIANG, W.; ZHANG, D. Association of 14-bp insertion/deletion polymorphism of HLA-G gene with unexplained recurrent spontaneous abortion: a meta-analysis. *Tissue Antigens*, v. 81, n. 2, p. 108–115, 2013.
- WATERHOUSE, M.; WÄSCH, R.; BERTZ, H.; FINKE, J. Soluble HLA-G Molecules and HLA-G 14–Base Pair Polymorphism After Allogeneic Hematopoietic Cell Transplantation. *TPS*, v. 45, n. 1, p. 397–401, 2013.
- WIŚNIEWSKI A, BILIŃSKA M, KLIMCZAK A, WAGNER M, MAJORCZYK E, NOWAK I, POKRYSZKO-DRAGAN A, K. P. Association of the HLA-G gene polymorphism with multiple sclerosis in a Polish population. *Int J Immunogenet*, v. 37, n. 4, p. 307–311, 2010.
- WU, F.; WU, L.; LUO, X.; TANG, Z.; YANG, M.; XIE, C.; LIU, N. Lack of association between HLA-G 14-bp polymorphism and systemic lupus erythematosus in a Han Chinese population. *Lupus*, v. 18, n. 14, p. 1259–1266, 2009.
- XU, H.; SHI, W.; LIN, A.; YAN, W. HLA-G 3' untranslated region polymorphisms influence the susceptibility for human papillomavirus infection. *Tissue Antigens*, v. 84, n. 2, p. 216–222, 2014.
- XUE, S.; YANG, J.; YAO, F.; XU, L.; FAN, L.; FAN, L. Recurrent spontaneous abortions patients have more -14 bp/+14 bp heterozygotes in the 3'NT region of the HLA-G gene in a Chinese Han population. *Tissue Antigens*, v. 69, n. Suppl 1, p. 153–155, 2007.
- YAN, W. H.; LIN, A.; CHEN, X. J.; DAI, M. Z.; GAN, L. H.; ZHOU, M. Y.; ZHU, M.; SHI, W. W.; LIU, J. M. Association of the maternal 14-bp insertion polymorphism in the HLA-G gene in women with recurrent spontaneous abortions. *Tissue Antigens*, v. 68, n. 6, p. 521–523, 2006.
- YANG, Y.; CHANG, T.; CHEN, T.; LIN, W.; CHANG, S.; LEE, Y. Human leucocyte antigen-G polymorphisms are associated with cervical squamous cell carcinoma risk in Taiwanese women. *European Journal of Cancer*, v. 50, n. 2, p. 469–474, 2014.
- ZAMBRA, F. M. B.; BIOLCHI, V.; CERQUEIRA, C. C. S. De; BRUM, I. S.; CASTELLI, E. C.; CHIES, J. A. B. Immunogenetics of prostate cancer and benign hyperplasia – the potential use of an HLA-G variant as a tag SNP for prostate cancer risk. *HLA*, v. 87, n. 2, p. 79–88, 2016.

ZHANG, X.; LI, S.; ZHANG, Y.; LU, Y.; WANG, J.; XU, J.; LI, X.; QIN, X. Meta-analysis of the relationship between 14bp insertion/deletion polymorphism of HLA-G gene and susceptibility to systemic lupus erythematosus. *Hum Immunol*, v. 75, n. 12, p. 1171–1176, 2014.

ZHANG, Z.; LI, Y.; ZHANG, L. L.; JIA, L. T.; YANG, X. Q. Association of 14 bp insertion / deletion polymorphism of the HLA-G gene in father with severe preeclampsia in Chinese. *Tissue Antigens*, v. 80, n. 2, p. 158–164, 2012.

ZHENG, X.; LI, C.; XU, D.; LIN, A.; BAO, W.; YANG, G. Analysis of the plasma soluble human leukocyte antigen – G and interleukin-10 levels in childhood atopic asthma. *HIM*, v. 71, n. 10, p. 982–987, 2010.

ZHENG, X.; ZHU, F.; SHI, W.; LIN, A.; YAN, W. The HLA-G 14 bp insertion/deletion polymorphism is a putative susceptible factor for active human cytomegalovirus infection in children. *Tissue antigens*, v. 74, n. 4, p. 317–21, out. 2009.

ZHU, Y.; HUO, Z.; LAI, J.; LI, S.; JIAO, H.; DANG, J.; JIN, C. Case-control study of a HLA-G 14-bp insertion-deletion polymorphism in women with recurrent miscarriages. *Scandinavian journal of immunology*, v. 71, n. 1, p. 52–4, jan. 2010.

ZIDI, I.; KHARRAT, N.; ABDELHEDI, R.; BEN, A.; BALIGH, A.; BEN, H.; BOUAYED, N.; ABID, L.; REBAI, A.; RIZZO, R. Nonclassical human leukocyte antigen (HLA-G, HLA-E, and HLA-F) in coronary artery disease. *Human Immunology*, v. 77, n. 4, p. 325–329, 2016.

ZIDI, I.; YAHIA, H. Ben; BORTOLOTTI, D.; MOUELHI, L.; LAARIBI, A. B.; AYADI, S.; ZIDI, N.; HOUISSA, F.; DEBBECH, R.; BOUDABOUS, A.; NAJJAR, T.; LUCA, D. Di; RIZZO, R. Association between sHLA-G and HLA-G 14-bp deletion / insertion polymorphism in Crohn's disease. *Int Immunol*, v. 27, n. 6, p. 289–296, 2015.

ANEXO A

Artigo Publicado - de Almeida BS, Muniz YCN, Prompt AH, Castelli EC, Mendes-Junior CT, Donadi EA.,
Genetic association between HLA-G 14-bp polymorphism and diseases: a systematic review and meta-analysis. Hum Immunol. 2018 Oct;79(10):724-735.

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

Genetic association between *HLA-G* 14-bp polymorphism and diseases: A systematic review and meta-analysis

Bibiana Sgorla de Almeida^{a,b,1,*}, Yara Costa Netto Muniz^{c,1}, Alice Heidrich Prompt^c,
 Erick C. Castelli^d, Celso Teixeira Mendes-Junior^e, Eduardo Antonio Donadi^a

^a Divisão de Imunologia Clínica, Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto (FMRP), Universidade de São Paulo (USP), 14049-900 Ribeirão Preto, SP, Brazil

^b Laboratório Multiusuário de Estudos em Biologia, Centro de Ciências Biológicas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brazil

^c Departamento de Biologia Celular, Embriologia e Genética, Centro de Ciências Biológicas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brazil

^d Departamento de Patologia, Faculdade de Medicina de Botucatu, Unesp – Univ. Estadual Paulista, 13618-970 Botucatu, SP, Brazil

^e Faculdade de Filosofia Ciências e Letras de Ribeirão Preto (FFCLRP), Universidade de São Paulo (USP), 14049-900 Ribeirão Preto, SP, Brazil

ARTICLE INFO

Keywords:

Non-classical MHC
 HLA-G
 14-bp polymorphism
 Indel
 3'UTR

ABSTRACT

Background: HLA-G is an immune checkpoint molecule. Since a differential molecule expression has been reported even for healthy individuals, many studies have focused on polymorphisms at *HLA-G* regulatory regions, particularly the 3' untranslated region (3'UTR). The presence/absence of a 14-bp sequence was the first polymorphism described and it is the most studied in association between *HLA-G* and disorders.

Methods: In this study, we performed a systematic review and meta-analysis of all association studies published regarding the *HLA-G* 14-bp.

Results: We verified association between 14-bp alleles and diseases in the following situations: (1) presence of 14-bp (insertion) conferred susceptibility to preeclampsia (child alleles evaluated) and systemic lupus erythematosus (OR = 1.42; 95%CI = 1.04–1.93; p = 0.026 and OR = 1.13; 95%CI = 1.01–1.27, p = 0.028); (2) 14-bp absence (deletion) was associated with increased risk to breast cancer (OR = 1.23; 95%CI = 1.06–1.43; p = 0.006) and human Cytomegalovirus infection (OR = 2.06; 95%CI = 1.60–2.64; p < 0.0001); and (3) a risk association was observed between the group of reproductive disorders and the 14-bp insertion (OR = 1.12; 95%CI = 1.01–1.24; p = 0.034).

Conclusions: Considering that others 14-bp associations were inconclusive and that other variation sites observed at *HLA-G* 3'UTR exhibit a proven role on post-transcriptional regulation of *HLA-G* expression, the complete 3'UTR segment should be analyzed in terms of disease susceptibility, instead of a single polymorphism.

ANEXO B

Processo Comitê de Ética em Pesquisa HCPR nº 11237/2009.

Ribeirão Preto, 09 de fevereiro de 2010

Ofício nº 343/2010
CEP/MGV

CÓPIA

Prezados Senhores,

O trabalho intitulado **"CORRELAÇÃO ENTRE OS POLIMORFISMOS E A EXPRESSÃO DE CTLA-4, PD-1 E HLA-G E MANIFESTAÇÕES CLÍNICAS EM PACIENTES PORTADORES DE DOENÇA DE CHAGAS"** foi analisado pelo Comitê de Ética em Pesquisa, em sua 301ª Reunião Ordinária realizada em 08/02/2010 e enquadrado na categoria: **APROVADO, bem como o Termo de Consentimento Livre e Esclarecido**, de acordo com o Processo HCRP nº 11237/2009.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (ICH-GCP), bem como a Resolução nº 196/96 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente,

DRª MARCIA GUIMARÃES VILLANOVA
Vice-Coordenadora do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores
FABRÍCIO CÉSAR DIAS
PROF. DR. EDUARDO ANTONIO DONADI (Supervisor)
Depto. de Clínica Médica

APÊNDICE A

Dados da análise populacional da região Promotora-Distal

Diversidade Haplotípica				Teste de Neutralidade - Ewens-Watterson				
		Diversidade Haplotípica	Desvio Padrão			Valor de p	Valor F observado	Valor F esperado
Americanas	BSP	0,7648	0,0095	Americanas	BSP	0,92400	0,23635	0,21472
	CLM	0,7480	0,0190		CLM	0,73900	0,25594	0,25863
	MXL	0,7787	0,0197		MXL	0,38600	0,22711	0,28415
	PEL	0,7271	0,0208		PEL	0,46200	0,27718	0,30021
	PUR	0,7766	0,0192		PUR	0,31200	0,22712	0,29059
Europeias	CEU	0,6692	0,0281	Europeias	CEU	0,73400	0,33415	0,28541
	FIN	0,5506	0,0354		FIN	0,51000	0,45215	0,37763
	GBR	0,7156	0,0235		GBR	0,83800	0,28828	0,26489
	IBS	0,7131	0,0214		IBS	0,46500	0,29024	0,31686
	TSI	0,7202	0,0207		TSI	0,97000	0,28309	0,21296
Africanas	ACB	0,7943	0,0115	Africanas	ACB	0,91300	0,20985	0,23922
	ASW	0,7992	0,0187		ASW	0,71800	0,20684	0,20946
	ESN	0,7930	0,0126		ESN	0,54100	0,21100	0,28185
	GW							
	D	0,7880	0,0103		GWD	0,76500	0,21544	0,27479
	MSL	0,7922	0,0161		MSL	0,55000	0,21246	0,25590
	YRI	0,7753	0,0125		YRI	0,95400	0,22822	0,22862
	LWK	0,7895	0,0168		LWK	0,50100	0,21444	0,24048
Sul-Asiáticas	BEB	0,7795	0,0098	Sul-Asiáticas	BEB	0,94600	0,22505	0,26058
	GIH	0,7385	0,0194		GIH	0,32600	0,26500	0,35398
	ITU	0,7868	0,0133		ITU	0,38800	0,21699	0,31490
	PJL	0,7814	0,0137		PJL	0,73900	0,22266	0,28644
	STU	0,7663	0,0108		STU	0,84700	0,23739	0,29252
Leste-Asiáticas	CHB	0,7581	0,0127	Leste-Asiáticas	CHB	0,73900	0,24555	0,31297
	CHS	0,6794	0,0164		CHS	0,46000	0,32373	0,48102
	CDX	0,6324	0,0166		CDX	0,15800	0,37083	0,61495
	KHV	0,7664	0,0140		KHV	<i>0,54600</i>	0,23738	0,34397
	JPT	0,6673	0,0208		JPT	0,42000	0,33594	0,41394

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

AMOVA			AMOVA		
Variância	% de variação	Valor de <i>p</i>	Variância	% de variação	Valor de <i>p</i>
Entre grupos	3,12	0,0000	Entre grupos	3,47	0,0000
Entre populações dentro dos grupos	1,59	0,0000	Entre populações dentro dos grupos	1,73	0,0000
Dentro das populações	95,29	0,0000	Dentro das populações	94,8	0,0000

5 Grupos - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,03120
FSC	0,01646
FST	0,04714

5 Grupos 1000 Genomes - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,03473
FSC	0,01790
FST	0,05201

AMOVA		
Variância	% de variação	Valor de <i>p</i>
Entre grupos	4,17	0,0000
Entre populações dentro dos grupos	1,87	0,0000
Dentro das populações	93,96	0,0000

4 Grupos 1000 Genomes - Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,04169
FSC	0,01956
FST	0,06044

Distância Genética – Valores de F_{ST} (parte 1)

		Americanas					Europeias					Africanas						
		BSP	CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
Americanas	BSP	*																
	CLM	0,0065	*															
	MXL	-0,0029	0,0019	*														
	PEL	0,0005	0,0103	0,0002	*													
	PUR	0,0108	0,0009	0,0056	0,0185	*												
Europeias	CEU	0,0205	0,0207	0,0241	0,0164	0,0200	*											
	FIN	0,0621	0,0599	0,0757	0,0595	0,0570	0,0094	*										
	GBR	0,0103	0,0205	0,0118	0,0041	0,0189	0,0023	0,0330	*									
	IBS	0,0072	0,0114	0,0072	0,0002	0,0135	0,0038	0,0345	-0,0023	*								
	TSI	0,0032	0,0045	0,0036	-0,0004	0,0108	0,0040	0,0362	0,0019	-0,0011	*							
Africanas	ACB	0,0205	0,0150	0,0130	0,0352	0,0220	0,0701	0,1286	0,0602	0,0502	0,0373	*						
	ASW	0,0028	0,0120	-0,0019	0,0079	0,0189	0,0481	0,1079	0,0302	0,0234	0,0175	0,0063	*					
	ESN	0,0166	0,0095	0,0106	0,0284	0,0176	0,0557	0,1084	0,0488	0,0397	0,0274	-0,0021	0,0073	*				
	GWD	0,0222	0,0220	0,0140	0,0360	0,0291	0,0818	0,1453	0,0658	0,0548	0,0439	-0,0005	0,0047	0,0069	*			
	MSL	0,0116	0,0045	0,0060	0,0242	0,0055	0,0411	0,0873	0,0383	0,0300	0,0208	0,0020	0,0070	0,0013	0,0099	*		
	YRI	0,0226	0,0154	0,0145	0,0345	0,0279	0,0738	0,1340	0,0629	0,0515	0,0375	-0,0018	0,0085	-0,0002	0,0012	0,0093	*	
	LWK	0,0008	0,0059	-0,0017	0,0026	0,0100	0,0282	0,0742	0,0161	0,0098	0,0074	0,0145	-0,0009	0,0103	0,0171	0,0068	0,0181	*
Sul-Asiáticas	BEB	0,0321	0,0367	0,0249	0,0435	0,0329	0,0601	0,1227	0,0430	0,0484	0,0409	0,0414	0,0384	0,0383	0,0458	0,0423	0,0391	0,0429
	GIH	0,0277	0,0549	0,0251	0,0273	0,0658	0,0836	0,1557	0,0499	0,0508	0,0461	0,0491	0,0209	0,0492	0,0397	0,0608	0,0422	0,0340
	ITU	0,0103	0,0255	0,0078	0,0148	0,0312	0,0514	0,1137	0,0289	0,0300	0,0224	0,0257	0,0092	0,0240	0,0231	0,0304	0,0230	0,0155
	PJL	0,0183	0,0377	0,0154	0,0227	0,0424	0,0609	0,1274	0,0346	0,0386	0,0318	0,0368	0,0180	0,0348	0,0338	0,0430	0,0330	0,0262
	STU	0,0238	0,0191	0,0165	0,0348	0,0228	0,0537	0,1119	0,0430	0,0419	0,0302	0,0233	0,0274	0,0205	0,0286	0,0266	0,0186	0,0320
Leste-Asiáticas	CHB	0,0315	0,0219	0,0244	0,0450	0,0208	0,0509	0,1034	0,0458	0,0462	0,0344	0,0315	0,0403	0,0272	0,0408	0,0294	0,0292	0,0408
	CHS	0,0711	0,0523	0,0674	0,0893	0,0409	0,0618	0,0970	0,0708	0,0747	0,0644	0,0811	0,0969	0,0710	0,0994	0,0646	0,0826	0,0854
	CDX	0,1117	0,1094	0,1142	0,1341	0,0835	0,0924	0,1263	0,0963	0,1115	0,1080	0,1436	0,1501	0,1315	0,1617	0,1195	0,1499	0,1331
	KHV	0,0448	0,0318	0,0355	0,0573	0,0215	0,0546	0,1036	0,0483	0,0498	0,0451	0,0503	0,0567	0,0443	0,0601	0,0418	0,0502	0,0513
	JPT	0,0820	0,0586	0,0727	0,1029	0,0767	0,1410	0,2082	0,1350	0,1200	0,0974	0,0388	0,0720	0,0428	0,0427	0,0617	0,0256	0,0859

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 1)

		Americanas					Europeias					Africanas						
		BSP	CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
Americanas	BSP	*																
	CLM	0,0083	*															
	MXL	0,1924	0,1143	*														
	PEL	0,2396	0,0464	0,1182	*													
	PUR	0,0009	0,2062	0,1334	0,0015	*												
Europeias	CEU	0,0091	0,0000	0,0019	0,0339	0,0000	*											
	FIN	0,0000	0,0000	0,0000	0,0000	0,0000	0,4112	*										
	GBR	0,0010	0,0000	0,0011	0,1141	0,0001	0,5176	0,0013	*									
	IBS	0,0030	0,0014	0,0063	0,4477	0,0017	0,0638	0,0000	0,5633	*								
	TSI	0,0444	0,0311	0,0873	0,7525	0,0005	0,6538	0,0023	0,1446	0,2000	*							
Africanas	ACB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*						
	ASW	0,0002	0,0175	0,2644	0,0059	0,0001	0,0000	0,0000	0,0000	0,0000	0,0001	0,0349	*					
	ESN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,3255	0,0007	*				
	GWD	0,0000	0,0003	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0718	0,0335	0,0005	*			
	MSL	0,0000	0,0000	0,0010	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,2294	0,1015	0,0151	0,0057	*		
	YRI	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,6320	0,0129	0,5429	0,0279	0,0288	*	
	LWK	0,0000	0,0061	0,0008	0,0031	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1131	0,0000	0,0004	0,0008	0,0004	*
Sul-Asiáticas	BEB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	GIH	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	ITU	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	PJL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	STU	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Leste-Asiáticas	CHB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CHS	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CDX	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	KHV	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	JPT	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 2)

Sul-Asiáticas					Leste-Asiáticas				
BEB	GIH	ITU	PJL	STU	CHB	CHS	CDX	KHV	JPT
*									
0,0000	*								
0,0001	0,1977	*							
0,0078	0,2492	0,6711	*						
0,5763	0,0000	0,0005	0,0061	*					
0,3747	0,0000	0,0000	0,0001	0,7507	*				
0,0000	0,0000	0,0000	0,0000	0,0000	0,0011	*			
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	*		
0,0032	0,0000	0,0000	0,0000	0,0004	0,0000	0,0000	0,0000	*	
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

APÊNDICE B

Dados da análise populacional da região Promotora-Proximal

Diversidade Haplotípica				Teste de Neutralidade - Ewens-Watterson				
		Diversidade Haplotípica	Desvio Padrão			Valor de p	Valor F observado	Valor F esperado
Americanas	BSP	0,8124	0,0091	Americanas	BSP	0,01900	0,18893	0,30475
	CLM	0,8607	0,0123		CLM	0,15000	0,14390	0,20738
	MXL	0,8363	0,0186		MXL	0,18200	0,16997	0,22663
	PEL	0,7485	0,0217		PEL	0,73000	0,25591	0,22190
	PUR	0,8621	0,0113		PUR	0,10500	0,14200	0,21638
Europeias	CEU	0,7341	0,0193	Europeias	CEU	0,83100	0,26962	0,26760
	FIN	0,7562	0,0155		FIN	0,60300	0,24758	0,31706
	GBR	0,7406	0,0207		GBR	0,83400	0,26347	0,23859
	IBS	0,7809	0,0173		IBS	0,75100	0,22277	0,21956
	TSI	0,7813	0,0161		TSI	0,78000	0,22235	0,21850
Africanas	ACB	0,8409	0,0143	Africanas	ACB	0,48000	0,16352	0,19571
	ASW	0,8391	0,0201		ASW	0,49100	0,16724	0,17830
	ESN	0,8375	0,0120		ESN	0,12200	0,16677	0,26006
	GWD	0,8254	0,0132		GWD	0,62600	0,17828	0,21358
	MSL	0,8633	0,0097		MSL	0,16900	0,14173	0,23645
	YRI	0,8189	0,0143		YRI	0,06600	0,18483	0,26909
	LWK	0,8366	0,0145		LWK	0,12100	0,16753	0,24484
Sul-Asiáticas	BEB	0,6894	0,0216	Sul-Asiáticas	BEB	0,91300	0,31456	0,30840
	GIH	0,6187	0,0311		GIH	0,69500	0,38426	0,34741
	ITU	0,6606	0,0250		ITU	0,79100	0,34259	0,31987
	PJL	0,6365	0,0301		PJL	0,90500	0,36685	0,28411
	STU	0,7274	0,0159		STU	0,79200	0,27618	0,31600
Leste-Asiáticas	CHB	0,7477	0,0150	Leste-Asiáticas	CHB	0,66900	0,25596	0,31537
	CHS	0,7086	0,0174		CHS	0,91500	0,29467	0,31954
	CDX	0,6294	0,0178		CDX	0,09500	0,37374	0,61225
	KHV	0,7421	0,0138		KHV	0,39300	0,26159	0,38004
	JPT	0,7384	0,0126		JPT	0,42100	0,26516	0,41528

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

AMOVA			AMOVA		
Variância	% de variação	Valor de <i>p</i>	Variância	% de variação	Valor de <i>p</i>
Entre grupos	4,55	0,0000	Entre grupos	3,15	0,0000
Entre populações dentro dos grupos	2,83	0,0000	Entre populações dentro dos grupos	1,29	0,0000
Dentro das populações	92,62	0,0000	Dentro das populações	95,56	0,0000

5 Grupos - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,04547
FSC	0,02967
FST	0,07379

5 Grupos 1000 Genomes - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,03146
FSC	0,01335
FST	0,04439

AMOVA		
Variance	% of variation	P-value
Entre grupos	3,58	0,0000
Entre populações dentro dos grupos	1,30	0,0000
Dentro das populações	95,12	0,0000

4 Grupos 1000 Genomes - Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,03585
FSC	0,01345
FST	0,04882

Distância Genética – Valores de F_{ST} (parte 1)

		Americanas					Europeias					Africanas						
		BSP	CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
Americanas	BSP	*																
	CLM	<i>0,0960</i>	*															
	MXL	<i>0,1066</i>	0,0030	*														
	PEL	<i>0,1327</i>	<i>0,0200</i>	0,0102	*													
	PUR	<i>0,1027</i>	0,0049	0,0070	<i>0,0288</i>	*												
Europeias	CEU	<i>0,1497</i>	<i>0,0273</i>	<i>0,0236</i>	<i>0,0107</i>	<i>0,0246</i>	*											
	FIN	<i>0,1359</i>	<i>0,0350</i>	<i>0,0306</i>	<i>0,0336</i>	<i>0,0331</i>	<i>0,0175</i>	*										
	GBR	<i>0,1604</i>	<i>0,0241</i>	<i>0,0170</i>	0,0077	<i>0,0260</i>	0,0035	<i>0,0303</i>	*									
	IBS	<i>0,1444</i>	<i>0,0152</i>	<i>0,0083</i>	<i>0,0073</i>	<i>0,0114</i>	0,0056	<i>0,0288</i>	-0,0004	*								
	TSI	<i>0,1433</i>	<i>0,0157</i>	<i>0,0107</i>	<i>0,0151</i>	<i>0,0113</i>	0,0042	<i>0,0237</i>	0,0023	0,0005	*							
Africanas	ACB	<i>0,1404</i>	<i>0,0332</i>	<i>0,0303</i>	<i>0,0699</i>	<i>0,0152</i>	<i>0,0705</i>	<i>0,0789</i>	<i>0,0620</i>	<i>0,0400</i>	<i>0,0360</i>	*						
	ASW	<i>0,1402</i>	<i>0,0164</i>	0,0075	<i>0,0349</i>	<i>0,0098</i>	<i>0,0428</i>	<i>0,0562</i>	<i>0,0280</i>	<i>0,0137</i>	<i>0,0157</i>	0,0069	*					
	ESN	<i>0,1355</i>	<i>0,0280</i>	<i>0,0242</i>	<i>0,0591</i>	<i>0,0107</i>	<i>0,0581</i>	<i>0,0605</i>	<i>0,0529</i>	<i>0,0315</i>	<i>0,0294</i>	-0,0012	0,0050	*				
	GWD	<i>0,1483</i>	<i>0,0329</i>	<i>0,0291</i>	<i>0,0626</i>	<i>0,0163</i>	<i>0,0650</i>	<i>0,0784</i>	<i>0,0550</i>	<i>0,0350</i>	<i>0,0317</i>	-0,0018	0,0054	0,0034	*			
	MSL	<i>0,1217</i>	<i>0,0252</i>	<i>0,0262</i>	<i>0,0602</i>	<i>0,0090</i>	<i>0,0604</i>	<i>0,0617</i>	<i>0,0587</i>	<i>0,0346</i>	<i>0,0353</i>	0,0054	<i>0,0101</i>	0,0054	0,0049	*		
	YRI	<i>0,1498</i>	<i>0,0393</i>	<i>0,0340</i>	<i>0,0748</i>	<i>0,0211</i>	<i>0,0728</i>	<i>0,0754</i>	<i>0,0655</i>	<i>0,0444</i>	<i>0,0373</i>	-0,0022	<i>0,0098</i>	-0,0017	0,0017	<i>0,0109</i>	*	
	LWK	<i>0,1342</i>	<i>0,0159</i>	<i>0,0097</i>	<i>0,0271</i>	<i>0,0101</i>	<i>0,0365</i>	<i>0,0484</i>	<i>0,0228</i>	<i>0,0094</i>	<i>0,0182</i>	<i>0,0185</i>	0,0015	<i>0,0144</i>	<i>0,0160</i>	<i>0,0133</i>	<i>0,0246</i>	*
Sul-Asiáticas	BEB	<i>0,1934</i>	<i>0,0475</i>	<i>0,0280</i>	<i>0,0308</i>	<i>0,0424</i>	<i>0,0341</i>	<i>0,0617</i>	<i>0,0213</i>	<i>0,0198</i>	<i>0,0142</i>	<i>0,0445</i>	<i>0,0228</i>	<i>0,0406</i>	<i>0,0356</i>	<i>0,0576</i>	<i>0,0405</i>	<i>0,0329</i>
	GIH	<i>0,2442</i>	<i>0,0890</i>	<i>0,0628</i>	<i>0,0750</i>	<i>0,0909</i>	<i>0,0888</i>	<i>0,1144</i>	<i>0,0544</i>	<i>0,0590</i>	<i>0,0540</i>	<i>0,0800</i>	<i>0,0479</i>	<i>0,0809</i>	<i>0,0676</i>	<i>0,1030</i>	<i>0,0744</i>	<i>0,0622</i>
	ITU	<i>0,2147</i>	<i>0,0612</i>	<i>0,0374</i>	<i>0,0399</i>	<i>0,0595</i>	<i>0,0475</i>	<i>0,0817</i>	<i>0,0258</i>	<i>0,0287</i>	<i>0,0250</i>	<i>0,0617</i>	<i>0,0310</i>	<i>0,0603</i>	<i>0,0501</i>	<i>0,0772</i>	<i>0,0591</i>	<i>0,0421</i>
	PJL	<i>0,2290</i>	<i>0,0743</i>	<i>0,0502</i>	<i>0,0531</i>	<i>0,0763</i>	<i>0,0638</i>	<i>0,0947</i>	<i>0,0359</i>	<i>0,0418</i>	<i>0,0373</i>	<i>0,0749</i>	<i>0,0412</i>	<i>0,0744</i>	<i>0,0617</i>	<i>0,0921</i>	<i>0,0712</i>	<i>0,0523</i>
	STU	<i>0,1826</i>	<i>0,0473</i>	<i>0,0314</i>	<i>0,0538</i>	<i>0,0342</i>	<i>0,0489</i>	<i>0,0635</i>	<i>0,0407</i>	<i>0,0315</i>	<i>0,0196</i>	<i>0,0223</i>	<i>0,0179</i>	<i>0,0203</i>	<i>0,0188</i>	<i>0,0400</i>	<i>0,0153</i>	<i>0,0345</i>
Leste-Asiáticas	CHB	<i>0,1468</i>	<i>0,0203</i>	<i>0,0120</i>	<i>0,0120</i>	<i>0,0230</i>	<i>0,0184</i>	<i>0,0404</i>	<i>0,0190</i>	<i>0,0135</i>	<i>0,0088</i>	<i>0,0413</i>	<i>0,0231</i>	<i>0,0346</i>	<i>0,0354</i>	<i>0,0435</i>	<i>0,0402</i>	<i>0,0289</i>
	CHS	<i>0,1462</i>	<i>0,0356</i>	<i>0,0332</i>	<i>0,0160</i>	<i>0,0349</i>	<i>0,0150</i>	<i>0,0411</i>	<i>0,0322</i>	<i>0,0249</i>	<i>0,0234</i>	<i>0,0729</i>	<i>0,0556</i>	<i>0,0603</i>	<i>0,0676</i>	<i>0,0641</i>	<i>0,0739</i>	<i>0,0523</i>
	CDX	<i>0,1885</i>	<i>0,0623</i>	<i>0,0503</i>	<i>0,0169</i>	<i>0,0587</i>	<i>0,0131</i>	<i>0,0498</i>	<i>0,0261</i>	<i>0,0267</i>	<i>0,0282</i>	<i>0,1015</i>	<i>0,0723</i>	<i>0,0874</i>	<i>0,0911</i>	<i>0,0930</i>	<i>0,1018</i>	<i>0,0647</i>
	KHV	<i>0,1521</i>	<i>0,0294</i>	<i>0,0217</i>	<i>0,0197</i>	<i>0,0172</i>	<i>0,0171</i>	<i>0,0419</i>	<i>0,0223</i>	<i>0,0108</i>	<i>0,0114</i>	<i>0,0357</i>	<i>0,0256</i>	<i>0,0254</i>	<i>0,0335</i>	<i>0,0382</i>	<i>0,0360</i>	<i>0,0259</i>
	JPT	<i>0,1542</i>	<i>0,0517</i>	<i>0,0541</i>	<i>0,0800</i>	<i>0,0414</i>	<i>0,0786</i>	<i>0,0995</i>	<i>0,0877</i>	<i>0,0665</i>	<i>0,0531</i>	<i>0,0326</i>	<i>0,0492</i>	<i>0,0326</i>	<i>0,0348</i>	<i>0,0469</i>	<i>0,0298</i>	<i>0,0673</i>

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 1)

		Americanas					Europeias					Africanas						
		BSP	CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
Americanas	BSP	*																
	CLM	0,0000	*															
	MXL	0,0000	0,0212	*														
	PEL	0,0000	0,0000	0,0366	*													
	PUR	0,0000	0,0168	0,0088	0,0000	*												
Europeias	CEU	0,0000	0,0000	0,0000	0,0000	0,0000	*											
	FIN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*										
	GBR	0,0000	0,0000	0,0000	0,0000	0,0000	0,3162	0,0000	*									
	IBS	0,0000	0,0000	0,0000	0,0000	0,0008	0,0026	0,0000	0,4130	*								
	TSI	0,0000	0,0000	0,0000	0,0000	0,0000	0,2173	0,0000	0,1345	0,0643	*							
Africanas	ACB	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	*						
	ASW	0,0000	0,0000	0,0006	0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0948	*					
	ESN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1460	0,0328	*				
	GWD	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,1914	0,0004	0,0003	*			
	MSL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0342	0,0020	0,0007	0,1420	*		
	YRI	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,3721	0,0098	0,8919	0,0046	0,0001	*	
	LWK	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*
Sul-Asiáticas	BEB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	GIH	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	ITU	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	PJL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	STU	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Leste-Asiáticas	CHB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CHS	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CDX	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	KHV	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	JPT	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 2)

Sul-Asiáticas					Leste-Asiáticas				
BEB	GIH	ITU	PJL	STU	CHB	CHS	CDX	KHV	JPT
*									
0,0554	*								
0,1898	0,0880	*							
0,1757	0,2937	0,4003	*						
0,1984	0,0022	0,0010	0,0026	*					
0,0064	0,0000	0,0000	0,0000	0,0000	*				
0,0000	0,0000	0,0000	0,0000	0,0000	0,0511	*			
0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0097	*		
0,0004	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*	
0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	*

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

APÊNDICE C

Dados da análise populacional da região *Enhancer L*

Diversidade Haplotípica				Teste de Neutralidade - Ewens-Watterson				
		Diversidade Haplotípica	Desvio Padrão			Valor de <i>p</i>	Valor F observado	Valor F esperado
Americanas	CLM	0,3488	0,0337	Americanas	CLM	0,20000	0,65301	0,83557
	MXL	0,4094	0,0340		MXL	0,20300	0,59367	0,81316
	PEL	0,4534	0,0230		PEL	0,11100	0,54928	0,81356
	PUR	0,2727	0,0344		PUR	0,28200	0,72862	0,83374
Europeias	CEU	0,3608	0,0329	Europeias	CEU	0,58200	0,64101	0,71942
	FIN	0,3241	0,0341		FIN	0,22500	0,67758	0,83515
	GBR	0,4029	0,0296		GBR	0,16900	0,59930	0,82792
	IBS	0,4054	0,0271		IBS	0,15200	0,59647	0,84034
	TSI	0,4112	0,0264		TSI	0,14900	0,59071	0,83652
Africanas	ACB	0,3742	0,0325	Africanas	ACB	0,61600	0,62777	0,70907
	ASW	0,4372	0,0299		ASW	0,13400	0,56612	0,82164
	ESN	0,3474	0,0329		ESN	0,21100	0,65432	0,83154
	GWD	0,3917	0,0277		GWD	0,18300	0,60999	0,82969
	MSL	0,3147	0,0372		MSL	0,23700	0,68713	0,83431
	YRI	0,3835	0,0289		YRI	0,16500	0,61821	0,84246
	LWK	0,4027	0,0282		LWK	0,17500	0,59925	0,83020
Sul-Asiáticas	BEB	0,3329	0,0362	Sul-Asiáticas	BEB	0,23000	0,66901	0,82930
	GIH	0,4894	0,0113		GIH	0,06700	0,51286	0,82727
	ITU	0,4472	0,0246		ITU	0,40700	0,55495	0,71440
	PJL	0,4468	0,0229		PJL	0,13500	0,55556	0,82665
	STU	0,3431	0,0325		STU	0,21600	0,65854	0,83329
Leste-Asiáticas	CHB	0,2704	0,0348	Leste-Asiáticas	CHB	0,29700	0,73094	0,82410
	CHS	0,1054	0,0277		CHS	0,46900	0,89506	0,83961
	CDX	0,0300	0,0168		CDX	0,74900	0,97016	0,82838
	KHV	0,1872	0,0340		KHV	0,36600	0,81369	0,82878
	JPT	0,3391	0,0339		JPT	0,50300	0,66254	0,69988

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

AMOVA		
Variância	% de variação	Valor de <i>p</i>
Entre grupos	2,58	0,0000
Entre populações dentro dos grupos	1,82	0,0000
Dentro das populações	95,60	0,0000

**5 Grupos - Americanas 1000 Genomes, Europeias 1000 Genomes,
Africanas 1000 Genomes, Sul-Asiáticas 1000 Genomes e Leste-
Asiáticas 1000 Genomes**

Índices de Fixação

FCT	0,02582
FSC	0,01873
FST	0,04406

Distância Genética – Valores de F_{ST} (parte 1)

		Americanas				Europeias					Africanas						
		CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
Americanas	CLM	0,0000															
	MXL	0,0033	0,0000														
	PEL	0,0294	0,0015	0,0000													
	PUR	0,0072	0,0378	0,0804	0,0000												
Europeias	CEU	-0,0050	0,0014	0,0257	0,0097	0,0000											
	FIN	-0,0038	0,0122	0,0443	0,0005	-0,0028	0,0000										
	GBR	0,0023	-0,0064	0,0045	0,0336	0,0007	0,0103	0,0000									
	IBS	0,0036	-0,0061	0,0039	0,0353	0,0019	0,0118	-0,0051	0,0000								
	TSI	0,0056	-0,0061	0,0021	0,0394	0,0037	0,0145	-0,0048	-0,0046	0,0000							
Africanas	ACB	-0,0043	-0,0019	0,0189	0,0149	-0,0047	-0,0004	-0,0021	-0,0012	0,0002	0,0000						
	ASW	0,0166	-0,0047	-0,0053	0,0624	0,0135	0,0293	-0,0025	-0,0027	-0,0038	0,0080	0,0000					
	ESN	-0,0052	0,0038	0,0304	0,0068	-0,0049	-0,0039	0,0028	0,0041	0,0062	-0,0040	0,0174	0,0000				
	GWD	-0,0001	-0,0052	0,0092	0,0268	-0,0013	0,0064	-0,0046	-0,0040	-0,0034	-0,0034	0,0008	0,0003	0,0000			
	MSL	-0,0030	0,0155	0,0493	-0,0018	-0,0017	-0,0053	0,0133	0,0149	0,0178	0,0012	0,0339	-0,0031	0,0090	0,0000		
	YRI	-0,0019	-0,0042	0,0126	0,0222	-0,0028	0,0036	-0,0040	-0,0032	-0,0023	-0,0043	0,0032	-0,0015	-0,0043	0,0059	0,0000	
	LWK	0,0026	-0,0061	0,0048	0,0335	0,0009	0,0104	-0,0052	-0,0048	-0,0046	-0,0019	-0,0022	0,0031	-0,0044	0,0135	-0,0037	0,0000
Sul-Asiáticas	BEB	-0,0050	0,0084	0,0382	0,0022	-0,0042	-0,0053	0,0068	0,0082	0,0107	-0,0024	0,0241	-0,0050	0,0035	-0,0052	0,0011	0,0070
	GIH	0,0793	0,0332	0,0071	0,1450	0,0738	0,0999	0,0387	0,0373	0,0334	0,0634	0,0156	0,0808	0,0473	0,1059	0,0533	0,0391
	ITU	0,0183	-0,0029	-0,0041	0,0613	0,0152	0,0305	-0,0006	-0,0008	-0,0020	0,0099	-0,0061	0,0191	0,0027	0,0347	0,0052	-0,0004
	PJL	0,0245	-0,0006	-0,0053	0,0720	0,0211	0,0382	0,0021	0,0017	0,0001	0,0149	-0,0059	0,0254	0,0062	0,0429	0,0092	0,0023
	STU	-0,0050	0,0053	0,0330	0,0056	-0,0046	-0,0042	0,0041	0,0054	0,0077	-0,0034	0,0196	-0,0049	0,0014	-0,0036	-0,0007	0,0044
Leste-Asiáticas	CHB	0,0078	0,0391	0,0819	-0,0048	0,0104	0,0009	0,0347	0,0364	0,0405	0,0158	0,0639	0,0075	0,0278	-0,0014	0,0232	0,0346
	CHS	0,1102	0,1861	0,2385	0,0526	0,1140	0,0887	0,1651	0,1623	0,1686	0,1269	0,2263	0,1078	0,1456	0,0842	0,1385	0,1614
	CDX	0,1863	0,2794	0,3191	0,1183	0,1875	0,1612	0,2438	0,2353	0,2413	0,2022	0,3199	0,1818	0,2166	0,1617	0,2105	0,2365
	KHV	0,0464	0,0998	0,1520	0,0096	0,0504	0,0317	0,0895	0,0902	0,0957	0,0598	0,1337	0,0453	0,0772	0,0270	0,0707	0,0883
	JPT	-0,0041	0,0101	0,0407	0,0020	-0,0032	-0,0047	0,0084	0,0099	0,0125	-0,0013	0,0262	-0,0040	0,0050	-0,0047	0,0024	0,0087

Distância Genética – Valores de F_{ST} (parte 2)

Sul-Asiáticas					Leste-Asiáticas				
BEB	GIH	ITU	PJL	STU	CHB	CHS	CDX	KHV	JPT
0,0000									
0,0912	0,0000								
0,0254	0,0162	0,0000							
0,0325	0,0109	-0,0045	0,0000						
-0,0051	0,0846	0,0212	0,0278	0,0000					
0,0027	0,1467	0,0627	0,0735	0,0062	0,0000				
0,0986	0,3075	0,2020	0,2219	0,1032	0,0514	0,0000			
0,1772	0,3777	0,2749	0,2984	0,1756	0,1173	0,0184	0,0000		
0,0369	0,2229	0,1246	0,1399	0,0425	0,0088	0,0113	0,0629	0,0000	
-0,0051	0,0949	0,0275	0,0349	-0,0042	0,0025	0,0905	0,1599	0,0341	0,0000

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 1)

		Americanas				Europeias					Africanas						
		CLM	MXL	PEL	PUR	CEU	FIN	GBR	IBS	TSI	ACB	ASW	ESN	GWD	MSL	YRI	LWK
América	MXL	0,2455															
	PEL	0,0144	0,3187														
	PUR	0,1293	0,0109	0,0000													
Europeias	CEU	1,0000	0,3581	0,0152	0,0909												
	FIN	0,6224	0,0870	0,0016	0,3177	0,5605											
	GBR	0,2257	0,9013	0,2053	0,0068	0,2663	0,0911										
	IBS	0,2114	1,0000	0,2288	0,0028	0,2166	0,0633	1,0000									
	TSI	0,1656	1,0000	0,2609	0,0029	0,1702	0,0543	0,9102	0,9138								
Africanas	ACB	0,7630	0,5895	0,0382	0,0530	0,8285	0,3511	0,4803	0,3721	0,2454							
	ASW	0,0743	0,5844	0,7105	0,0008	0,1063	0,0198	0,4467	0,4667	0,5495	0,1513						
	ESN	1,0000	0,2436	0,0113	0,1343	0,9039	0,7100	0,2367	0,2117	0,1532	0,6794	0,0556					
	GWD	0,3580	0,7108	0,0854	0,0091	0,3696	0,1455	0,8230	0,7464	0,6707	0,4788	0,3223	0,3016				
	MSL	0,5131	0,0749	0,0032	0,4233	0,4788	0,8990	0,0784	0,0540	0,0339	0,3260	0,0149	0,5171	0,1239			
	YRI	0,4831	0,6210	0,0663	0,0170	0,5026	0,2053	0,6534	0,5910	0,5061	0,6476	0,2118	0,4079	0,9137	0,1826		
	LWK	0,2454	0,9006	0,1858	0,0053	0,2880	0,0823	1,0000	1,0000	0,8239	0,4265	0,4478	0,2576	0,8247	0,0612	0,6619	
Sul-Asiáticas	BEB	0,8028	0,1443	0,0084	0,2932	0,8455	0,8940	0,1402	0,1150	0,1004	0,5794	0,0365	0,7994	0,2365	0,7893	0,2981	0,1513
	GIH	0,0000	0,0114	0,1389	0,0000	0,0000	0,0000	0,0014	0,0037	0,0027	0,0000	0,0710	0,0001	0,0001	0,0000	0,0004	0,0024
	ITU	0,0229	0,5115	0,5297	0,0000	0,0240	0,0045	0,3062	0,2776	0,3173	0,0578	0,7465	0,0255	0,1406	0,0103	0,1016	0,2798
	PJL	0,0245	0,3959	0,9116	0,0000	0,0284	0,0025	0,2651	0,2813	0,3253	0,0582	0,8115	0,0181	0,1228	0,0024	0,1046	0,2345
	STU	1,0000	0,2153	0,0066	0,1636	0,7541	0,7125	0,1913	0,1787	0,1144	0,5863	0,0429	1,0000	0,2607	0,6109	0,3538	0,2097
Leste-Asiáticas	CHB	0,1235	0,0088	0,0000	1,0000	0,0868	0,2991	0,0086	0,0026	0,0033	0,0552	0,0010	0,1292	0,0089	0,4084	0,0165	0,0057
	CHS	0,0000	0,0000	0,0000	0,0005	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CDX	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	KHV	0,0005	0,0000	0,0000	0,1108	0,0012	0,0080	0,0000	0,0000	0,0000	0,0004	0,0000	0,0029	0,0000	0,0199	0,0000	0,0000
	JPT	0,5596	0,1240	0,0011	0,1924	0,4218	0,6280	0,0720	0,0601	0,0356	0,3068	0,0209	0,5189	0,0891	0,6299	0,1452	0,0664

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 2)

Sul-Asiáticas					Leste-Asiáticas			
BEB	GIH	ITU	PJL	STU	CHB	CHS	CDX	KHV
<i>0,0000</i>								
<i>0,0167</i>	<i>0,0273</i>							
<i>0,0098</i>	0,0812	0,5614						
0,9017	<i>0,0000</i>	<i>0,0161</i>	<i>0,0149</i>					
0,2417	<i>0,0000</i>	<i>0,0001</i>	<i>0,0002</i>	0,1664				
<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0010</i>			
<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	0,0350		
<i>0,0054</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0022</i>	0,1003	0,0710	<i>0,0003</i>	

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

APÊNDICE D

Dados da análise populacional da região 3'NT

Diversidade Haplotípica				Teste de Neutralidade Ewens-Watterson				
		Diversidade Haplotípica	Desvio Padrão			Valor de <i>p</i>	Valor F observado	Valor F esperado
Americanas	BSC	0,8082	0,0142	Americanas	BSC	0,00800	0,19580	0,37699
	BRS	0,8050	0,0051		BRS	0,02100	0,19559	0,43378
	BPR	0,8000	0,0095		BPR	0,00500	0,20208	0,41138
	BSP	0,8150	0,0100		BSP	0,71000	0,18764	0,26334
	BPE	0,8192	0,0114		BPE	0,00100	0,18404	0,39918
	CLM	0,8099	0,0103		CLM	0,18200	0,19438	0,34063
	MXL	0,8234	0,0155		MXL	0,31500	0,18278	0,27035
	PEL	0,7279	0,0202		PEL	0,34000	0,27637	0,33515
	PUR	0,8421	0,0109		PUR	0,16600	0,16195	0,26910
	Europeias	FRM	0,7972		0,0132	Europeias	FRM	0,03700
FRP		0,8100	0,0094	FRP	0,04600		0,19217	0,37854
POR		0,7714	0,0185	POR	0,01600		0,23500	0,42973
CYP		0,8155	0,0100	CYP	0,41700		0,18672	0,27474
ITA		0,7512	0,0161	ITA	0,09400		0,25250	0,42421
CEU		0,7542	0,0185	CEU	0,10700		0,24962	0,38573
FIN		0,7306	0,0147	FIN	0,19700		0,27308	0,42291
GBR		0,7795	0,0167	GBR	0,39300		0,22472	0,31089
IBS		0,7725	0,0167	IBS	0,12400		0,23116	0,35409
TSI		0,7774	0,0144	TSI	0,46400		0,22621	0,29691
Africanas	MAL	0,7989	0,0078	Africanas	MAL	0,00500	0,20288	0,45825
	YAN	0,8308	0,0081		YAN	0,00900	0,17161	0,40839
	SER	0,7775	0,0106		SER	0,00200	0,22410	0,46891
	GUI	0,7410	0,0212		GUI	0,06100	0,26514	0,44214
	TOR	0,8158	0,0214		TOR	0,06000	0,19778	0,33719
	ACB	0,8186	0,0122		ACB	0,39600	0,18566	0,28360
	ASW	0,8104	0,0193		ASW	0,35700	0,19571	0,24010
	ESN	0,8200	0,0104		ESN	0,16500	0,18411	0,31159
	GWD	0,8093	0,0128		GWD	0,54400	0,19430	0,24804
	MSL	0,8222	0,0118		MSL	0,01300	0,18263	0,36761
	YRI	0,7975	0,0134		YRI	0,20900	0,20613	0,31421
	LWK	0,8079	0,0140		LWK	0,04200	0,19611	0,34347
Sul-Asiáticas	MDE	0,7648	0,0139	Sul-Asiáticas	MDE	0,00500	0,23777	0,49113
	BEB	0,8060	0,0078		BEB	0,05400	0,19869	0,41858
	GIH	0,7336	0,0234		GIH	0,02300	0,26985	0,42085
	ITU	0,7650	0,0185		ITU	0,08000	0,23867	0,37576
	PJL	0,7572	0,0178		PJL	0,28100	0,24674	0,38034
	STU	0,8054	0,0096		STU	0,20300	0,19851	0,34834
Leste-Asiáticas	CHB	0,7731	0,0114	Leste-Asiáticas	CHB	0,30300	0,23065	0,37556
	CHS	0,6948	0,0165		CHS	0,30300	0,30838	0,46822
	CDX	0,6324	0,0166		CDX	0,16700	0,37083	0,60348
	KHV	0,7899	0,0117		KHV	0,00200	0,21400	0,46816
	JPT	0,6509	0,0216		JPT	0,36100	0,35226	0,48481

AMOVA

Variância	% de variação	Valor de <i>p</i>
Entre grupos	2,56	0,00000
Entre populações dentro dos grupos	1,74	0,00000
Dentro das populações	95,7	0,00000

5 Grupos - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,02565
FSC	0,01783
FST	0,04302

AMOVA

Variância	% de variação	Valor de <i>p</i>
Entre grupos	1,75	0,00000
Entre populações dentro dos grupos	1,39	0,00000
Dentro das populações	96,87	0,00000

4 Grupos 1000 Genomes - Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,01749
FSC	0,01411
FST	0,03135

AMOVA

Variância	% de variação	Valor de <i>p</i>
Entre grupos	3,65	0,00000
Entre populações dentro dos grupos	2,02	0,00000
Dentro das populações	94,33	0,00000

5 Grupos 1000 Genomes - Americanas, Europeias, Africanas, Sul-Asiáticas e Leste-Asiáticas

Índices de Fixação

FCT	0,03652
FSC	0,02098
FST	0,05674

AMOVA

Variância	% de variação	Valor de <i>p</i>
Entre grupos	1,18	0,00098
Entre populações dentro dos grupos	1,28	0,00000
Dentro das populações	97,55	0,00000

4 Grupos - Revisão Literatura - Americanas, Europeias, Africanas e Leste-Asiáticas

Índices de Fixação

FCT	0,01177
FSC	0,01291
FST	0,02452

Distância Genética – Valores de F_{ST} (parte 1)

		Americanas									Europeias															
		BSC	BRS	BPR	BSP	BPE	CLM	MXL	PEL	PUR	FRM	FRP	POR	CYP	ITA	CEU	FIN	GBR	IBS	TSI						
Americanas	BSC	*																								
	BRS	0,0099	*																							
	BPR	-0,0004	0,0092	*																						
	BSP	-0,0010	0,0080	0,0005	*																					
	BPE	-0,0011	0,0062	-0,0006	0,0023	*																				
	CLM	0,0100	0,0159	0,0076	0,0043	0,0093	*																			
	MXL	0,0026	0,0085	0,0001	-0,0019	0,0041	-0,0003	*																		
	PEL	0,0122	0,0184	0,0109	0,0154	0,0182	0,0277	0,0126	*																	
	PUR	0,0069	0,0160	0,0096	0,0050	0,0068	0,0011	0,0016	0,0304	*																
Europeias	FRM	0,0009	0,0181	0,0113	0,0056	0,0075	0,0138	0,0130	0,0258	0,0087	*															
	FRP	-0,0005	0,0114	0,0021	-0,0014	0,0042	0,0095	0,0018	0,0166	0,0113	0,0064	*														
	POR	0,0334	0,0130	0,0278	0,0341	0,0245	0,0510	0,0343	0,0319	0,0529	0,0516	0,0319	*													
	CYP	0,0001	0,0112	0,0007	-0,0013	0,0043	0,0092	-0,0008	0,0128	0,0100	0,0091	-0,0011	0,0310	*												
	ITA	0,0055	0,0137	0,0111	0,0065	0,0150	0,0186	0,0098	0,0054	0,0193	0,0071	0,0081	0,0410	0,0071	*											
	CEU	0,0074	0,0188	0,0180	0,0119	0,0171	0,0217	0,0170	0,0144	0,0163	0,0026	0,0149	0,0544	0,0152	-0,0006	*										
	FIN	0,0313	0,0389	0,0476	0,0303	0,0438	0,0361	0,0418	0,0542	0,0327	0,0139	0,0326	0,0894	0,0390	0,0199	0,0106	*									
	GBR	0,0087	0,0175	0,0171	0,0092	0,0196	0,0226	0,0103	0,0101	0,0174	0,0093	0,0101	0,0439	0,0090	-0,0006	0,0013	0,0196	*								
	IBS	0,0068	0,0151	0,0119	0,0096	0,0143	0,0167	0,0076	0,0028	0,0140	0,0104	0,0116	0,0397	0,0101	0,0010	0,0017	0,0270	-0,0009	*							
	TSI	0,0027	0,0104	0,0055	0,0009	0,0098	0,0076	0,0024	0,0073	0,0124	0,0060	0,0018	0,0364	0,0022	-0,0017	0,0044	0,0211	0,0023	0,0021	*						
Africanas	MAL	0,0307	0,0357	0,0183	0,0232	0,0232	0,0241	0,0179	0,0516	0,0342	0,0529	0,0236	0,0470	0,0229	0,0575	0,0729	0,1009	0,0633	0,0571	0,0396						
	YAN	0,0346	0,0373	0,0271	0,0297	0,0224	0,0206	0,0276	0,0724	0,0280	0,0440	0,0312	0,0529	0,0327	0,0650	0,0714	0,0881	0,0708	0,0644	0,0465						
	SER	0,0423	0,0369	0,0257	0,0306	0,0336	0,0272	0,0222	0,0535	0,0448	0,0642	0,0311	0,0419	0,0285	0,0597	0,0801	0,1082	0,0676	0,0616	0,0413						
	GUI	0,0654	0,0343	0,0577	0,0623	0,0544	0,0822	0,0627	0,0634	0,0859	0,0853	0,0578	-0,0012	0,0572	0,0719	0,0900	0,1221	0,0719	0,0713	0,0651						
	TOR	-0,0029	0,0061	-0,0067	-0,0036	-0,0057	-0,0013	-0,0035	0,0163	0,0047	0,0090	-0,0027	0,0252	-0,0016	0,0153	0,0226	0,0482	0,0224	0,0154	0,0039						
	ACB	0,0265	0,0346	0,0172	0,0179	0,0205	0,0099	0,0132	0,0551	0,0215	0,0399	0,0198	0,0574	0,0200	0,0511	0,0622	0,0811	0,0569	0,0514	0,0318						
	ASW	0,0090	0,0119	0,0021	0,0032	0,0087	0,0068	-0,0027	0,0128	0,0141	0,0260	0,0050	0,0276	0,0029	0,0182	0,0311	0,0598	0,0210	0,0162	0,0077						
	ESN	0,0187	0,0331	0,0131	0,0159	0,0139	0,0123	0,0153	0,0493	0,0216	0,0301	0,0148	0,0507	0,0170	0,0459	0,0543	0,0746	0,0518	0,0453	0,0283						

	GW D	0,0231	0,0273	0,0114	0,0158	0,0177	0,0085	0,0061	0,0331	0,0195	0,0405	0,0196	0,0453	0,0158	0,0393	0,0519	0,0825	0,0449	0,0356	0,0243
	MSL	0,0075	0,0165	0,0049	0,0068	0,0029	0,0032	0,0047	0,0310	0,0031	0,0162	0,0129	0,0481	0,0122	0,0275	0,0282	0,0522	0,0334	0,0245	0,0178
	YRI	0,0357	0,0443	0,0255	0,0258	0,0304	0,0135	0,0209	0,0632	0,0303	0,0466	0,0267	0,0672	0,0272	0,0571	0,0696	0,0858	0,0644	0,0587	0,0363
	LWK	0,0018	0,0097	0,0006	0,0008	0,0032	0,0086	0,0005	0,0118	0,0121	0,0144	0,0005	0,0264	0,0020	0,0135	0,0211	0,0447	0,0156	0,0112	0,0046
Sul-Asiáticas	MDE	0,0138	0,0170	0,0093	0,0063	0,0196	0,0162	0,0030	0,0116	0,0248	0,0271	0,0073	0,0357	0,0041	0,0096	0,0258	0,0516	0,0138	0,0158	0,0046
	BEB	0,0401	0,0573	0,0416	0,0325	0,0451	0,0357	0,0300	0,0731	0,0340	0,0402	0,0336	0,0880	0,0283	0,0490	0,0575	0,0744	0,0455	0,0566	0,0420
	GIH	0,0521	0,0550	0,0431	0,0382	0,0588	0,0531	0,0337	0,0530	0,0670	0,0705	0,0329	0,0607	0,0293	0,0511	0,0783	0,1036	0,0504	0,0597	0,0398
	ITU	0,0235	0,0302	0,0164	0,0158	0,0288	0,0283	0,0109	0,0226	0,0375	0,0407	0,0140	0,0396	0,0089	0,0238	0,0447	0,0772	0,0262	0,0299	0,0170
	PJL	0,0297	0,0380	0,0241	0,0224	0,0368	0,0374	0,0173	0,0281	0,0436	0,0451	0,0205	0,0482	0,0136	0,0266	0,0474	0,0819	0,0270	0,0338	0,0227
	STU	0,0433	0,0562	0,0402	0,0317	0,0451	0,0257	0,0278	0,0771	0,0332	0,0448	0,0330	0,0876	0,0300	0,0548	0,0656	0,0764	0,0552	0,0622	0,0411
Leste-Asiáticas	CHB	0,0264	0,0461	0,0257	0,0297	0,0266	0,0276	0,0258	0,0465	0,0236	0,0287	0,0353	0,0783	0,0270	0,0366	0,0378	0,0749	0,0426	0,0384	0,0357
	CHS	0,0658	0,0901	0,0723	0,0782	0,0660	0,0781	0,0804	0,0943	0,0606	0,0576	0,0857	0,1370	0,0768	0,0781	0,0653	0,1063	0,0837	0,0769	0,0851
	CDX	0,1128	0,1362	0,1256	0,1274	0,1171	0,1432	0,1363	0,1479	0,1097	0,0999	0,1334	0,1915	0,1227	0,1239	0,1060	0,1492	0,1199	0,1223	0,1386
	KHV	0,0308	0,0596	0,0367	0,0394	0,0314	0,0422	0,0413	0,0682	0,0308	0,0283	0,0415	0,0902	0,0368	0,0522	0,0474	0,0791	0,0539	0,0526	0,0506
	JPT	0,0954	0,1025	0,0748	0,0901	0,0822	0,0642	0,0808	0,1137	0,0832	0,1064	0,0967	0,1392	0,0881	0,1171	0,1272	0,1685	0,1357	0,1158	0,1004

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Distância Genética – Valores de F_{ST} (parte 2)

[illegible]

Teste de Diferenciação Populacional (parte 1)

		Americanas									Europeias										
		BSC	BRS	BPR	BSP	BPE	CLM	MXL	PEL	PUR	FRM	FRP	POR	CYP	ITA	CEU	FIN	GBR	IBS	TSI	
Americanas	BSC	*																			
	BRS	0,0000	*																		
	BPR	0,0985	0,0000	*																	
	BSP	0,6593	0,0000	0,0912	*																
	BPE	0,0140	0,0000	0,2399	0,0040	*															
	CLM	0,0000	0,0000	0,0000	0,0000	0,0000	*														
	MXL	0,0046	0,0000	0,0000	0,0326	0,0000	0,1878	*													
	PEL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0185	*												
	PUR	0,0000	0,0000	0,0000	0,0000	0,0000	0,1208	0,4845	0,0000	*											
Europeias	FRM	0,1796	0,0000	0,0000	0,0418	0,0000	0,0000	0,0000	0,0000	0,0000	*										
	FRP	0,7446	0,0000	0,0215	0,6007	0,0000	0,0000	0,0004	0,0000	0,0000	0,0987	*									
	POR	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	*								
	CYP	0,3428	0,0000	0,0030	0,5510	0,0000	0,0000	0,0044	0,0000	0,0000	0,0025	0,1182	0,0000	*							
	ITA	0,0000	0,0000	0,0000	0,0012	0,0000	0,0000	0,0004	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000	*						
	CEU	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0017	0,0006	0,0000	0,0018	0,0000	0,0000	0,0000	0,4048	*					
	FIN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0115	*				
	GBR	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0021	0,0003	0,0002	0,0000	0,0000	0,0000	0,0000	0,0085	0,2953	0,0002	*			
	IBS	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0245	0,0475	0,0000	0,0000	0,0000	0,0000	0,0000	0,0013	0,0838	0,0000	0,3639	*		
	TSI	0,0049	0,0000	0,0000	0,0421	0,0000	0,0085	0,0661	0,0026	0,0009	0,0078	0,0044	0,0000	0,0110	0,2394	0,3084	0,0002	0,0120	0,0165	*	
Africanas	MAL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	YAN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	SER	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	GUI	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,5534	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	TOR	0,2369	0,0082	0,7158	0,4610	0,7698	0,0333	0,0934	0,0017	0,0259	0,0097	0,2415	0,0001	0,2248	0,0000	0,0000	0,0000	0,0000	0,0000	0,0282	
	ACB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
	ASW	0,0000	0,0000	0,0001	0,0004	0,0000	0,0177	0,1611	0,0023	0,0005	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0000	
	ESN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	

	GWD	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	MSL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	YRI	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	LWK	0,0006	0,0000	0,0000	0,0001	0,0000	0,0001	0,0011	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Sul-Asiáticas	MDE	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0029	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0146	0,0000	0,0000	0,0000	0,0000	0,0051
	BEB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	GIH	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	ITU	0,0000	0,0000	0,0000	0,0016	0,0000	0,0000	0,0008	0,0000	0,0000	0,0000	0,0000	0,0000	0,0095	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003
	PJL	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	STU	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Leste-Asiáticas	CHB	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CHS	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	CDX	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	KHV	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	JPT	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Valores significativos estão em negrito ($p < 0,05$) ou em negrito e itálico ($p < 0,01$).

Teste de Diferenciação Populacional (parte 2)

[illegible]